INTERACTIVE CONTROL ENGINEERING

Control in Petroleum Refining

Interactive Control Engineering: Principles and Applications

Edited by John D. Hedengren

Control technology

1 Process Dynamics, Modeling, and Identification
2 Practical PID Control
3 Feedforward, Cascade, and Ratio Control
4 Override, Split-Range, and Constraint Control
5 Delays, Slow Measurements, and State Estimation
6 Multivariable and Plantwide Control
7 Nonlinear Dynamics, Oscillations, and Stabilization
8 Batch, Run-to-Run, and Hybrid Control
9 Model Predictive Control and Economic Optimization

Industry applications

10 Control in Upstream and Midstream Oil and Gas
11 Control in Petroleum Refining
12 Control in Petrochemicals and Polymers
13 Control in Chemical Plants
14 Control in Semiconductor Manufacturing
15 Control in Mining and Mineral Processing
16 Control in Iron, Steel, and Metals Production
17 Control of Energy Conversion and Storage
18 Control of Nuclear and Fusion Systems
19 Control of Biological Systems and Human Performance
20 Motion Control, Robotics, and Autonomous Systems

INTERACTIVE CONTROL ENGINEERING
Principles and Applications

Volume 11

Control in Petroleum Refining

Throughput, quality and limits from crude unit to blending

John D. Hedengren
Editor

Advanced Process Solutions, LLC

Interactive Control Engineering: Principles and Applications
Volume 11: Control in Petroleum Refining
Edited by John D. Hedengren

First edition, 2026

Copyright © 2026 Advanced Process Solutions, LLC. All rights reserved.

Produced by Advanced Process Solutions, LLC.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without the prior written permission of the publisher, except for brief quotations in a review or scholarly work.

The simulators described in this book are teaching models. Their equations, parameters and results are simplified for instruction and are not validated for the design, operation or safety assessment of any real plant, vehicle or medical device. The editor and the publisher accept no liability for use of this material outside that purpose.

Names of products and companies are used for identification only and may be trademarks of their owners.

Every chapter has a companion simulator that runs in a web browser at https://apmonitor.com/apps/. This volume online, with chapter files and citation records: https://apmonitor.com/apps/books/v11-refining/

How to cite this volume. John D. Hedengren, Ed., *Control in Petroleum Refining*, vol. 11 of *Interactive Control Engineering: Principles and Applications*. Advanced Process Solutions, LLC, 2026.

Typeset in Times and Helvetica-family type with LaTeX and the memoir class.

Preface

This book is one of 20 volumes of *Interactive Control Engineering: Principles and Applications*. Each chapter is about one plant, machine or system and one control lesson, and each has a simulator that runs in a web browser with nothing to install. The simulators are collected at https://apmonitor.com/apps/. Nine volumes are organized by control technology and eleven by industry or application area; the last pages of this volume list them all.

How to read this book with the apps

Chapter 1 teaches the theory this volume relies on and explains, once, the simulator framework that the apps share. Each chapter after it opens with the control question it answers, its learning objectives, what to read first, and the address of its app, also given as a code to scan. The chapter then develops the model, the controllers and the measured results. Every number reported was measured on the app, so a reader can open the app, load the named scenario and reproduce it.

Every chapter ends with exercises. Each is labeled with its kind and the time it should take, and none takes more than ten minutes. Concept checks and calculations can be done on paper. App challenges name a scenario, the settings and the run length. One exercise in each chapter, labeled *Review the agent*, presents a tuning, a model or a claim proposed by an AI assistant and asks what is wrong with it. Short answers to the odd-numbered exercises are at the back of the book, and worked solutions are available to instructors.

Who this book is for and three ways to read it

The book has three readers in mind. A student of engineering learns the method and works the problems. An engineer or operator in industry wants to know whether a result applies to the unit at hand, with which hardware, and what goes wrong in service. A curious reader wants to see what the system is, why it is hard to control and how people solved it, without the mathematics. Each chapter is built in layers so that all three can use it, and its first page names a path for each.

Explore	The picture and the plain words on the first page, the shaded history and fact segments, the first task in the app, and the key points. About a quarter of an hour, with no equations.
Study	The whole chapter in order: objectives, model, controller design, results, the pause points that send the reader to an exercise, and the exercises at the end.
Practice	The first page, the results, the section *Technology in Practice*, the limitations and the open challenges. About twenty minutes.

Every chapter opens with a picture of the system. Numbered badges mark the hardware on it: a green square is something the controller moves, a blue circle is something the controller reads, and an amber triangle is something that disturbs the system. The picture is an illustration and is not to scale. The schematic later in the chapter is the exact statement of what the model contains.

Shaded segments carry material that can be read on its own. Each kind has its own label and mark.

In plain terms	The main idea and the challenge of the chapter in everyday words.
History	Where the plant, the instrument or the control idea came from.
Did you know?	One number or event worth remembering, with its source.
Technology	The real sensor or final element behind the simulated one.
In practice	What people who run such a loop know that the model does not say.
Open challenge	What the field has not solved, and why it is hard.
Try it in the app	A task in the simulator at the point where the text describes it.
Key points	What to remember, at the end of the chapter.

The illustrations were generated by an AI image model from a written brief for each chapter and were checked against the chapter's model. The labels on them are set by the book.

How AI agents were used

AI agents did much of the development work behind this series. Under the editor's direction they wrote and tested the simulators, drafted the papers from which the chapters are drawn, measured the reported results on the running apps, converted the papers to book form, and drafted the introductory and closing chapters and the exercises. Other agents then checked that work: they re-measured reported numbers, audited what each app displays against the state of its model, reviewed the papers as referees, and solved the exercises without seeing the answers. The editor set the scope, chose the plants and the lessons, reviewed the results and is responsible for the content.

That way of working is also a subject of the book. The last chapter of each volume describes what the agents got wrong while this volume was made and how the errors were caught, and it sets out what the editor believes this means for engineers: more of the routine work can be delegated, and more of the engineer's value lies in checking, supervising, choosing the problem and knowing the process.

About this volume

This volume follows crude oil through a refinery, from the desalter to the gasoline blender. Chapters 2 to 5 cover crude preparation and distillation. Chapters 6 to 11 cover the conversion and upgrading units, where reactor heat balances and catalyst limits decide the control structure. Chapters 12 to 15 cover hydrogen, sulfur recovery, the fuel gas system and product blending.

A refinery is a chain of heat-integrated units, each pushed against a different limit, with product quality that is analyzed slowly or inferred. Each chapter names the technology volume that teaches its main technique.

John D. Hedengren
2026

Contents

Preface v

Notation xi

1 Introduction: Throughput, Quality and Limits in a Refinery 1
- 1.1 The Question This Volume Answers . . . 1
- 1.2 A Short History . . . 2
- 1.3 Notation and Conventions . . . 4
- 1.4 From Crude Tank to Product Tank on One Train . . . 4
- 1.5 Instruments and Final Elements in This Volume . . . 11
- 1.6 The Simulator Framework . . . 14
- 1.7 Working With AI Agents While Studying . . . 18
- 1.8 Chapter Roadmap . . . 18
- 1.9 Exercises . . . 21

2 Crude Desalter Simulator With a Mixing Optimum and a Misread Rag Layer 23
- 2.1 Introduction . . . 24
- 2.2 Water, Salt and the Electrical Treater . . . 25
- 2.3 The Desalter as Modeled . . . 26
- 2.4 An Integrator, an Optimum and a Drifting Reading . . . 32
- 2.5 The Control System as Implemented . . . 35
- 2.6 Ten Cases, Measured . . . 37
- 2.7 Technology in Practice . . . 42
- 2.8 Limitations . . . 44
- 2.9 Conclusion . . . 45
- 2.10 Exercises . . . 47

3 Fouling Crude Preheat Train Simulator for Override Control on Throughput 50
- 3.1 Introduction . . . 51
- 3.2 Background . . . 52
- 3.3 Process Model . . . 54
- 3.4 Controller Design . . . 59
- 3.5 Results . . . 64
- 3.6 Technology in Practice . . . 71
- 3.7 Limitations . . . 73
- 3.8 Discussion . . . 74
- 3.9 Conclusion . . . 75
- 3.10 Exercises . . . 77

4 A Crude Atmospheric Tower Simulator for Degrees of Freedom and Triangular Interaction 81
- 4.1 Introduction . . . 82
- 4.2 Background . . . 83
- 4.3 Process Model . . . 84
- 4.4 Controller Design . . . 87
- 4.5 Results . . . 90
- 4.6 Technology in Practice . . . 96
- 4.7 Limitations . . . 98

4.8 Discussion . . . 98
4.9 Conclusion . . . 99
4.10 Exercises . . . 101

5 Vacuum Tower Cut Points Held by an Ejector Curve That Breaks 104
5.1 Introduction . . . 105
5.2 Vacuum, Ejectors and the Heavy Gas Oil . . . 106
5.3 The Unit as Modeled . . . 107
5.4 Loops and Maximizer as Implemented . . . 113
5.5 Nine Cases, Measured . . . 114
5.6 Technology in Practice . . . 119
5.7 Limitations . . . 121
5.8 Conclusion . . . 122
5.9 Exercises . . . 124

6 Fluid Catalytic Cracking Simulator for Internal Heat Balance Feedback and Selective Control 127
6.1 Introduction . . . 128
6.2 How the Cat Cracker Came To Be Controlled . . . 129
6.3 Process Model . . . 130
6.4 What the Heat Balance Does . . . 140
6.5 Controller Design . . . 142
6.6 Results . . . 147
6.7 Technology in Practice . . . 152
6.8 Limitations . . . 154
6.9 Discussion . . . 155
6.10 Conclusion . . . 155
6.11 Exercises . . . 157

7 A Hydrocracker Reactor Simulator for Open-Loop Instability and Layered Protection 161
7.1 Introduction . . . 162
7.2 Background . . . 163
7.3 Process Model . . . 164
7.4 Dynamic Analysis . . . 167
7.5 Controller Design . . . 168
7.6 Results . . . 171
7.7 Technology in Practice . . . 177
7.8 Limitations . . . 179
7.9 Discussion . . . 179
7.10 Conclusion . . . 180
7.11 Exercises . . . 181

8 A Diesel Hydrotreater With Two Inlet Handles and a Half-Hour Sulfur Analyzer 184
8.1 Introduction . . . 185
8.2 From Cobalt Molybdate to Ten Parts per Million . . . 186
8.3 The Unit as Simulated . . . 187
8.4 Three Behaviors Before Any Feedback . . . 192
8.5 The Control Layers in the Code . . . 194
8.6 Nine Cases Measured . . . 196
8.7 Technology in Practice . . . 200
8.8 Limitations . . . 202
8.9 Conclusion . . . 203

8.10 Exercises . . . 205

9 Interheater Coordination and Inferred Octane on a Continuous Catalytic Reformer 208
9.1 Introduction . . . 209
9.2 Platinum, Rhenium and the Moving Bed . . . 210
9.3 Four Beds, Four Heaters and a Moving Catalyst . . . 211
9.4 What the Plant Does Before Any Controller . . . 217
9.5 Controllers, Estimator and Optimizer as Coded . . . 219
9.6 What the Nine Cases Show . . . 222
9.7 Technology in Practice . . . 225
9.8 Limitations . . . 227
9.9 Conclusion . . . 228
9.10 Exercises . . . 230

10 Three Clocks Behind One Octane in a Sulfuric Acid Alkylation Unit 233
10.1 Introduction . . . 234
10.2 Isoparaffin Alkylation, Acid Strength and Octane . . . 236
10.3 The Unit as Modeled . . . 236
10.4 Three Clocks, a Floor and a Trip . . . 242
10.5 Loops, a Soft Sensor and a Supervisor . . . 244
10.6 Nine Cases and a Day . . . 248
10.7 Technology in Practice . . . 254
10.8 Limitations . . . 256
10.9 Conclusion . . . 256
10.10 Exercises . . . 258

11 A Delayed Coker Whose Drum Switch Is a Disturbance Known in Advance 262
11.1 Introduction . . . 263
11.2 Delayed Coking and Its Drum Cycle . . . 264
11.3 The Unit as Modeled . . . 265
11.4 The Control Layers as Implemented . . . 270
11.5 Nine Cases, Measured . . . 272
11.6 Technology in Practice . . . 276
11.7 Limitations . . . 278
11.8 Conclusion . . . 279
11.9 Exercises . . . 281

12 A Hydrogen Plant Whose Steam Must Lead Its Feed and Whose Waste Is Its Fuel 284
12.1 Introduction . . . 285
12.2 Hydrogen From Natural Gas, and the Adsorber Behind It . . . 286
12.3 The Plant as Simulated . . . 287
12.4 The Control Layers . . . 293
12.5 Nine Cases, Measured . . . 295
12.6 Technology in Practice . . . 300
12.7 Limitations . . . 302
12.8 Implementation History . . . 302
12.9 Conclusion . . . 303
12.10 Exercises . . . 304

13 A Claus Sulfur Recovery Simulator for Feedforward, Dead Time and an Optimum Turned Into a Zero 308
13.1 Introduction . . . 309
13.2 Background . . . 310

13.3 Process Model . . . 311
13.4 Controller Design . . . 313
13.5 Results . . . 315
13.6 Technology in Practice . . . 321
13.7 Limitations . . . 323
13.8 Discussion . . . 323
13.9 Conclusion . . . 324
13.10 Exercises . . . 325

14 A Refinery Fuel Gas Header That Should Never Import and Flare at Once 329
14.1 Introduction . . . 330
14.2 Fuel Gas, Flares and the Wobbe Index . . . 331
14.3 The Fuel Gas System as Modeled . . . 332
14.4 Controllers as Implemented . . . 338
14.5 Nine Cases, Measured . . . 340
14.6 Technology in Practice . . . 345
14.7 Limitations . . . 347
14.8 Conclusion . . . 348
14.9 Exercises . . . 349

15 Gasoline Blending Simulator Whose Octane Authority Shrinks as the Tank Fills 353
15.1 Introduction . . . 354
15.2 Knock Ratings, Blending Indices and Blend Optimizers . . . 355
15.3 The Blender as Modeled . . . 356
15.4 Why the Specification Does Not Add Up . . . 361
15.5 Implemented Controller . . . 362
15.6 Scenarios and Measured Results . . . 366
15.7 Technology in Practice . . . 371
15.8 Limitations . . . 373
15.9 Conclusion . . . 374
15.10 Exercises . . . 375

16 Summary and Outlook 378
16.1 What the Chapters Showed . . . 378
16.2 Choosing a Method . . . 382
16.3 Where Practice Is Heading . . . 382
16.4 Engineering With AI Agents . . . 384
16.5 Gaps in Refinery Control . . . 387
16.6 Exercises . . . 391

Answers to Selected Exercises 392

Glossary 397

Index of Control Topics 400

Index of Applications 402

About the Editor 405

Volumes in This Series 406

Notation

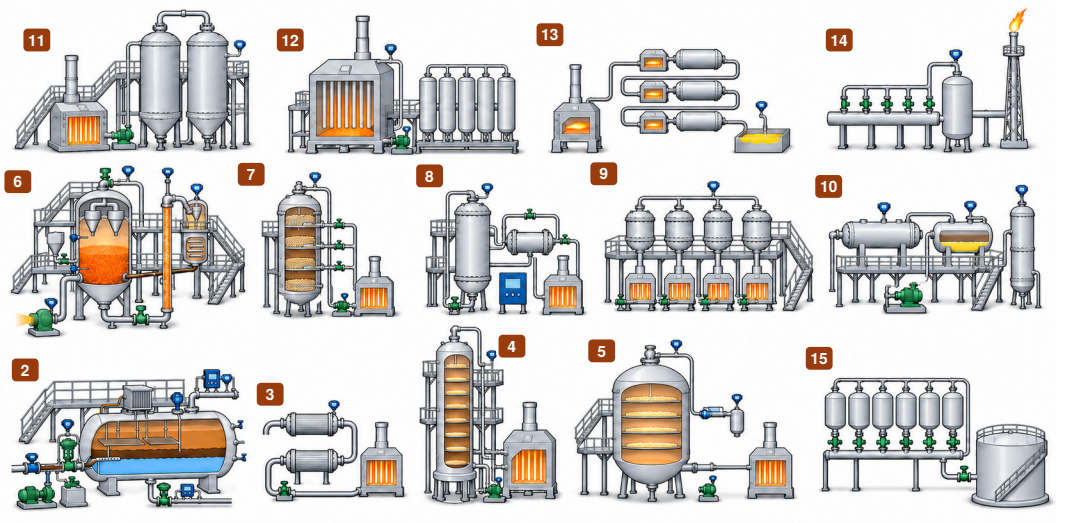
The chapters share the symbols below. A symbol used by one chapter only is defined where it first appears, and each chapter states its own units.

Symbol	Meaning
SP	setpoint, the value a controller is asked to hold
PV	process variable, the measurement a controller acts on
CV	controlled variable
MV	manipulated variable; also OP, the controller output
DV	disturbance variable
e	control error, SP $-$ PV
K_c	controller gain
τ_I, τ_D	integral (reset) time and derivative time
K_p	process gain, the steady-state change in PV per unit change in MV
τ_p	process time constant
θ_p	process dead time
τ_c	closed-loop time constant chosen in a tuning rule
Δt	controller sample time
s	Laplace variable
x, u, y, d	state, input, output and disturbance of a state-space model
A, B, C	matrices of the linear model $\dot{x} = Ax + Bu$, $y = Cx$
Q, R	weights on outputs (or states) and on input moves in an optimal controller
$\hat{x}$	an estimate of x

Abbreviations used throughout: PID, proportional-integral-derivative; PI, proportional-integral; FOPDT, first-order plus dead time; MPC, model predictive control; MHE, moving horizon estimation; IAE, integral of the absolute error; RGA, relative gain array. Times are in seconds unless a chapter says otherwise. Controller output and valve position are in percent of span. A scenario or case number refers to the numbered list in the chapter's app.

CHAPTER 1

Introduction: Throughput, Quality and Limits in a Refinery



The chapters 2 desalter 3 preheat train 4 atmospheric tower 5 vacuum tower 6 cat cracker 7 hydrocracker 8 diesel hydrotreater 9 catalytic reformer 10 alkylation unit 11 delayed coker 12 hydrogen plant 13 sulfur recovery 14 fuel gas header 15 gasoline blender

Notice that the crude is washed, heated and distilled along the front row, where the gasoline is also blended, that the middle row converts and treats the cuts, and that the rear row holds the coker and the three plants that serve the others. Illustration, not to scale.

IN PLAIN TERMS **Running a refinery against its limits**

The main idea. A refinery heats crude oil, splits it into parts by how easily they boil, changes some parts chemically and blends the results into fuels. Each unit earns more the more oil it takes, so it is pushed until a product or a machine says stop.

The challenge. The limit that stops a unit moves with the oil, the weather and the wear of the equipment. The quality of a product is often known only minutes or hours after the oil has passed.

Why it matters. A unit held short of its limits wastes money, and one pushed past them makes fuel that cannot be sold.

THREE WAYS TO READ THIS CHAPTER

Explore	7 min	This page, Section 1.2 with its timeline and two History boxes, Figure 1.2 and the Did you know? boxes in Sections 1.4.1, 1.4.4 and 1.4.7, and Key points in Section 1.8. No equations.
Study	45 min	The whole chapter in order, working each number of the example R-1, then Exercises 1 to 10.
Practice	10 min	This page, Section 1.4.7 with Table 1.2, Section 1.5 with Figures 1.6 and 1.10, and Section 1.8.

1.1 The Question This Volume Answers

A refinery turns crude oil into fuels by heating it, cutting it into fractions, reacting the fractions over catalysts and blending what comes out. Each unit earns more the more it processes, each product is sold against a specification, and each piece of equipment has a limit that ends the run if it is crossed. This volume answers one question: how does each unit hold its products on specification and its

equipment inside its limits while it processes as much as the market will take, and which control techniques recur from the crude tank to the gasoline tank?

Six control problems define the industry, and each chapter meets at least two of them.

- *Heat is bought with fuel and recovered from products.* Fired heaters supply the heat the units need, exchangers recover what the products carry away, and fouling moves the load from the exchangers to the heaters.
- *Products are cut from a boiling curve.* A tower makes several products from one feed, so it has more specifications than handles, and half of the specifications follow from the other half.
- *Reactors feed back on their own heat.* An exothermic bed amplifies a change in its inlet, a cracking unit burns the coke it made a minute earlier, and an exchanger around a reactor returns part of every degree added.
- *Quality is known late or inferred.* Product analyzers report minutes to half an hour after the oil they describe, laboratories report a few times a day, and much of what the controllers read is computed from temperatures, pressures and densities.
- *Throughput is pushed against limits that move.* Firing, tube metal, compressors, blowers and catalyst all bind at some throughput, and which binds first changes with the crude, the weather and the season.
- *Units are coupled through shared streams.* Off-gases fuel the heaters, hydrogen passes from the units that make it to the units that consume it, acid gas goes to sulfur recovery, and the gasoline sold is the average of everything blended into a tank.

Fourteen chapters answer the question on fourteen plants, each with an app that the reader runs in a browser. Chapters 2 to 5 follow the crude through washing, preheating and atmospheric and vacuum distillation. Chapters 6 to 11 are the conversion and treating units: a catalytic cracker, a hydrocracker, a diesel hydrotreater, a catalytic reformer, an alkylation unit and a delayed coker. Chapters 12 to 15 are the plants that serve the others and finish the products: hydrogen, sulfur recovery, the fuel gas header and gasoline blending.

A reader who has worked through the volume can do the following on a unit that is new to the reader.

- Write a fired heater's heat balance, turn it into a feedforward and say how much throughput a fouled preheat train costs once the firing is at its limit.
- Count a tower's degrees of freedom, place its cut points on the assay and say which draw moves which cut.
- Compute the runaway index of an adiabatic bed and the loop gain of a feed and effluent exchanger, and say how close the unit sits to losing its operating point.
- Correct a temperature for pressure, update an inferred property from a laboratory result and say what an analyzer's delay costs.
- Place each controller in its layer, from regulatory loops to the optimizer, and name the limit that each override or selector protects.

This chapter develops each of these on one small example and names the chapter that uses it. PID forms and tuning are taken as known from Volume 2, cascade and feedforward from Volume 3, selectors and overrides from Volume 4, dead time and estimation from Volume 5, and predictive control and optimization from Volume 9.

1.2 A Short History

The first refineries made one product. A plant of the early 1860s distilled crude oil at atmospheric pressure to make kerosene for lamps, and naphtha and tar were its by-products [1]. The electric light then cut the demand for kerosene, the internal combustion engine created one for gasoline and diesel, and distillation alone gave only a certain amount of gasoline from each barrel [1]. Each later process

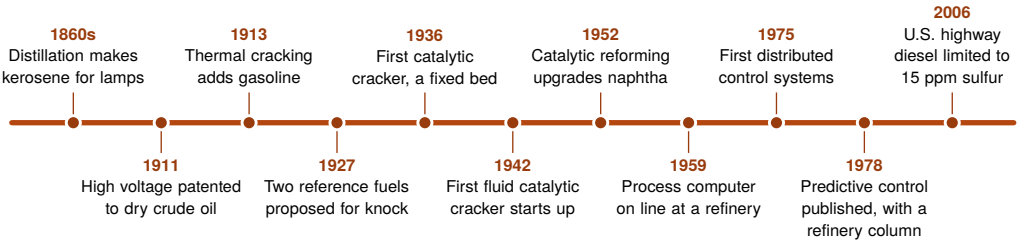


Figure 1.1 Milestones of refining and of its control. The process dates are from [1] (the 1860s, 1913 and 1952), [2] (1911), [3] (1927), [4] (1936 and 1942) and [5] (2006). The control dates are from [6] (1959 and 1975) and [7] (1978). Notice that the three control milestones fall within twenty years, about a century after the first unit. Events are evenly spaced, not to a time scale.

on the timeline of Figure 1.1 answers one of three pressures: more gasoline from a barrel, a better gasoline, or a cleaner product.

Thermal cracking, introduced in 1913, broke heavy fractions into gasoline under heat and pressure [1, 4]. In 1927 two pure hydrocarbons were proposed as reference fuels for knock, and the octane tests still rest on mixtures of those two [3]. Octane is the number that Chapter 9 holds and Chapter 15 blends to. Catalytic cracking arrived in the 1930s, alkylation in 1940 and catalytic reforming in 1952, and each raised the octane or the yield of gasoline [1]. Hydrodesulfurization followed in 1954 and hydrocracking in 1960 [1], the two processes that the plant of Chapter 12 supplies. A rule on highway diesel in the United States then cut the sulfur limit from 500 to 15 parts per million for refiners from June 1, 2006 [5]. Chapter 8 works under that limit with an analyzer half an hour late.

> ⧖ HISTORY **1913 to 1942: Three Ways to Crack a Heavy Oil**
>
> Thermal cracking, developed around 1913, heated distillate fuels and heavy oils under pressure in large drums until they cracked into smaller molecules, and it left large amounts of solid, unwanted coke [1]. Its gasoline was of relatively poor quality [4]. The first catalytic cracker started up in 1936 with its catalyst in a fixed bed and processed 15 000 barrels a day. The first commercial fluid unit followed in mid-1942, with a powdered catalyst that a gas flow lifts until it moves like a liquid between a reactor and a regenerator, and 34 such units ran in the United States at the end of the war [4]. Thermal cracking did not disappear. It evolved into visbreaking, steam cracking and coking [1]. The fluid unit is the plant of Chapter 6, and coking that of Chapter 11.

Control followed the processes at a distance. The first process computer went fully on line on a refinery unit in March 1959, and the first distributed control systems were introduced in 1975 [6]. Predictive control was first described at a conference in 1976 and published in 1978 [7], with a distillation column of a catalytic cracker among its three applications [8]. Figure 1.10 sets today's layers on their platforms.

> ⧖ HISTORY **1959: A Computer Goes on Line at a Refinery**
>
> The first computer sold for process use went fully on line in March 1959, on a catalytic polymerization unit of a Texas refinery [6]. It had a memory of 8K, and it raised the conversion of propylene to gasoline from 80 to 91 %. The first process computers were supervisory. They adjusted the setpoints of pneumatic or electric controllers, so the plant kept working when the computer was off. This one closed the loop on product quality every hour and adjusted 19 parameters. The first distributed control system, introduced in 1975, also adjusted independent controllers [6]. This volume still uses that arrangement, a slow layer that moves the setpoints of fast loops (Section 1.4.7).

1.3 Notation and Conventions

Temperatures are in °C, and differences of temperature in kelvin. Mass flows are in t/h, duties in MW, and a stream's heat capacity flow $C = \dot{m}\,c_p$ in MW/K. The coil outlet temperature of a heater is COT, the crude temperature leaving the preheat train T_{pre}, and an exchanger's effectiveness ε. A cut point is the temperature on the crude's true boiling point (TBP) curve at which one product ends and the next begins, and v is the cumulative fraction of the crude distilled. In a reactor, x is the conversion of the reacting lump, ΔT_{ad} the adiabatic rise at complete conversion, r the runaway index and L the gain of a loop of positive feedback. The framework's h is its integration step and its a the controller action, so neither is used here for anything else.

The chapters keep their own symbols, and several collide. Chapter 7 writes the adiabatic rise as K, and Chapter 14 uses K for the integrating gain of its header. Chapter 4 writes the crude's heaviness as h and the internal reflux ratio as R_i, while Chapter 13 uses R for the air ratio and Chapter 2 h for layer heights. Chapter 6, Chapter 5 and Chapter 8 write the controller action as σ. The weighted average bed temperature of Chapter 8 is WABT and the weighted average inlet temperature of Chapter 9 is WAIT. Octane is RON or the anti-knock index AKI, and vapor pressure is Reid vapor pressure (RVP) in psi, the units of the trade in Chapter 15.

1.4 From Crude Tank to Product Tank on One Train

1.4.1 The Refinery Flow Scheme

Figure 1.2 draws a refinery and marks where each chapter's plant sits in it [9]. Crude is washed free of salt, preheated against the hot products and heated in a furnace, and the atmospheric tower cuts it into naphtha, kerosene, diesel, gas oil and a residue. The residue is distilled again under vacuum. Its gas oil feeds the cracking units, and the vacuum residue goes to the coker. Naphtha is reformed into a high-octane component, diesel is hydrotreated to remove sulfur, and the light olefins of the cracker are alkylated. Gasoline is blended from several of these streams. Three plants serve all the others: the hydrogen plant supplies the hydroprocessing units, sulfur recovery takes the hydrogen sulfide that treating removes, and the fuel gas header carries the units' off-gases to every fired heater.

> **DID YOU KNOW? Where the Barrels Go**
>
> On January 1, 2026 the 130 operable refineries of the United States could charge 19.2 million barrels of crude oil per stream day. Their hydrotreaters could take 17.0 million, their vacuum towers 8.4 million, their catalytic crackers 5.3 million, their reformers 3.5 million and their hydrocrackers 2.3 million [10].

The chapters take the plants one at a time, and the techniques they share are developed below on one running example, a small train called R-1. A crude unit charges 720 t/h of crude through a preheat train and a fired heater to an atmospheric tower. The tower's diesel cut feeds a hydroprocessing reactor with a feed and effluent exchanger, a charge heater and one adiabatic bed. The example ends at a gasoline tank. Every number of R-1 is declared here or computed from these declarations by the script kept with this chapter's record. The numbers are chosen for clear arithmetic and are not taken from any one chapter.

1.4.2 Fired Heaters

A fired heater raises a process stream from the temperature at which it arrives to its coil outlet temperature, and the heat it must absorb follows from a balance on the stream, Equation (1.1).

$$Q = C\,(\mathrm{COT} - T_{pre}) + Q_{vap}, \qquad C = \dot{m}\,c_p \tag{1.1}$$

where Q_{vap} is the heat that vaporizes part of the stream in the coil. R-1's crude has $c_p = 2.5\,\mathrm{kJ/(kg\,K)}$, so 720 t/h is $C = 0.5\,\mathrm{MW/K}$. The example leaves out Q_{vap}, which Chapter 4 and Chapter 5 keep.

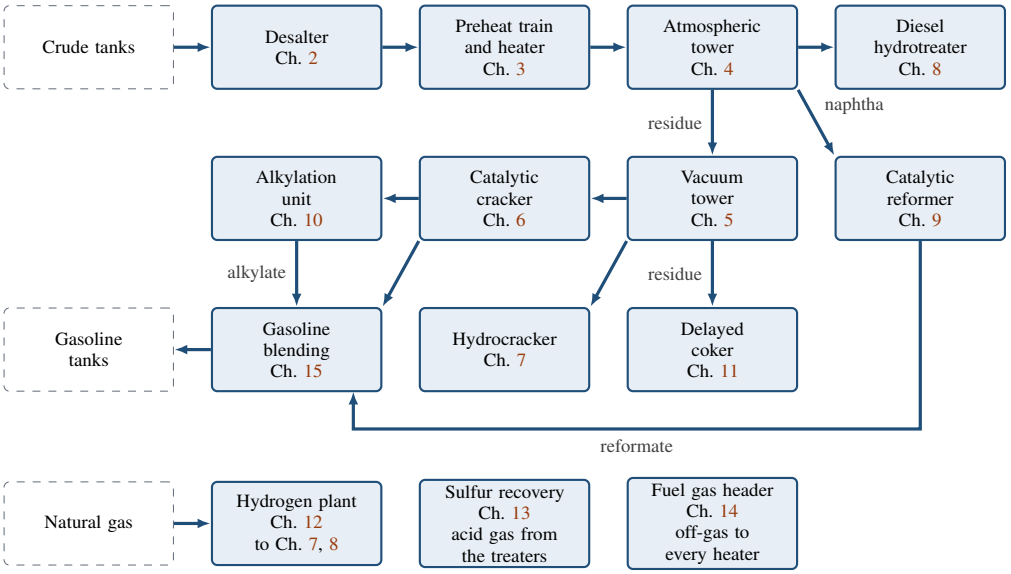


Figure 1.2 A Refinery Flow Scheme. Filled Blocks Are the Plants of Chapters 2 to 15, Dashed Blocks Lie Outside the Volume. Vacuum Gas Oil Feeds the Cracker and the Hydrocracker, and the Cracker's Light Olefins the Alkylation Unit. The Three Plants in the Bottom Row Serve All the Others. Kerosene, Light Ends and Most Product Pools Are Omitted.

The crude arrives at 265 °C and leaves at a COT of 365 °C, so the heater absorbs 50 MW. Its burners can deliver 60 MW, and the heater runs at 83.3 % of full firing.

The balance is also the most useful feedforward in a refinery. A heater whose firing follows Equation (1.2) meets a change in rate or in preheat before the outlet temperature moves.

$$Q_{ff} = C_m \left(\mathrm{COT}_{SP} - T_{pre,m}\right) \tag{1.2}$$

where the subscript m marks a measured value. On R-1 each kelvin of preheat lost needs 0.5 MW more, and a tenth more crude needs 5 MW more at the same preheat. The outlet loop then trims only what the balance does not know. Chapter 3 adds this feedforward to its outlet loop, Chapter 12 subtracts the heat its tail gas brings, and Chapter 11 and Chapter 9 drive their heaters from the duty their balances predict.

The fuel is the balance's least certain term. Refinery fuel gas is a mixture of off-gases whose heating value changes with what the units send. A heater whose fuel valve is held at one flow loses heat in proportion to the heating value. R-1's heater adds 100 K, so each percent of heat lost is a kelvin of COT, and a gas 5 % leaner takes 5 K off the outlet before the loop acts. A burner held at a fixed pressure instead passes heat in proportion to the Wobbe index, the heating value divided by the square root of the specific gravity. Hydrogen lowers both, so 10 % hydrogen in the gas of Chapter 14 lowers the heating value 7.11 % but the Wobbe index only 2.66 %. On R-1 that is 7.1 K against 2.7 K at the coil outlet. Chapter 14 feeds the Wobbe index forward to its burners, and Chapter 3 meets an unmeasured change in heating value through its outlet loop alone.

A heater has two limits. The firing is limited by its burners and its fuel valve. The tubes are limited by their metal temperature, which runs above the fluid by an amount that grows with the heat flux, as shown in Equation (1.3).

$$T_{metal} = T_{fluid} + \theta_m \frac{Q}{Q_{max}} + \Delta T_{coke} \tag{1.3}$$

where θ_m is the metal's rise at full firing and ΔT_{coke} grows as coke lays down inside the tubes. The tube metal is computed, or read from a few thermocouples welded to the tubes, and it binds before the

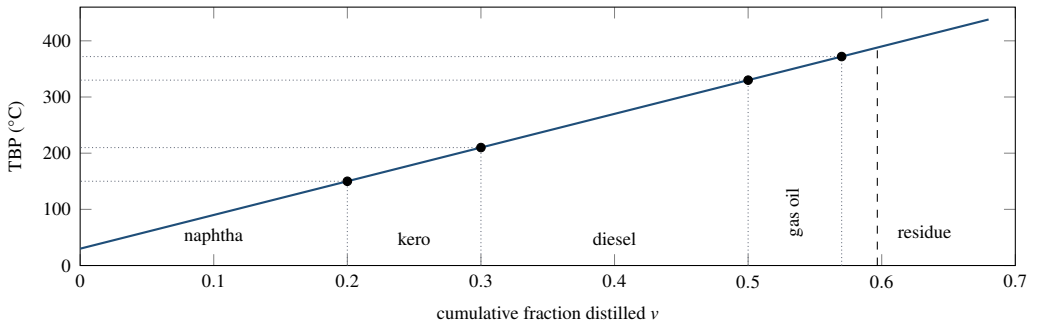


Figure 1.3 R-1's Assay (Thick Line) and Its Four Cuts (Dots), Placed by the Material Balance of Equation (1.4). The Dashed Line Is the Fraction Vaporized at the Flash Zone, and the Gap Between It and the Last Cut Is the Overflash.

firing does once the tubes are coked. Chapter 9 and Chapter 12 write the metal in this form, Chapter 5 and Chapter 11 add a coke layer that grows with temperature, and Chapter 12 protects its tubes with an override on the outlet loop.

1.4.3 Distillation: Cut Points and Degrees of Freedom

A crude is described by its true boiling point curve, the temperature at which each cumulative fraction of it boils [11]. A tower that makes products in series from one feed cuts that curve, and to a first approximation the material balance alone places each cut. The cut below draw i lies at the share of the crude taken by that draw and every draw above it, as shown in Equation (1.4).

$$v_i = \frac{1}{F}\sum_{j\le i} D_j, \qquad T_i = T_{bp}(v_i), \qquad \frac{\partial T_i}{\partial D_j} = \begin{cases} T'_{bp}(v_i)/F, & j \le i \\ 0, & j > i \end{cases} \tag{1.4}$$

where F is the charge, D_j the product drawn at position j counted from the top and T'_{bp} the slope of the curve. R-1's assay is a straight line, $T_{bp} = 30 + 600\,v$, which no crude is, but its slope is in the range of a real curve's middle. Naphtha, kerosene, diesel and gas oil are drawn at 144, 72, 144 and 50.4 t/h, so the cuts fall at 150, 210, 330 and 372 °C, as shown in Figure 1.3. Each tonne an hour of a draw moves every cut below it by 0.83 K and leaves every cut above it alone.

The structure of Equation (1.4) is the lesson. Each product has two ends, and neighbors share a cut, so four cuts fix the eight end specifications of four products. Four of the eight are choices and four are consequences, and which four is an operating decision. The kerosene's front end and the naphtha's back end are the same cut, so a loop on each fights over one variable. The interaction is also one way. A draw moves its own cut and every cut below it, so the gain matrix is lower triangular. To lift the kerosene's end point by 12 K without moving the diesel's, the kerosene draw rises 14.4 t/h and the diesel draw falls by as much.

The tower's heat sets what the material balance does not. Internal reflux sharpens each cut, and a draw that takes liquid off a tray thins the reflux below it. Chapter 4 measures that second effect at about 4 % of the first. The heater sets how much of the crude vaporizes at the flash zone. R-1's flash zone, 2 K below the COT and read 25 K higher on the assay for the stripping steam, vaporizes 59.67 % of the crude, and the draws take 57 %. The difference, 2.67 %, is the overflash, the liquid that wets the wash zone above the flash zone. With a floor of 2 %, the gas oil draw can rise only 4.8 t/h. Each kelvin of COT vaporizes 1.2 t/h more and costs 0.5 MW, so the lowest draw is held by the heater as much as by its own valve.

Chapter 4 counts these degrees of freedom on a four-product tower, measures the triangular interaction and protects the overflash with an override. Chapter 5 cuts the residue by pressure as well as by heat, and holds its wash bed wet with an overflash override. Chapter 11 cuts a fractionator

whose feed swings with the drum cycle, and Chapter 10 sets the purity of its recycle with a column built on the column module shared across the series.

1.4.4 Reactor Heat Balances and Runaway

Most refinery reactors are adiabatic beds of catalyst. The heat of reaction stays in the stream, so the outlet temperature is the inlet plus the reaction's rise, and the rise depends on the temperature. A lumped bed that reacts at the mean of its inlet and outlet obeys Equation (1.5), the form Chapter 7 uses.

$$\tau_b \frac{\mathrm{d}T_{out}}{\mathrm{d}t} = T_{in} + \Delta T_{ad}\, x(\bar{T}) - T_{out}, \qquad \bar{T} = \tfrac{1}{2}(T_{in} + T_{out}) \tag{1.5}$$

where τ_b is the bed's thermal time constant and x the conversion, which rises with temperature. Linearizing about a steady state gives Equation (1.6), in which one group appears twice.

$$\tau_b \frac{\mathrm{d}\,\delta T_{out}}{\mathrm{d}t} = (r-1)\,\delta T_{out} + (1+r)\,\delta T_{in}, \qquad r = \frac{\Delta T_{ad}}{2}\frac{\mathrm{d}x}{\mathrm{d}\bar{T}} \tag{1.6}$$

The runaway index r decides two things. The bed is stable while $r < 1$, and while stable it amplifies its inlet by the steady gain of Equation (1.7) and slows its own response by $1/(1-r)$.

$$G = \frac{\partial T_{out}}{\partial T_{in}} = \frac{1+r}{1-r} \tag{1.7}$$

Studies of parametric sensitivity give the same warning for beds with axial profiles [12].

R-1's bed has $\Delta T_{ad} = 100\,\mathrm{K}$ and a conversion that grows by a factor e every 30 K of mean temperature, which doubles it every 20.8 K, up to complete conversion. At the design inlet of 320 °C it converts $x = 0.2$, rises 20 K to 340 °C and has $\mathrm{d}x/\mathrm{d}\bar{T} = x/30$, so $r = 50 \times 0.2/30 = 1/3$ and $G = 2$ exactly. A kelvin at the inlet is two at the outlet, and the bed's time constant grows from 3 to 4.5 min. Because r grows with the rise, the bed loses its cool operating point where the rise reaches 60 K. That happens at an inlet of 332.96 °C, only 12.96 K above design, as shown in Figure 1.5a. Above it the only steady state is complete conversion, 100 K above the inlet, and the bed stays there until the inlet is cut below 328.28 °C.

A cut in feed moves the bed toward the same edge. Conversion rises as the space velocity falls, and the heat released has less oil to warm. At the design inlet a tenth less feed lifts the rise to 23.6 K and G to 2.30, a fifth less to 29.1 K and 2.88, and a cut of 35.1 % leaves no cool state at all. Chapter 7 reports the same mechanism on its four beds, where the first sits at $r = 0.538$, amplifies its inlet 3.3 times and loses its cool state four degrees above design. Its quench between beds breaks the chain of amplification, and its protection must read a bed rise whose transient sign is wrong for a few minutes after every correction.

An endothermic bed has $r < 0$ and $G < 1$. A hotter bed reacts faster and cools itself more, so it absorbs most of any heat given to it. Chapter 9 measures this on its first bed, which passes 6.3 K of a 15.7 K rise in its inlet to the next bed, which is why a reformer needs a heater before each bed.

The same arithmetic governs any unit that feeds back its own heat. If a fraction L of any change returns to its own input, the steady change is amplified as in Equation (1.8), and a first-order element inside the loop slows by the same factor.

$$A = \frac{1}{1-L}, \qquad \tau_{eff} = \frac{\tau}{1-L} \tag{1.8}$$

The cat cracker of Chapter 6 has no furnace. Its catalyst is reheated by burning the coke the cracking laid on it, and its measured loop gain is 0.523, so a slide valve move arrives 2.09 times larger than it first looks and settles over about 90 min. The loop gain grows with severity, and once it reaches one the unit has more than one steady state [13, 14]. A classic analysis of autothermic processes sets the heat a reaction generates against the heat its stream removes in the same way [15], and Table 1.1 collects the loop gains of the volume.

Table 1.1 Heat Fed Back in the Example and the Chapters

Plant	Path of the feedback	Index or L	Gain
R-1 bed	reaction heat in the bed	$r = 1/3$	2.00
R-1 bed and exchanger	effluent heat returned to feed	$L = 0.60$	2.50
Chapter 7, bed 1	reaction heat in the bed	$r = 0.538$	3.3
Chapter 8	effluent heat returned to feed	$L = 0.728$	3.7
Chapter 6	coke burned to reheat catalyst	$L = 0.523$	2.09
Chapter 9, bed 1	endothermic reaction	$r < 0$	below 1

✦ DID YOU KNOW? Three Units That Multiply Their Own Disturbances

From the simulator (Table 1.1). The first bed of the hydrocracker passes on a change at its inlet 3.3 times larger. The exchanger loop of the diesel hydrotreater multiplies a change by 3.7, and the heat balance of the cat cracker by 2.09.

1.4.5 Heat Integration

A refinery recovers most of its heat. R-1's preheat train heats the crude from 40 °C against products that leave the tower at 340 °C. Because the products are the crude, their heat capacity flow equals the crude's, and the train behaves as one balanced counter-current exchanger with the effectiveness of Equation (1.9).

$$\varepsilon = \frac{N}{1+N}, \qquad N = \frac{UA}{C}, \qquad T_{pre} = T_{cold} + \varepsilon\,(T_{hot} - T_{cold}) \tag{1.9}$$

Clean, the train has $UA = 1.5\,\mathrm{MW/K}$, so $N = 3$, $\varepsilon = 0.75$ and $T_{pre} = 265\,^{\circ}\mathrm{C}$. It recovers 112.5 MW, which is 69.2 % of the heat the crude takes, and the heater supplies the rest.

Fouling lays a deposit on the exchanger walls and lowers UA over weeks [16, 17]. The outlet loop hides it. The COT stays on setpoint while the firing climbs, and only the position of the fuel valve records what happened. On R-1 the firing reaches 95 % when 21 % of the clean UA is lost and the preheat has fallen to 251 °C, as shown in Figure 1.4a. Beyond that point no tuning helps, and the only handle left is the charge. A valve position override takes the charge from the operator's demand when the firing passes its target, as shown in Equation (1.10).

$$F_{cmd} = \min\left(F_{dem},\ \mathrm{PI}(x_{v,SP} - x_v)\right) \tag{1.10}$$

where x_v is the firing and $x_{v,SP}$ its target, 95 % here. The override's upper output limit is the demand, so it rests there until the firing crosses its target and cannot wind up while it waits.

The override costs throughput, and less than a first estimate suggests. At 40 % of UA lost, the charge held at 720 t/h would need 66.1 MW and would deliver a COT of 352.9 °C with the burners wide open. The override cuts the charge to 653.6 t/h, a cut of 9.2 %, and holds the COT. An estimate at fixed preheat would allow only 621.1 t/h. The difference is the train itself: less crude through the same surface has a larger N, so the preheat rises from 232.9 °C to 239.4 °C and each tonne needs less firing. Chapter 3 measures this trade on a two-exchanger train, where holding temperature through one exchanger's cleaning costs 555 t of crude, and Chapter 4 ties its preheat to the three pumparound duties that recover the tower's heat.

Around a reactor, heat integration closes a loop. R-1's diesel enters its feed and effluent exchanger at 200 °C, and the exchanger with its bypass returns $\varepsilon_u = 0.3$ of the effluent's temperature above the feed. The charge heater adds the rest of the way to the bed, as shown in Equation (1.11).

$$T_{in} = T_f + \varepsilon_u\,(T_{out} - T_f) + \frac{Q_h}{C}, \qquad L = \varepsilon_u\,G \tag{1.11}$$

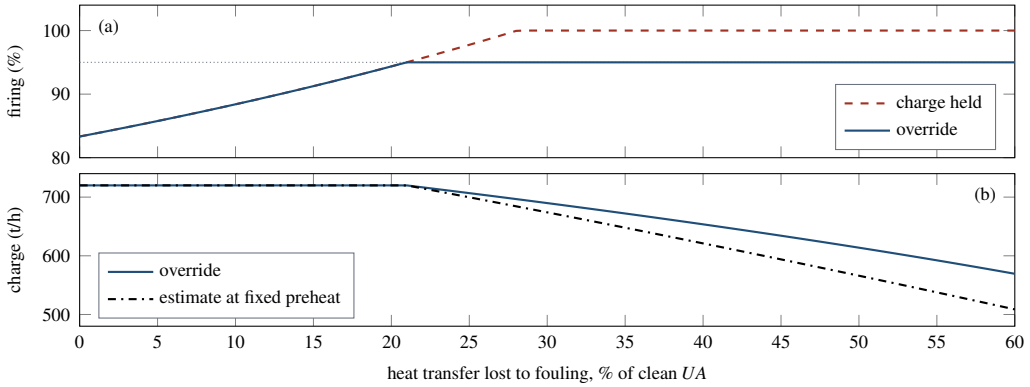


Figure 1.4 R-1's Crude Unit as Its Preheat Train Fouls. (a) Firing With the Charge Held at 720 t/h (Dashed), Which Reaches Its Limit and Then Loses COT, and With the Valve Position Override (Solid), Which Holds the 95 % Target (Dotted). (b) The Charge the Override Allows (Solid) and the Estimate That Ignores the Preheat Gained at Lower Charge (Dash-Dotted).

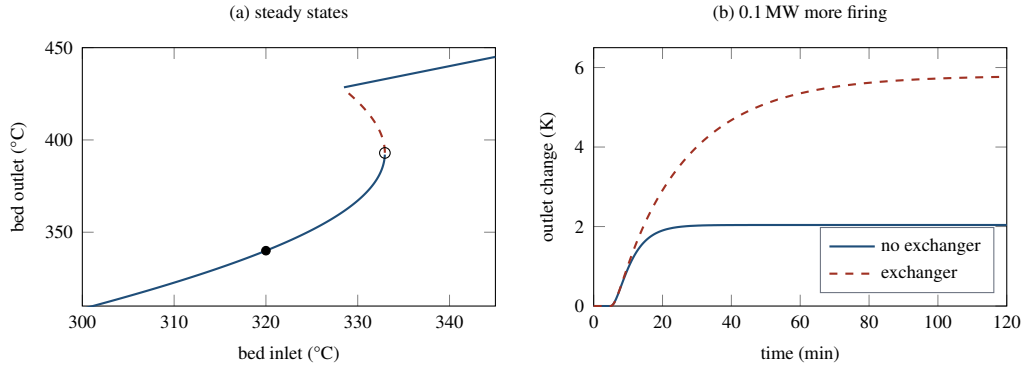


Figure 1.5 R-1's Bed. (a) Steady Outlet Against Inlet: Stable States (Solid) and Unstable States (Dashed). The Dot Is the Design Point and the Circle the Fold Where $r = 1$. (b) Outlet After a 0.1 MW Step in Firing at 5 min, With the Heater Feeding the Bed Directly (Solid) and Through the Exchanger Loop (Dashed).

At design the exchanger delivers 242 °C and the heater adds 78 K, 7.8 MW at $C = 0.1\,\mathrm{MW/K}$. A kelvin more at the inlet is two more at the outlet, and 0.3 of those two come back, so the loop gain is $L = 0.6$ and Equation (1.8) gives an amplification of 2.5. An exchanger that returned half the effluent's heat would bring L to one, the edge of open-loop instability [18]. The loop also brings the bed's own edge closer. With the exchanger in place the unit loses its cool operating point at an inlet of 328.23 °C, 8.23 K above design instead of 12.96 K, and only 0.21 MW more firing takes it there.

Figure 1.5b shows both effects of the loop. With heater and bed lags of 2 and 3 min and an exchanger lag of 3 min, 0.1 MW more firing raises the inlet 1 K and the outlet 2.04 K when the exchanger's return is held, and the outlet reaches 63 % of its change in 7.0 min. Through the loop the outlet settles 5.8 K higher, more than the linear 5.0 because r grows with the rise, and takes 21.5 min to reach 63 %. An operator or a loop tuned on the first ten minutes sees a plant less than half as strong as the one underneath.

Chapter 8 measures its exchanger loop at $L = 0.728$, rising to 0.899 at a cooler inlet, and moves the inlet with a fast bypass while mid-ranging returns the slow heater to the middle of its range. Its bed also shows wrong-way behavior, a cooler inlet that first warms the outlet [19]. Chapter 9 returns

its last bed's effluent to its first heater, and Chapter 12 burns the adsorber's tail gas in the furnace that makes the hydrogen, so what the adsorber recovers changes what the furnace burns.

1.4.6 Analyzers and Inferred Properties

A refinery sells its products on properties that no transmitter reads: a sulfur content, an octane number, a vapor pressure, a 95 percent point. Three sources supply them, and each has a cost. An on-line analyzer takes a sample, carries it down a line and reports after a cycle, so its reading is always old. A sample-and-hold report with transport delay θ and cycle T_s is between θ and $\theta + T_s$ old, and for tuning it acts as the effective delay of Equation (1.12).

$$\theta_{eff} \approx \theta + \tfrac{1}{2} T_s \tag{1.12}$$

R-1's sulfur analyzer, with a 20 min delay and a 10 min cycle, is 25 min late on average. A loop on it can be no faster than that delay allows, as Volume 5 develops. A laboratory is more accurate and far rarer, a few results a day.

An inferred property is computed now from what is measured now. The oldest inference is a tray temperature corrected for pressure, because a cut point is a boiling point and a boiling point moves with pressure. The atmospheric equivalent temperature of Chapter 5 gives the normal boiling point T_b of what boils at a temperature T and pressure p, as shown in Equation (1.13).

$$T_b = T \left(1 + \frac{1}{A} \ln \frac{p_{atm}}{p} \right) \tag{1.13}$$

with temperatures in kelvin and A a constant of the fraction, 16 on the residue of Chapter 5 and declared 12 for R-1's diesel. R-1's diesel draw tray at 2 bar boils at 366.2 °C for a 330 °C cut. A rise of 0.1 bar raises the tray 2.8 K with no change in the cut, so a loop holding the raw temperature would lighten the diesel cut by 2.6 K. Chapter 5 infers its heavy gas oil's 95 percent point this way, and Chapter 4 reads its product properties from inferred cut points.

A model of the property from several measurements is a soft sensor [20, 21]. It has no delay, but it drifts as the plant moves away from where the model was fitted. The laboratory corrects it with the bias update of Equation (1.14).

$$b_k = b_{k-1} + \lambda \left(y_{lab,k} - \hat{y}_{raw}(t_{s,k}) - b_{k-1} \right) \tag{1.14}$$

where $\hat{y}_{raw}(t_{s,k})$ is the model's value at the instant the sample was drawn, not when the result came back, and $0 < \lambda \leq 1$. With $\lambda = 0.5$, half of a step in the true bias remains after one result and an eighth after three. Comparing against the wrong instant adds error of its own. A property drifting by 0.1 an hour, compared at a result that returns 2 h after sampling, puts 0.2 into the bias. Chapter 10 corrects a density soft sensor for acid strength from a titration every four hours, Chapter 9 corrects an analyzer and an inferred octane from the laboratory and gates a reading it cannot trust, and Chapter 15 updates the bias of its blend model.

Three chapters turn on what a measurement does to the problem. Chapter 13 cannot measure recovery, which has a maximum and therefore no sign. Its analyzer reads the excess of hydrogen sulfide over twice the sulfur dioxide, which is zero at the maximum, and an ordinary PI loop holds it there. The instrument of Chapter 2 reads part of the rag layer as water, so a loop that holds the reading lets the true water fall as the rag grows, with no error on the trend. In Chapter 15 the product is certified on the average of the whole tank. With $\bar{y}$ the tank's average so far and f the share of the batch still to blend, the rest must average the value of Equation (1.15) to finish on the target y^*.

$$y_h = y^* + \left(y^* - \bar{y} \right) \frac{1 - f}{f} \tag{1.15}$$

A tank 40 % full at 87.0 against a target of 87.2 needs the header at 87.33. At 90 % full it needs 89.0, which the 1.5-unit clip of Chapter 15 does not allow. The authority of the header over the tank shrinks as the tank fills.

1.4.7 The Layers From Regulatory Control to Optimization

Refinery control is built in layers, each faster than the one above and each setting the targets of the one below [22, 23]. Regulatory loops hold flows, pressures, levels and temperatures in seconds to minutes. Constraint control decides which loop owns a shared actuator when a limit binds: overrides, low and high selectors, valve position control and split ranges, acting in minutes. Quality control trims the regulatory setpoints from analyzers, laboratories and soft sensors over tens of minutes to hours. Multivariable predictive control moves several handles together against several limits [7], and optimization sets the targets from prices and models over hours to days [24].

> ✦ DID YOU KNOW? **Refining Holds the Largest Block of Predictive Controllers**
>
> A vendor survey taken in mid-1999 estimated 4542 applications of linear predictive control by five suppliers. Refining held 1985 of them, the largest single block. The count leaves out the controllers that operating companies built for themselves [7].

The layers share one rule. Most economic optima of a refinery lie on a limit, because throughput and severity both earn money until some piece of equipment says stop. The exceptions sit inside the range, such as the salt minimum of Chapter 2 and the recovery maximum of Chapter 13, and each is found by an optimizer or turned into a setpoint. A layer that pushes throughput must therefore know which limit binds, and the binding limit changes. A low selector over one throughput handle is the classical answer, with external reset so that the loops not selected do not wind up, and its regions can be designed systematically [25]. A coordinating layer does the same across units, raising the plant's rate until its tightest unit binds [26]. R-1's override of Equation (1.10) is the smallest case: one handle, one limit, and a throughput that falls only when the limit binds.

Table 1.2 lists, for each chapter, the handle that sets throughput, the quality the unit holds and how it is known, and the equipment limit with the structure that guards it. Chapter 16 adds the measured results and asks how each unit balances throughput, product quality and equipment limits.

1.5 Instruments and Final Elements in This Volume

Every loop in this volume begins at a sensor and ends at a final element, the device that changes a flow of material or of energy. Figure 1.6 lists what the controllers of each chapter read and what they move, taken from the chapter's variables table and from its instrument model. A flow is read in all fourteen chapters, an on-line analyzer in twelve and a temperature in eleven. A valve or a flow loop is moved in thirteen, and the firing or duty of a heater in ten.

Each device sets a limit that no tuning removes. A thermocouple reaches the stream through a thermowell, drawn in Figure 1.7, and every thermowell adds a first-order lag through its own mass and specific heat [27]. The chapters carry that lag as a time constant, 36 s on the column inlet of Chapter 3. A flow read across an orifice plate loses accuracy below about a third of the meter's range [27].

The analyzers bring the longest delays. A chromatograph separates the components of a sample in time, as shown in Figure 1.8, so cycle times of minutes are not uncommon and some dead time is inevitable [27]. Equation (1.12) turns a cycle into an effective delay. The cycles in this volume include 10 s on the Wobbe meter of Chapter 14 and 25 min on the conversion analyzer of Chapter 7. Where no analyzer serves, a laboratory sample or an inferred property stands in, and eight chapters carry an inference.

The final element is most often a control valve, drawn in Figure 1.9. Static and dynamic friction turn a smooth signal into stair-steps of the stem, an effect called stiction [27]. A valve at the end of its travel has no authority left, which is the case that the override of Equation (1.10) guards.

The platform sets the last limit. In a distributed control system a supervisory controller sends setpoints to distributed field controllers, which move their actuators from sensor feedback, and cycle times run from milliseconds to minutes [28]. Figure 1.10 places the layers of Section 1.4.7 on that

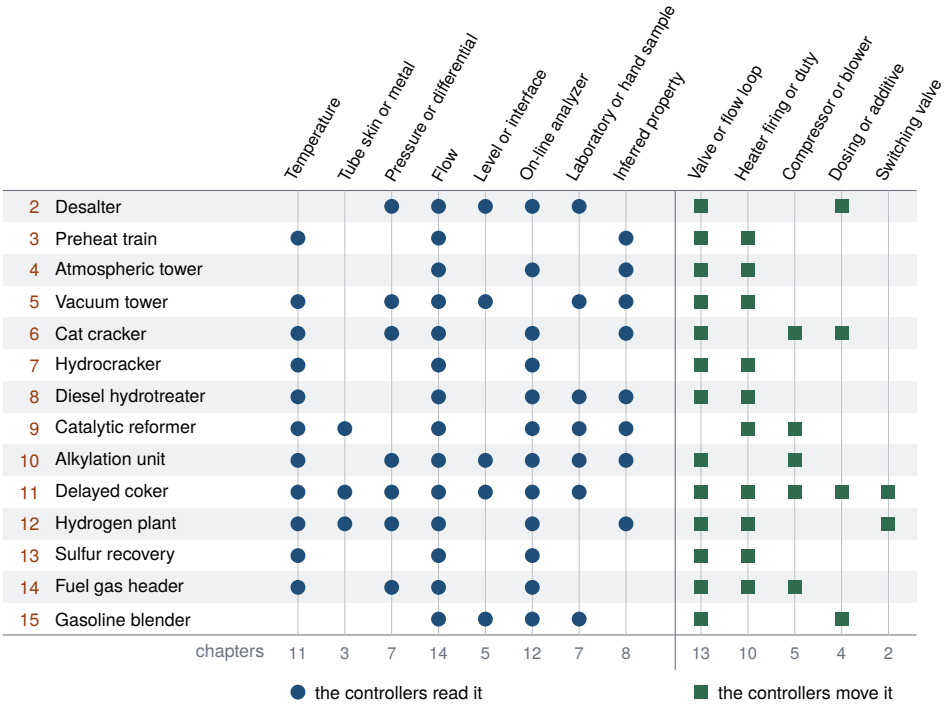


Figure 1.6 Sensors and final elements by chapter. A round mark means that the chapter's controllers or operator read the device, a square mark that they move it. A mark is set only where the chapter's variables table or its instrument section names the device or its reading, so a blank says that the model leaves the device out and not that a real unit lacks it. Notice that every chapter reads a flow and that all but two read an analyzer.

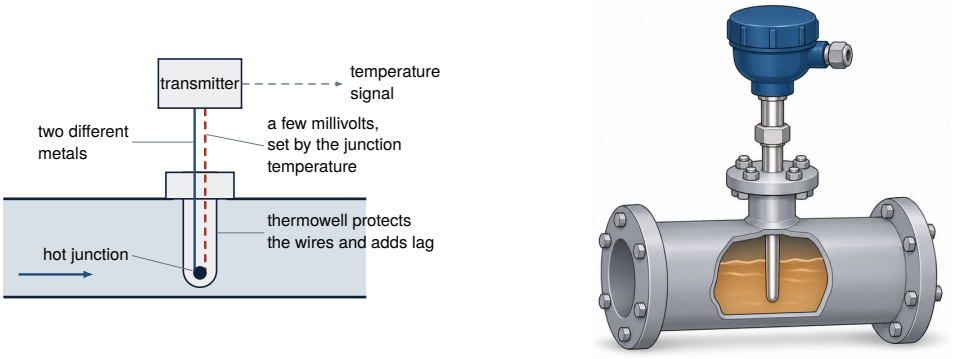


Figure 1.7 Principle of a thermocouple in a thermowell, with a view of the device as installed. Notice that the junction sits inside a metal well, so the reading follows the stream through the well's lag.

Table 1.2 For Each Chapter, the Throughput Handle, the Quality Held and How It Is Known, and the Equipment Limit With Its Guard

Ch.	Throughput	Quality, and how it is known	Limit, and its guard
2	Charge, set upstream	Salt and water in the crude, sampled, and a misread interface	Grid current and trip, an optimizer on the mixing pressure drop
3	Charge	Column inlet temperature, measured	Firing, valve position override on the charge
4	Charge, set upstream	End points and cold properties from inferred cuts	Overflash and heater duty, overflash override
5	Lift by pressure and COT	Heavy gas oil 95 percent point, inferred	Ejector break and tube skin, overrides in a maximizer
6	Feed, at a low selector	Severity, as the riser outlet temperature	Blower, compressor, cyclone and valve limits, four constraint controllers
7	Feed	Conversion on a 25 min analyzer	Bed rise and temperature, override on a compensated rise, trips
8	Feed	Sulfur on a 30 min analyzer	Outlet temperature and heater duty, quench override, mid-ranging
9	Feed	Octane from an analyzer, the laboratory and an inference	Tube metal, regenerator capacity, the operator and the MPC
10	Olefin feed	Acid strength titrated every four hours, a density soft sensor	Refrigeration capacity and the strength floor, a supervisor
11	Charge on a drum cycle	Fractionator cuts	Tube skin and foam, schedule feedforward and MPC
12	Plant rate on header pressure	Steam to carbon, purity from the adsorber	Steam floor and tube metal, lead-lag and an override
13	Acid gas, set upstream	Tail gas ratio on a 3 min analyzer	Sulfur dew point and stack permit, dew point override
14	Off-gas, set by other units	Header pressure and Wobbe index, measured	Import and flare at once, split range with a dead band
15	Blend rate into a tank	Tank-average octane and RVP, analyzer and laboratory	Component stocks and giveaway, an LP and average targeting

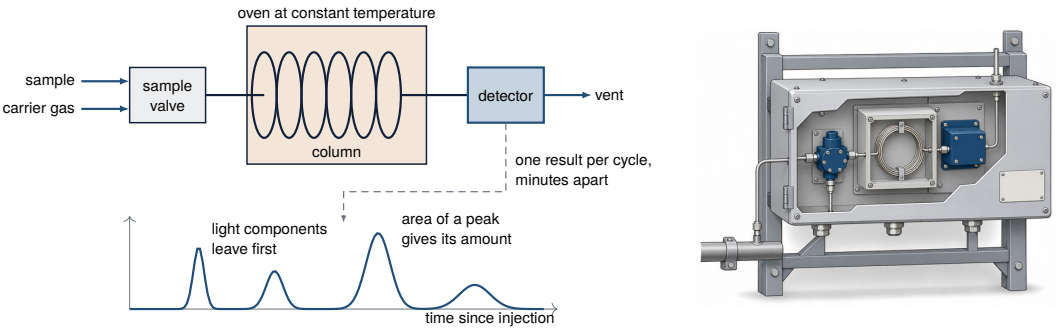


Figure 1.8 Principle of a process gas chromatograph. Notice that one sample gives one result at the end of the cycle, so a loop on this instrument acts on a value that is always old.

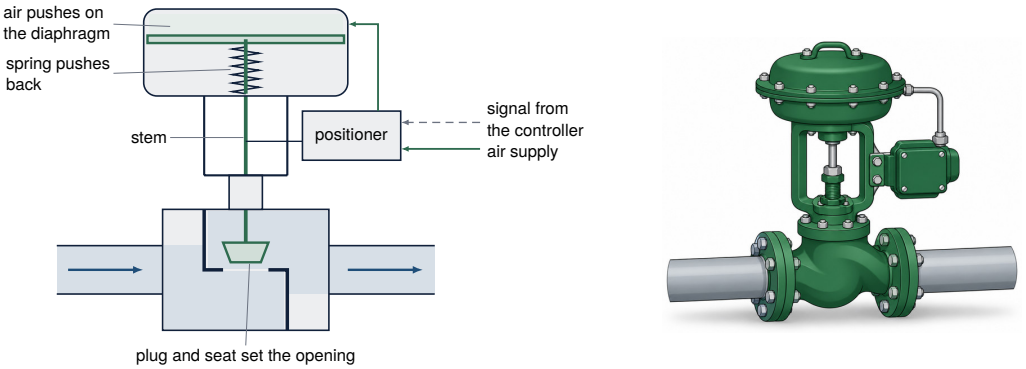


Figure 1.9 Principle of a globe control valve with a spring-and-diaphragm actuator and a positioner, with a view of the device as installed. Notice that the positioner compares the controller's signal with the position of the stem and adjusts the air until the two agree [27].

Layer	What it does	Where it runs	How often
Optimization	sets targets from prices and models	**Supervisory computers** send setpoints to the controllers below	every hour for a unit, every day for the plant
Predictive control	moves several handles together against several limits		every minute
Quality control	trims setpoints from analyzers, the laboratory and soft sensors	**Distributed control system** field controllers with PID, lead-lag and select blocks	at each analyzer or laboratory report
Constraint control	overrides, selectors, split ranges, valve position control		every second to every minute
Regulatory control	holds flows, pressures, levels and temperatures		every second
Field devices	transmitters, analyzers, valves and drives at the unit	point-to-point wiring or a fieldbus	continuously

Figure 1.10 The control layers of Section 1.4.7, the platform that hosts each and how often each acts. The periods of the regulatory, predictive and optimization layers are those of [7, 23], and the platforms those of [28]. Notice that each layer sends setpoints to the layer below and acts less often than it.

hardware. Regulatory loops execute about every second, constraint and predictive control every minute, a unit optimizer every hour and a plant-wide optimizer every day [7, 23]. The apps execute their loops every 6 s in Chapter 11 and every 10 s in Chapter 8.

1.6 The Simulator Framework

Every chapter of this book has an app that runs in a web browser, collected with the others at https://apmonitor.com/apps/. About 240 apps exist across the series. All but a few are built on one shared framework, so the engine, the PID algorithm, the trends and the data export behave the same way in every chapter. The exceptions are standalone studios, such as the cascade studio and the two Temperature Control Lab apps, which follow the same ideas in their own code. This section describes the framework once and states where the standalone apps differ. Table 1.3 lists what a reader meets.

Table 1.3 What Every App Provides. Standalone Apps Are Described in Their Own Chapters.

Element	What it does
Run, Pause, Reset	Start, hold and restart the run. Reset keeps the reader's settings.
Speed	Simulated time per second of real time. It changes the pace and never the result.
Manual and Auto	Hand each actuator to the reader or to the controller, without a bump.
Scenario list	Numbered cases, each with a briefing, a completion rule and a reload button.
Control panel	Setpoints, manual outputs, tuning, options and disturbances, at a Beginner or an Advanced level.
Trends	SP and PV above the MV for each loop, with limits, event marks and a cursor.
Scorecard	IAE, overshoot, settling time, MV travel and time at a limit for each loop.
Learn card	Predict, Try It, two hints and Explain, opened by the reader in that order.
Show Model	Equations, live variable values, editable parameters and assumptions.
Export CSV	Every logged sample of the run as a text file with a header row.
Instructions	A separate page with background, an operating guide and the model's limitations.

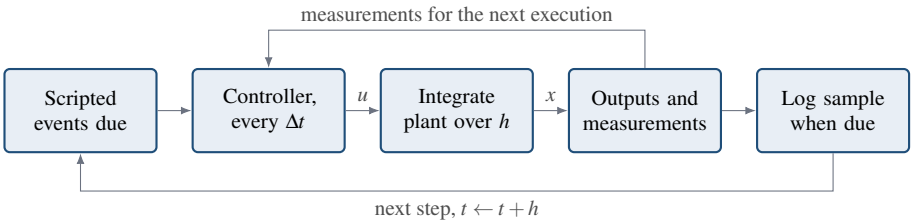


Figure 1.11 Work of the Engine in One Integration Step

1.6.1 The Simulation Engine

Each app carries a plant model written as ordinary differential equations $\frac{dx}{dt} = f(x,u,d,p)$ in the states x, the inputs u, the disturbances d and the parameters p. The engine integrates the model with the classical fourth-order Runge–Kutta method on a fixed step h set for the fastest dynamics of the app. Some apps supply their own step, such as explicit Euler for a discontinuous right-hand side or an exact update for a model driven by events.

Three clocks run at separate rates. The plant advances every step h. The controller executes every control period Δt, equal to h unless the app declares a slower controller, and holds its outputs between executions. The log takes one sample at a fixed interval of simulated time for the trends, the scorecard and the exported file. The order of work in one step is shown in Figure 1.11. The controller also executes once when a scenario is loaded, so every run starts from an initialized controller.

Drawing the screen is kept apart from these clocks. A speed of 5 means five units of the app's time base per second of real time. Speed, frame rate and pauses change how fast a run is shown and never what it computes, so a phone and a workstation produce the same trajectory and the same file.

Noise comes from a deterministic pseudorandom generator. Its seed is declared by the app, or by a scenario that needs its own noise record, and it is restored whenever a scenario is loaded or Reset is pressed. The framework gives the reader no seed setting, so *default seed* in this book means the seed the app ships with. A run repeats bit for bit when the same scenario is loaded and the same settings are made before Run. A change made by hand during a run lands on a different step each time, which is one reason numeric answers carry a tolerance. The standalone simulators draw noise from the browser's own generator, so a noisy run there repeats in its statistics and not sample for sample.

Run and Pause start and hold the clock. Reset restarts the run at $t = 0$, paused, and keeps the reader's settings: every setpoint, tuning value, option, model parameter, disturbance slider and mode the reader changed is put back before the plant and the controller are initialized. Scripted events replay, and a one-shot disturbance button pressed during the run is cleared. The reload button beside

the scenario list restores the scenario's own settings instead.

1.6.2 The Velocity-Form PID Algorithm

One PID implementation serves the PID loops of every framework app. It is the velocity (incremental) form: each execution adds the change in the proportional and derivative terms to a running sum, and the integral is a separate term, as shown in Equations (1.16) to (1.19) for execution k and control period Δt.

$$e_k = a\,(\mathrm{SP}_k - \mathrm{PV}_k), \qquad P_k = K_c\,e_k, \qquad D_k = -a\,K_c\,\tau_D\,\frac{\mathrm{PV}^f_k - \mathrm{PV}^f_{k-1}}{\Delta t} \tag{1.16}$$

$$S_k = S_{k-1} + (P_k - P_{k-1}) + (D_k - D_{k-1}) \tag{1.17}$$

$$v_k = u_b + S_k + I_{k-1}, \qquad u_k = \mathrm{sat}(v_k) \tag{1.18}$$

$$I_k = I_{k-1} + \left[\frac{K_c}{\tau_I}\,e_k + \frac{u_k - v_k}{T_t}\right]\Delta t \tag{1.19}$$

where $a = \pm 1$ is the controller action, PV^f is the measurement after an optional first-order filter, u_b is the output bias, S is the running sum of proportional and derivative increments, v is the output before limiting, u is the output sent to the actuator, I is the integral term and T_t is the tracking time constant, equal to τ_I unless a loop sets another value. The function sat clamps v_k to the actuator range and then applies a rate limit if the loop declares one.

The limit acts on the emitted output u_k only. The sum S_k is never clamped, so it equals $P_k + D_k$ and the output before limiting is the position form $u_b + K_c e_k + D_k + I_{k-1}$. Anti-windup touches the integral alone. While the output is limited, the last term of Equation (1.19) drives I toward the value that reproduces the output applied, which is back-calculation [29]. With no integral action, Equation (1.19) is skipped. Clamping the running sum would instead discard each increment that crosses a limit and keep each one that returns, so symmetric noise would push the output off the limit even in a proportional-only loop.

The derivative acts on the measurement, so a setpoint step gives a proportional kick and no derivative kick. A change of K_c acts on the whole error at the next execution, as in the position form. The action $a = +1$ suits a process whose PV rises with the MV, a reverse-acting controller in the ISA convention, and $a = -1$ gives the direct-acting one.

In Manual the controller keeps executing and sets $I_k = u_{\mathrm{man}} - u_b - S_k$, so the output it would send equals the manual value u_{man}. The first automatic output differs from it by one period's change in P and D. In Auto the engine copies each controller output to the manual station, so a switch to Manual is also bumpless. The standalone studios with PID loops use velocity-form code of their own, with the limit on the output and anti-windup on the integral alone, in which a gain change applies only to later increments.

1.6.3 The Predictive Control and Estimator Options

More than 100 apps offer model predictive control (MPC) as an option beside the structure they open with, which stays the default. Volume 9 develops the theory. In most of them the controller plans with a small linear model in discrete time as shown in Equation (1.20).

$$x_{k+1} = A\,x_k + B\,u_k + E\,w_k + c, \qquad y_k = C\,x_k \tag{1.20}$$

where w is a measured or scheduled disturbance and c is a constant left by linearization. The model comes from a step test at the design point, from the app's balances linearized there or again at every decision, or from the model the reader fits. It never reads the true states or parameters of the run, so a parameter the reader changes in Show Model is model mismatch, as on a real plant.

At each decision the controller minimizes the objective of Equation (1.21) over a horizon of N steps.

$$J = \sum_{k=1}^{N} (y_k - r_k)^{\top} Q (y_k - r_k) + \sum_{j} \Delta u_j^{\top} R \, \Delta u_j + \sum_{i} \rho_i s_i^2 \tag{1.21}$$

where r is the setpoint, Q weights the tracking error, R is the move suppression on each input change Δu_j and ρ_i is the price of a slack $s_i \geq 0$ that lets an output band be broken when it cannot be held. Actuator limits are hard constraints. The quadratic program is solved in the browser by a dual active-set method [30]. The first move is applied, and the plan is made again at the next decision.

The plan starts from an estimate of the model state. About half of these apps use a steady-state Kalman filter on the model with an added disturbance state, which absorbs what the model did not predict and removes steady offset without an integral term [31]. Others correct the model output with a filtered bias to the same end. About a dozen apps also offer moving horizon estimation (MHE), which solves the estimation problem over a window of recent samples and can hold bounds that a Kalman filter cannot, such as a concentration that is never negative [32].

With the MPC selected, the tuning fields of the loops it replaces leave the Control panel. In most of these apps the Advanced level shows up to three settings, each a multiple of the shipped value: the prediction horizon, the move suppression and the band price. The trend draws the accepted plan ahead of the present time, and a list of constraint prices shows which limits bind. A plan that is not optimal, not finite or outside its limits is rejected, and a fallback, in most apps the shipped controller, holds the actuator until a plan succeeds. The Model Fitting and Tuning Studio builds the same kind of controller from the reader's own data in its tab, MPC from the fit.

1.6.4 Scenarios, Progress and Data

An app teaches through scenarios, the numbered cases at the top of the Control panel, nine to thirteen in most apps. Loading one sets the model parameters, the controller settings, the disturbances, the mode, the initial state and any scripted events. The information button beside the list opens a briefing in four parts: what the case is, why it matters, what to do and one detail for the practicing engineer. The Learn card holds the lesson back, with Predict and Try It before the run, hints on request and Explain afterward.

Each scenario has a completion rule that the app judges from the run on the simulated clock: a run length, the scripted events passed, a setting made by the reader, a state held for a stated time or a target met. Choosing a case earns nothing. A completed case is marked in the list, and a badge row shows the count and a level, Explorer after one case, Practitioner at half and Completed at all. Progress is kept in the reader's browser with no account. A certificate of progress or of completion can be issued from it, as a record of self-paced work and not an assessed credential.

One case, usually the last, is the graded challenge. It states a target, such as an error band to hold through a sequence of upsets. Changing the model parameters or the disturbances voids the attempt in most apps, and a record made with Reveal suggested tuning says so.

Show Model opens the equations, live values of the variables, the editable parameters and the assumptions. The trends plot SP and PV above the MV for each loop, with actuator limits and marks at setpoint moves, disturbances, mode changes and scripted events. The scorecard reports for each loop the integral of the absolute error (IAE), overshoot, settling time, MV travel and share of time at a limit, with the app's own measures, and it does not combine them into one score. IAE and travel accumulate from Reset, so two runs are compared at the same clock reading.

Export CSV saves the log as a text file with a header row of column names that carry their units and one row per logged sample. In most apps the time column is in seconds, minutes or hours according to the run length. The log holds the most recent 60 000 samples unless the app sets another size. Apps with the MPC option begin the file with comment lines that start with #, which a script must skip.

An app exercise in this book names the scenario by its number and name as the app shows them, every setting that differs from the scenario's defaults, the run length in simulated time, the seed

(the default seed unless stated), what to read and where (a trend, a scorecard tile or a column of the exported file) and a tolerance on a numeric answer. Settings are made before Run unless the exercise gives a time.

1.7 Working With AI Agents While Studying

A reader of this book is likely to have an AI assistant within reach, and the apps make its work easy to test. This section gives the editor's view of how to use one while working through a chapter. In short, the assistant may do the algebra and the scripting, and the reader keeps the checking.

What to delegate. Tasks with a clear input and a result that can be tested are good candidates. An assistant can derive the linearization of a chapter's balance at an operating point and the transfer function that follows from it. It can write a short script that reads the exported CSV file and fits a first-order-plus-dead-time model to a step test. It can propose a first tuning from that fit with a named rule, convert units, explain a term in a model or outline a report. The assistant knows nothing about the app until it is told. Give it the equations and parameter values from Show Model, the column names of the exported file and the scenario that produced the data, as described in Section 1.6.4. Ask it to state its assumptions and the units of every result.

What to check, and how. An assistant returns a fluent answer whether or not the answer is right, so each result needs a test that it could fail. Four tests cover most of the errors a chapter will produce.

- *Units.* Every term of an equation carries the same units, and a process gain is in PV units per MV unit. A gain computed with the valve as a fraction is 100 times the gain of a loop whose valve is in percent.
- *Signs.* Decide from the physics which way the PV moves when the MV rises. Compare that with the sign of the fitted gain and with the controller action of Section 1.6.2.
- *Steady state.* Set the derivatives to zero. The model must return the operating point that Show Model reports, and a linearization must give zero deviation there.
- *A run on the app.* Load the named scenario at the default seed, enter the proposed numbers and run. The trend and the scorecard decide. A tuning that reads well on paper and holds the valve at a limit for half the run has failed.

A fitted model gets one more test: compare it with data it was not fitted to, such as a second step in the opposite direction. When a check fails, report the failure to the assistant with the evidence and ask for the cause, then check the correction the same way.

How the exercises train the habit. One exercise in each chapter is marked *Review the agent*. It presents a tuning, a model or a claim as an AI assistant might propose it, with a flaw, and asks the reader to find the flaw in about five minutes. The flaws are ordinary ones: a unit, a sign, a steady state that does not close, a result measured on the wrong scenario or at a different run length, or a limit that was ignored. Each is found with one of the tests above. The purpose is not distrust. An engineer who can say why an answer is right can hand an assistant much larger tasks and still stand behind the result. That larger picture, in which agents do most of the routine work and the engineer supervises and checks, is the subject of Chapter 16.

1.8 Chapter Roadmap

Table 1.4 lists the chapters with the lesson each adds to this one. The level says how much of Section 1.4 a chapter assumes. The time is an estimate for reading the chapter once and running each case of its app once, at about eight minutes a case.

Chapters 3, 13, 14 and 15 form a first course: a heater with one limit, a trim on a late analyzer, an integrating header and a blend. They rest on Sections 1.4.2, 1.4.5 and 1.4.6. Chapters 2, 4, 8 and 9 add

Table 1.4 Roadmap of the Volume

Ch.	Plant	The lesson	Level	Time
2	Crude desalter	A salt minimum inside the mixing range and an interface reading that drifts with the rag.	intermediate	2.5 h
3	Crude preheat train	Fouling hides in the fuel valve until the charge is the only handle left.	introductory	3 h
4	Atmospheric crude tower	Count the degrees of freedom before pairing loops.	intermediate	2.5 h
5	Vacuum tower	The cheapest handle on the cut rests on an ejector curve that breaks.	advanced	2 h
6	Fluid catalytic cracker	A unit that is its own furnace, and a selector over five limits.	advanced	3.5 h
7	Hydrocracker reactor	Beds that amplify their inlet, protected in layers.	advanced	2.5 h
8	Diesel hydrotreater	Two inlet handles on two clocks and a half-hour analyzer.	intermediate	2 h
9	Catalytic reformer	Four heaters on one octane that is measured late.	intermediate	2 h
10	Sulfuric acid alkylation	Three stocks on three clocks behind one octane.	advanced	2.5 h
11	Delayed coker	A drum switch is a disturbance known hours ahead.	intermediate	2 h
12	Hydrogen plant	The steam must lead the feed, and the waste is the fuel.	intermediate	2 h
13	Claus sulfur recovery	An optimum turned into a zero, behind three minutes of delay.	introductory	2.5 h
14	Fuel gas header	An integrator on a split range whose ends cost money both ways.	introductory	2 h
15	Gasoline blender	Authority over a tank average shrinks as the tank fills.	introductory	2 h

degrees of freedom, heat integration and inferred quality, and Chapters 11 and 12 add a scheduled disturbance and a ratio that must lead. Chapters 5, 6, 7 and 10 combine a moving limit with internal feedback or slow measurement and are best left until last. A reader who follows the crude can read the volume in order, from the desalter to the gasoline tank.

KEY POINTS

- A refinery unit earns more the more oil it processes, until a product specification or an equipment limit stops it.
- A fired heater's heat balance is also the most useful feedforward in a refinery.
- Four cuts fix the eight end specifications of four products, so half are choices and half are consequences.
- A unit that returns part of its own heat amplifies every change and slows its own response by the same factor.
- Most economic optima lie on a limit, so the layer that pushes throughput must know which limit binds.

Where else this applies. The same layers and limits return in every industry volume of the series. Selectors and overrides are developed in Volume 4 and predictive control in Volume 9.

References

[1] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[2] F. G. Cottrell and J. B. Speed, "Separating and collecting particles of one liquid suspended in another liquid," U.S. Patent 987 115, Mar. 21, 1911.

[3] D. Splitter, A. Pawlowski, and R. Wagner, "A historical analysis of the co-evolution of gasoline octane number and spark-ignition engines," *Frontiers in Mechanical Engineering*, vol. 1, p. 16, 2016.

[4] E. T. C. Vogt and B. M. Weckhuysen, "Fluid catalytic cracking: recent developments on the grand old lady of zeolite catalysis," *Chemical Society Reviews*, vol. 44, no. 20, pp. 7342–7370, 2015.

[5] U.S. Environmental Protection Agency, "Control of air pollution from new motor vehicles: Heavy-duty engine and vehicle standards and highway diesel fuel sulfur control requirements," *Federal Register*, vol. 66, pp. 5002–5193, Jan. 2001, document 01-2, published January 18, 2001.

[6] M. Pitt, "Taking a look back at control: Part 2," *The Chemical Engineer*, no. 989, Nov. 2023, thechemicalengineer.com/features/taking-a-look-back-at-control-part-2.

[7] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[8] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: an engineering perspective," *The International Journal of Advanced Manufacturing Technology*, vol. 117, no. 5–6, pp. 1327–1349, 2021.

[9] M. J. Kaiser, A. de Klerk, J. H. Gary, and G. E. Handwerk, *Petroleum Refining: Technology, Economics, and Markets*, 6th ed. CRC Press, 2019.

[10] U.S. Energy Information Administration, "Refinery capacity report 2026, table 1: Number and capacity of operable petroleum refineries by PAD district and state as of January 1, 2026," U.S. Department of Energy, Washington, DC, Tech. Rep., 2026.

[11] M. R. Riazi, *Characterization and Properties of Petroleum Fractions*. ASTM International, 2005.

[12] M. Morbidelli and A. Varma, "Parametric sensitivity and runaway in fixed-bed catalytic reactors," *Chemical Engineering Science*, vol. 41, no. 4, pp. 1063–1071, 1986.

[13] W. M. Edwards and H. N. Kim, "Multiple steady states in FCC unit operations," *Chemical Engineering Science*, vol. 43, no. 8, pp. 1825–1830, 1988.

[14] A. Arbel, I. H. Rinard, R. Shinnar, and A. V. Sapre, "Dynamics and control of fluidized catalytic crackers. 2. multiple steady states and instabilities," *Industrial & Engineering Chemistry Research*, vol. 34, no. 9, pp. 3014–3026, 1995.

[15] C. van Heerden, "Autothermic processes," *Industrial & Engineering Chemistry*, vol. 45, no. 6, pp. 1242–1247, 1953.

[16] H. Müller-Steinhagen, "Heat transfer fouling: 50 years after the Kern and Seaton model," *Heat Transfer Engineering*, vol. 32, no. 1, pp. 1–13, 2011.

[17] F. Coletti and S. Macchietto, "A dynamic, distributed model of shell-and-tube heat exchangers undergoing crude oil fouling," *Industrial & Engineering Chemistry Research*, vol. 50, no. 8, pp. 4515–4533, 2011.

[18] F. Reyes and W. L. Luyben, "Steady-state and dynamic effects of design alternatives in heat-exchanger/furnace/reactor processes," *Industrial & Engineering Chemistry Research*, vol. 39, no. 9, pp. 3335–3346, 2000.

[19] V. Pinjala, Y. C. Chen, and D. Luss, "Wrong-way behavior of packed-bed reactors: II. impact of thermal dispersion," *AIChE Journal*, vol. 34, no. 10, pp. 1663–1672, 1988.

[20] B. Joseph and C. B. Brosilow, "Inferential control of processes: Part I. steady state analysis and design," *AIChE Journal*, vol. 24, no. 3, pp. 485–492, 1978.

[21] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[22] M. Morari, Y. Arkun, and G. Stephanopoulos, "Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. process decomposition and the classification of the control tasks. analysis of the optimizing control structures," *AIChE Journal*, vol. 26, no. 2, pp. 220–232, 1980.

[23] S. Skogestad, "Plantwide control: The search for the self-optimizing control structure," *Journal of Process Control*, vol. 10, no. 5, pp. 487–507, 2000.

[24] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson, "RTO: An overview and assessment of current practice," *Journal of Process Control*, vol. 21, no. 6, pp. 874–884, 2011.

[25] A. Reyes-Lúa and S. Skogestad, "Systematic design of active constraint switching using classical advanced control structures," *Industrial & Engineering Chemistry Research*, vol. 59, no. 6, pp. 2229–2241, 2020.

[26] E. M. B. Aske, S. Strand, and S. Skogestad, "Coordinator MPC for maximizing plant throughput," *Computers & Chemical Engineering*, vol. 32, no. 1–2, pp. 195–204, 2008.

[27] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[28] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[29] K. J. Åström and L. Rundqwist, “Integrator windup and how to avoid it,” in *Proc. 1989 American Control Conference*, 1989, pp. 1693–1698.

[30] D. Goldfarb and A. Idnani, “A numerically stable dual method for solving strictly convex quadratic programs,” *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[31] G. Pannocchia and J. B. Rawlings, “Disturbance models for offset-free model-predictive control,” *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[32] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, “Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations,” *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246–258, 2003.

1.9 Exercises

1.1 Theory and calculation · 5 min

R-1’s crude unit is clean, with $UA = 1.5\,\mathrm{MW/K}$, crude entering the train at 40 °C and products at 340 °C, as in Section 1.4.5. The charge falls to 650 t/h and the COT setpoint rises to 370 °C, with $c_p = 2.5\,\mathrm{kJ/(kg\,K)}$.

(a) Find the preheat temperature. *Checkpoint:* 270.6 °C.

(b) Find the heater duty and the firing as a share of 60 MW. *Checkpoint:* 44.9 MW, 74.8 %.

1.2 Concept check · 2 min

The fuel gas of a refinery takes up hydrogen, and its heating value falls 7 % while its Wobbe index falls 3 %. Two heaters are on the header. Heater A holds its fuel flow on a flow loop, and heater B holds its burner pressure. Before their outlet loops act, which statement is right?

(a) both lose 7 % of their heat, because the heat is the heating value times the gas burned

(b) heater A loses about 3 % and heater B about 7 %, because a lighter gas flows faster through a burner

(c) heater A loses about 7 % and heater B about 3 %, because a burner at fixed pressure passes more of a lighter gas

(d) neither loses heat, because hydrogen burns hotter than the gas it replaces

1.3 Theory and calculation · 4 min

R-1’s tower draws naphtha, kerosene, diesel and gas oil at 144, 72, 144 and 50.4 t/h from 720 t/h of crude whose assay is $T_{bp} = 30 + 600\,v$.

(a) The kerosene draw rises 14.4 t/h with the others held. Find the new temperatures of the three cuts below the naphtha. *Checkpoint:* 222, 342 and 384 °C.

(b) Which draw should move, and by how much, so that only the kerosene and diesel cut moves? *Checkpoint:* diesel down 14.4 t/h.

1.4 Concept check · 2 min

An engineer pairs one loop to hold the naphtha end point on the top reflux and a second loop to hold the kerosene flash point on the kerosene draw of an atmospheric tower. What happens?

(a) the loops fight, because both specifications are set mostly by one and the same cut

(b) both loops settle, because each product has its own draw and therefore its own handle

(c) the kerosene loop moves every cut above the kerosene draw, so the naphtha loop saturates

(d) the loops settle only if the kerosene loop is tuned five times slower than the reflux loop

1.5 Theory and calculation · 4 min

R-1’s diesel draw tray boils at 366.23 °C at 2.0 bar for a 330 °C cut, with $A = 12$ and $p_{atm} = 1.013\,25$ bar in Equation (1.13). The tower pressure falls to 1.9 bar with the cut unchanged.

(a) Find the new tray temperature. *Checkpoint:* 363.35 °C.

(b) A loop holds the raw tray temperature at 366.23 °C. In which direction does it move the cut, and by about how much? *Checkpoint:* heavier, by about 2.7 K.

1.6 Theory and calculation · 4 min

A bed of the form of Equation (1.5) has $\Delta T_{ad} = 100\,\mathrm{K}$, converts $x = 0.3$ at design and its conversion grows by a factor e every 20 K of mean temperature.

(a) Find the runaway index r and the steady gain G. *Checkpoint:* $r = 0.75$, $G = 7$.

(b) By what factor is the bed's response slower than its time constant τ_b? *Checkpoint:* 4.

1.7 Theory and calculation · 3 min

A feed and effluent exchanger returns $\varepsilon_u = 0.4$ of the effluent's temperature above the feed to the bed inlet, as in Equation (1.11).

(a) Find the loop gain and the amplification of a heater move for a bed with $G = 1.5$. *Checkpoint:* $L = 0.6$, $A = 2.5$.

(b) Find the same for R-1's bed, $G = 2$, and the largest ε_u that keeps that bed open-loop stable. *Checkpoint:* $L = 0.8$, $A = 5$, $\varepsilon_u < 0.5$.

1.8 Concept check · 2 min

A slide valve on a catalytic cracker opens by three percent in Manual. The riser outlet rises 2.1 K in two minutes and about 7.2 K after ninety. Which statement explains the difference?

(a) the riser thermocouple has a ninety-minute lag, so the first reading is simply late

(b) the regenerator's air blower responds to the valve and adds heat after a delay

(c) the cracking reaction is exothermic, so a hotter riser releases more heat of its own

(d) the extra catalyst makes more coke, and burning it reheats the returning catalyst

1.9 Theory and calculation · 5 min

R-1's preheat train has fouled to 60 % of its clean UA, so $UA = 0.9\,\mathrm{MW/K}$. The valve position override holds the firing at 95 % of 60 MW and the COT at 365 °C.

(a) Estimate the charge the override allows if the preheat stayed where it is at 720 t/h. *Checkpoint:* 621 t/h.

(b) Check that 653.6 t/h meets the firing target once the preheat is recomputed at that charge. *Checkpoint:* preheat 239.4 °C, duty 57.0 MW.

1.10 Theory and calculation · 5 min

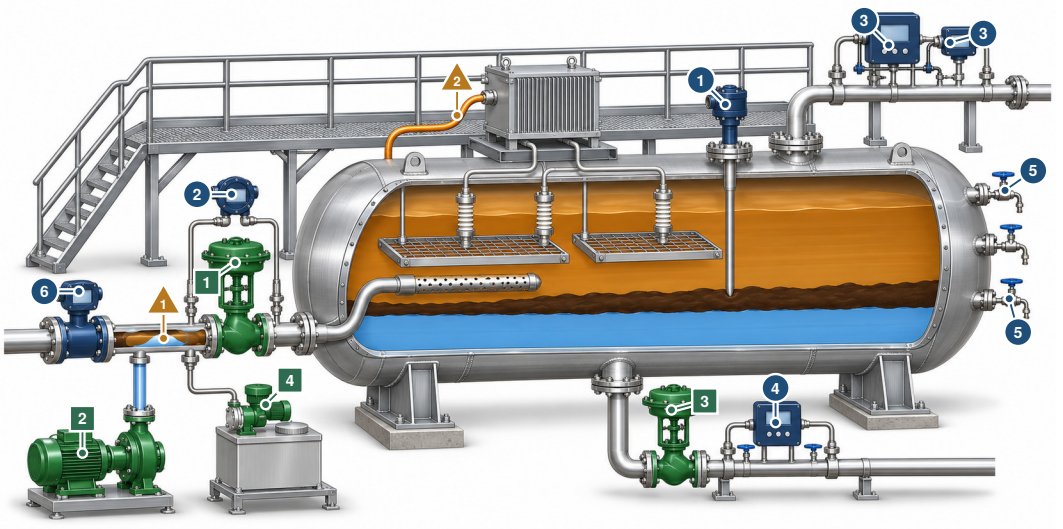
A gasoline tank's target is 87.2. Use Equation (1.15).

(a) The tank is 75 % full at an average of 87.0. At what value must the header run for the rest of the batch? *Checkpoint:* 87.8.

(b) The tank is 92 % full at 87.0, and the header may run at most 1.5 above target. With the header at that limit until 95 % full and at target after, where does the tank finish? *Checkpoint:* 87.06.

CHAPTER 2

Crude Desalter Simulator With a Mixing Optimum and a Misread Rag Layer



The controller moves 1 mixing valve 2 wash-water pump 3 brine valve 4 demulsifier pump

The controller reads 1 interface probe 2 pressure drop across the mixing valve 3 salt analyzer and water meter 4 oil in the brine 5 try-cocks for profile samples 6 crude flow

What disturbs it 1 a slug of free water 2 the transformer supply ▲ a heavier crude

The rod of the interface probe ends in the dark rag between the oil and the water, and the perforated inlet distributor lies in the oil between the rag and the grids. Illustration, not to scale.

IN PLAIN TERMS Washing the salt out of crude oil

The main idea. Crude oil arrives with a little salty water in it. A desalter mixes in fresh water to wash the oil, then lets the water sink and take the salt with it.

The challenge. Mixing harder washes better but makes drops too small to sink, so the best setting lies in between and moves with each oil. The gauge for the water level can be fooled by the rag, a layer of oil and water that will not come apart, and the controller then holds the wrong level.

Why it matters. Salt left in the oil turns to acid and scale in the units downstream.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the two History boxes in Section 2.2, the Did you know? boxes in Sections 2.1 and 2.6, the Try it boxes in Sections 2.3 and 2.4, and Key points in Section 2.9. No equations.
Study	60 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Section 2.6 with Figure 2.4, Section 2.7, Section 2.8 and its Open challenge box.

The control question Harder mixing contacts more brine but makes drops too small to settle, and the interface probe reads part of the rag as water. Where should the mixing pressure drop sit when the crude changes, and what does a loop on that reading do to the true water and the grid?

Learning objectives

- Explain the interior salt minimum, and why the demulsifier decides whether any pressure drop meets the water specification.
- Compute the integrating gain of the water layer and the offset of a proportional-only interface loop.
- Predict how the true water and the rag top move under a reading that counts part of the rag as water.
- Compute the grid voltage, current and trip gap.
- Compare tunings, the optimizer, the rag trim and a fixed dose on measured runs.

Prerequisites Sections 1.6.2 (PI loops), 1.4.6 (analyzers and the misread interface), 1.4.7 (optimizers) and 1.6.4 (cases).

The app https://apmonitor.com/apps/crude-desalter/

See also Volume 3, Chapter 6 and Volume 3, Chapter 7 for doses and ratios trimmed on a slow quality. Volume 5, Chapter 3 for a slow analyzer and Volume 5, Chapter 12 for a misleading sensor. Volume 9, Chapter 8 and Volume 9, Chapter 10 for setpoints chosen above the loops.

2.1 Introduction

Crude oil reaches a refinery with a few tenths of a percent of water in it, and that water is brine. In the generic unit simulated here, 500 m^3/h of crude at 130 °C arrives with 10 pounds of salt per thousand barrels (PTB) dissolved in 0.2 % of water. It takes in 5 % of wash water ahead of a mixing valve and leaves an electrostatic vessel with 0.67 PTB of salt and 0.104 % of water, so 93.3 % of its salt goes out with 25.5 m^3/h of brine. Raising the mixing valve's pressure drop from 1.0 to 1.29 bar lowers the salt to 0.54 PTB. Raising it further, to 2.0 bar, brings the salt back up to 0.89 PTB while the basic sediment and water (BS&W) left in the crude climbs to 0.45 %, more than twice its 0.2 % specification.

The salt matters to every unit downstream. Calcium and magnesium chlorides hydrolyze to hydrochloric acid when they meet steam at crude-unit temperatures, salts deposit as scale where water turns to steam, and metals carried with reservoir solids poison hydrotreating and catalytic cracking catalysts [1, 2]. The desalter is therefore the first process unit a refinery runs and the one whose upsets reach all the others.

> **DID YOU KNOW? Every Barrel Passes This Unit First**
>
> Desalting is the first step of the refining process [3]. On January 1, 2026 the 130 operable refineries of the United States could distill 18.16 million barrels of crude oil per calendar day [4]. The unit simulated in this chapter charges about 75 000 barrels per day.

The simulator is built around two beliefs that the unit contradicts. The first is that more mixing always washes out more salt. A mixing valve contacts more of the brine with wash water as its pressure drop rises, 83 % at 0.6 bar and 95 % at 1 bar in this model, but it also shears the water into drops that the field and gravity cannot recover, so the salt has an interior minimum and the optimum moves with every crude. The second belief is that an interface loop sitting on its setpoint holds the water where the setpoint says. The simulated capacitance probe counts half of the rag as water. With a 500 mm rag on a heavier crude the probe reads 1000 mm with the water at 750 mm, and over eight hours with the reading on its setpoint the water falls to 632 mm while the brine passes 200 mg/L of oil at 192 min.

Both features belong to classes a control engineer meets well outside a refinery. A manipulated variable with two opposing effects on one quality has an optimum that moves with the feed, the

situation in which a steady-state optimizer above the regulatory loops earns its place [5], as excess air does on a burner or reflux on a column. It is one of the few optima inside the operating range that Section 1.4.7 names, beside the recovery maximum of Chapter 13. A measurement whose bias grows with an unmeasured state of the process is followed faithfully by any loop that trusts it, and a level that integrates the difference of two flows is the textbook integrating process, with tuning rules of its own [6, 7]. The desalter adds a hazard above the level, since the energized grid trips when a conductive layer comes close to it.

The desalting literature treats these pieces apart. Electrostatic coalescence is reviewed as physics [8, 9], industrial desalters are simulated at steady state for their salt and water removal [10], and laboratory studies vary settling, chemical, wash water, heat and mixing one factor at a time [11], while a student who meets an interface loop meets it as a transfer function with a clean measurement. This chapter works with a dynamic simulator of a single-stage electrostatic desalter that computes the salt map, the rag, the interface instrument and the grid current together. It tests how loop tuning, ratio stations, a rag trim and a steady-state optimizer handle a crude switch, a misread interface and a water slug, and it states the simulator completely enough to be rebuilt. The correlations are teaching values, and proportional-integral (PI) control, ratio control and bias-updated optimization are established practice. The results are measurements on one generic unit at the default seed, with the challenge checked at five more. Among them are that the BS&W has a minimum of its own below the salt optimum and that the graded challenge is met by a fixed demulsifier ratio with no trim and no optimizer. The model has twenty states, its water and salt ledgers close to 6×10^{-13} of the throughput, and a measurement boundary keeps the true water, rag and salt hidden from every controller.

Section 2.2 follows desalting from the electrical treater to the modern unit. Section 2.3 sets out the model, from the vessel's geometry to the probe that misreads the rag, and Section 2.4 draws from it the three behaviors the cases turn on: an integrating water layer, an interior salt optimum and a reading that drifts with the rag. Section 2.5 describes the loops, the ratio stations, the rag trim, the correction by profile, the steady-state optimizer and model-free extremum seeking. Section 2.6 measures the ten cases of the app, and Sections 2.8 and 2.9 state the limitations and the conclusions.

2.2 Water, Salt and the Electrical Treater

Two patents mark the start of the methods a desalter combines, one electrical and one chemical. The first History segment gives the electrical patent of Cottrell and Speed [12]. Its warning against complete chains of drops between the electrodes is the grid current and the grid trip of the simulator.

> ⧖ HISTORY **1911: A High Voltage Breaks the Emulsion**
>
> Crude petroleum comes from the wells carrying water drops whose sizes run from easily visible to the limit of a microscope, and the patent of Cottrell and Speed, issued in 1911 and assigned to the Petroleum Rectifying Company of San Francisco, opens by listing what producers then did about them [12]. Some water settled on standing, faster when the oil was heated, and centrifuges had a limited place. Oils whose drops were finer than a thousandth of an inch could be dried only by distillation, where water made the still boil over and fouled its bottoms. The remedy was a high potential, preferably alternating, between electrodes immersed in the emulsion. Under a microscope the drops lined up in chains from one electrode toward the other and coalesced, and the claims require that the coalescing globules not form complete chains that short-circuit the electrodes.

The chemical route was patented three years later, as the second History segment tells [13]. The middle stratum of the three that the treatment left is what a desalter operator now calls the rag. By 1938 a survey of desalting in the Transactions of the AIME reported that settling with and without chemicals had given poor results, that acidized wells had loaded crudes with sodium, magnesium and calcium chlorides at anywhere from a few pounds to more than 1000 pounds per thousand barrels, and that severe corrosion had appeared in the topping and cracking units of many refineries [14].

> ⧖ HISTORY **1914: A Chemical Leaves Three Layers**
>
> Barnickel, of St. Louis, found that a small percentage of a soluble metallic sulfate, preferably copperas, broke the "roily" oils of Louisiana, Texas, Oklahoma and California into three strata, clear petroleum above, clear water below and a stratum of slimy substance between the two [13]. In some cases a hundredth of one percent of the oil treated was enough. The mixture was preferably heated to 100 to 140 °F, the water separated at once and the slimy stratum within eight to twelve hours. By the patent's own account the method had already recovered many thousands of barrels from wells whose production would otherwise have been waste.

The reasons a refinery desalts are now stated in terms of the units after the desalter. Chloride salts form hydrochloric acid in the presence of steam, and the hydrolysis of magnesium, calcium and sodium chlorides has been measured in flowing steam from 100 to 350 °C [1]. Salt also deposits as scale, and metals from reservoir solids adsorb on hydrotreating and cracking catalysts [2]. A current account of the process puts the refinery's target near 1.5 PTB, quotes about 90 % removal for one stage and 99 % for two, gives the mixing valve a pressure drop of 10 to 50 psi and the electrodes 12 to 35 kV, and states the trade the mixing valve makes, smaller drops that improve contact but can stabilize an emulsion the vessel then fails to break [2]. The simulated unit runs inside those ranges, with 1 bar (14.5 psi) of mixing, about 25 kV on the grids and 93.3 % of the salt removed in one stage.

The electrical stage has a literature of its own. A review of electrostatic phase separation finds alternating and pulsed fields preferred over a constant direct field, names chain formation and partial coalescence as the unwanted effects, and concludes that electrocoalescer design has not yet left empiricism [8], and the electrocoalescers used on crude oil have been assessed for their strengths and limits [9]. The emulsions are stabilized by asphaltenes, to a degree that depends on how well the rest of the crude solvates them [15]. Laboratory work on a Kuwaiti crude varied settling time, chemical treatment, fresh water, heating and mixing, and found that too much demulsifier hurt both desalting and dehydration [11], and an industrial desalter has been simulated with its mixing valve and electrostatic drum modeled together [10].

2.3 The Desalter as Modeled

The simulated unit is a generic single-stage desalter on a crude unit charging 500 m^3/h, about 75 000 barrels per day, as shown in Figure 2.1. Crude already heated to 130 °C takes in demulsifier and wash water ahead of the mixing valve, enters a horizontal vessel 3.6 m across and 24 m long through a distributor below two grids, and leaves from the top for the preheat train of Chapter 3. Water settles to the bottom and leaves as brine through a valve under interface control. A capacitance probe and its level transmitter (LT) read the interface, try-cock profile samples report the rag and the water height, an online analyzer and a meter report the salt and the BS&W of the desalted crude, and a third analyzer reports the oil in the brine. Crude flows and salt concentrations are per standard volume at 15 °C, while the oil occupies its volume at 130 °C inside the vessel. The clock is the minute, and flows in m^3/h are divided by 60 where they are integrated.

Each variable in Table 2.1 has a role in the loops, controlled (CV), manipulated (MV), disturbance (DV), state or computed output. Twenty states are integrated. Twelve are physical, the water layer's volume V_w and salt S_w, the rag's volume V_r and salt S_r, the water x_w and salt x_s leaving with the crude, the demulsifier at the drop surfaces D_{eff}, the brine and mixing valve positions u_b and u_m, the wash and dosing flows q_w and q_d, and the share x_h of the heavier crude in the charge. Eight are ledgers that integrate the transit holdups and every boundary flow of water and salt, so that both balances can be checked to round-off.

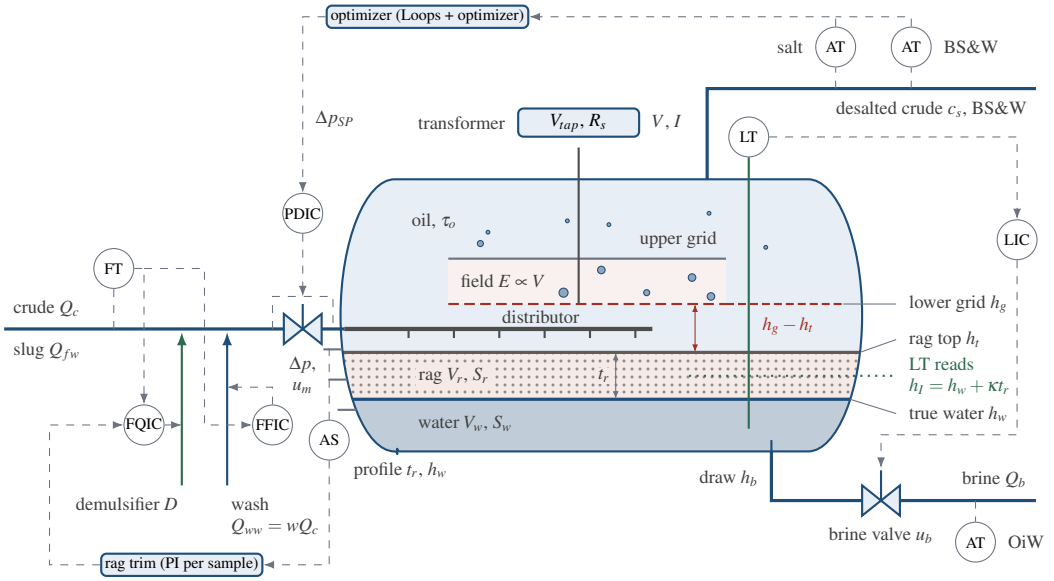


Figure 2.1 Single-stage electrostatic desalter as modeled, after the application's Vessel view. Crude Q_c (and any slug of free water Q_{fw}) takes in demulsifier D and wash water $Q_{ww} = wQ_c$ from ratio stations on the metered crude flow, crosses the mixing valve at pressure drop Δp and opening u_m, and enters the distributor. The water layer (V_w, S_w) rises to h_w, the rag (V_r, S_r, thickness t_r) to h_t, and the oil zone with residence τ_o fills the rest. The lower grid at h_g is fed through the reactor R_s from the tap V_{tap} at voltage V and current I. The lower grid is the dashed line under the field. The probe reads $h_I = h_w + \kappa t_r$ (dotted line), the true water is the solid line under the dotted rag, and the LIC moves the brine valve u_b to hold it. Brine Q_b leaves from the draw at h_b. The try-cock sample (AS) feeds the rag trim, and the salt analyzer and BS&W meter feed the optimizer, which writes the PDIC setpoint. Signals are dashed arrows. Heights at 1 cm per meter in the state of case 7 near 195 min, length not to scale. Process model only, not a piping or safety diagram.

2.3.1 One Volume Function for Three Layers

Every height in the vessel comes from one volume function of the horizontal cylinder and its inverse, as shown in Equation (2.1).

$$\begin{aligned} V_s(h) &= L\left[R^2 \cos^{-1}\frac{R-h}{R} - (R-h)\sqrt{2Rh-h^2}\right] \\ h_w &= V_s^{-1}(V_w), \qquad h_t = V_s^{-1}(V_w + V_r) \\ t_r &= h_t - h_w \end{aligned} \tag{2.1}$$

where $V_s(h)$ is the volume below height h, $R = D/2$ and L is the tangent length. The heads are drawn but carry no volume, so the vessel holds $V_{ves} = \pi R^2 L = 244.3\,\mathrm{m}^3$. The inverse is found by 18 bisection steps and 3 Newton steps, a fixed count, so a height is a function of its volume alone and repeated runs reproduce it exactly. The construction keeps $0 \le h_w \le h_t \le D$ without any test. At design the water stands at 1210 mm under an 80.4 mm rag whose top is at 1290 mm, 660 mm below the lower grid. At 1.25 m the interface is 3.43 m wide and 82.3 m^2 in area, so one cubic meter of water moves it 12.2 mm.

2.3.2 The Crude and Its Assay

Two assays are declared. The design crude has a density at 15 °C of 865 $\mathrm{kg/m}^3$, a viscosity at 130 °C of 1.4 mPa s, 10 PTB of salt in 0.2 % of brine, 5 PTB of filterable solids and 1.0 wt% of asphaltenes.

Table 2.1 Desalter Variables, Their Roles and Design Values (Case 3 at $t = 0$)

Variable	Role	Design	Range
h_I	CV, interface reading	1250 mm	SP 800–1700 mm
Δp	CV, mixing pressure drop	1.00 bar	SP 0.3–2.5 bar
u_b	MV, brine valve	42.5 %	0–100 %
u_m	MV, mixing valve	50.1 %	3–100 %
w	MV, wash ratio	5 % of crude	0–10 %
D	MV, demulsifier	5 ppm	0–60 ppm
Δp_{SP}	MV of the optimizer	1.00 bar	0.3–2.0 bar
assay	DV, crude charged	design	design, heavier
Q_c	DV, crude charge	500 m^3/h	60–110 %
Q_{fw}	DV, free water (slug)	0	0–10 % of crude
a_g	DV, transformer supply	100 %	40–100 %
b_I	DV, probe offset	0	±200 mm
a_w, a_D	DV, wash and dosing pumps	on	on, off
V_w, S_w	states, water layer	72.1 m^3, 37.6 kg	—
V_r, S_r	states, rag	6.61 m^3, 1.73 kg	—
x_w, x_s	states, carried water, salt	0.521 m^3/h, 0.956 kg/h	—
D_{eff}	state, demulsifier at drops	5.0 ppm	—
h_w, t_r, h_t	outputs, layers	1210, 80, 1290 mm	—
c_s	output, salt	0.670 PTB (1.91 mg/L)	—
BS&W	output, water in crude	0.104 %	—
V, I	outputs, grid	24.9 kV, 1.06 A	trip 3 A
Q_b, OiW	outputs, brine	25.5 m^3/h, 63 mg/L	limit 200 mg/L

The heavier crude has 898 kg/m^3, 1.9 mPa s, 10 PTB in 0.5 % of brine, 16 PTB of solids and 3.2 wt% of asphaltenes. Their properties enter the model as shown in Equation (2.2).

$$\begin{aligned} \mathrm{API} &= \frac{141.5}{\rho_{15}/999} - 131.5, \qquad \rho = \rho_{15}\,(1 - 115\,\beta) \\ s_e &= \frac{1}{2}\,\frac{\text{solids}}{5\,\text{PTB}} + \frac{1}{2}\,\frac{\text{asphaltenes}}{1\,\text{wt\%}} \\ \tau_{bl}\,\frac{\mathrm{d}x_h}{\mathrm{d}t} &= \chi_h - x_h \end{aligned} \tag{2.2}$$

where ρ is the oil density at 130 °C with the expansion coefficient β, s_e is an emulsion stability index of 1.0 for the design crude and 3.2 for the heavier one, and a crude switch, which sets χ_h from 0 to 1, blends in over $\tau_{bl} = 10$ min of line mixing. Blend properties are linear in x_h except the viscosity, which is log-linear. The two crudes are API 31.9 and 25.9. API 26 is a medium crude by the usual API gravity classes, in which heavy begins below 22.3, and it is heavy only beside the design crude. What makes it hard to desalt is its stability index, three times the design value. The viscosities lie 4.8 % above and 8.3 % below the Beggs–Robinson dead-oil correlation at API 32 and 26 [16], and the oil densities at 130 °C are 790.4 and 820.5 kg/m^3 against 935 kg/m^3 for the water.

2.3.3 Contact, Drop Size and Coalescence

The mixing valve has two effects, as shown in Equation (2.3). It brings a share η_c of the crude's own brine into contact with the wash water, and it shears the water into drops whose volume median d_v

falls with the pressure drop. The field between the grids then grows those drops to d_m.

$$\begin{aligned} \eta_c &= 1 - \exp\Big[-b\,\Delta p \Big(\frac{w}{w_0}\Big)^{c} \Big] \\ d_v &= d_0 \Big(\frac{\Delta p}{1\,\mathrm{bar}}\Big)^{-a} \\ d_m &= d_v \max\Big\{1,\; G_0 \Big[\Big(\frac{V}{V_0}\Big)^{2} \Big(\frac{\varphi}{\varphi_0}\Big)^{m} \psi \Big]^{1/3} \Big\} \\ \psi &= \frac{1+\alpha\,\theta_D}{1+\alpha\,\theta_{D,0}}\, s_e^{-q}, \qquad \theta_D = \frac{D_{eff}}{D_{eff}+K_D} \end{aligned} \tag{2.3}$$

where w is the wash and free water as a percentage of the crude, V is the grid voltage standing in for the field, φ is the water fraction of the treating zone with design value $\varphi_0 = 0.0494$, and $\theta_{D,0}$ is θ_D at the reference dose $D_0 = 5\,\mathrm{ppm}$. The exponent $a = 0.6$ is the breakup scaling of Hinze [17], the largest stable drop falling as the 0.4 power of the dissipation, with the dissipation in the valve taken to rise as $\Delta p^{1.5}$. The cube root follows from the dipole attraction between drops in a field, which grows as E^2 and as the sixth power of the drop size over the fourth power of their spacing [2]. At a fixed water fraction the spacing scales with the drop size, so the collision rate scales with E^2 and not with drop size, and the mean drop volume grows in proportion to the collisions. With the field 20 % weaker the carried-over share defined below rises from 0.018 to 0.033, where an exponential law in E^2 calibrated to the same design value would give about 0.60. At design $\eta_c = 0.950$, $d_v = 15.0\,\mu\mathrm{m}$ and $d_m = 695\,\mu\mathrm{m}$, a growth of 46.3 times at 24.9 kV.

2.3.4 Settling Against the Rising Oil

A drop settles if its hindered Stokes velocity beats the oil's upflow through the horizontal section at the rag top, which defines a cut size d_c, as shown in Equation (2.4).

$$\begin{aligned} u &= \frac{Q_c/(1-115\beta)}{3600\,L\,c(h_t)}, \qquad F = \Phi\Big(\frac{\ln(d_c/d_m)}{\ln\sigma_g}\Big) \\ d_c &= \sqrt{\frac{18\,\mu\,u}{(\rho_w-\rho)\,g\,(1-\varphi)^{4.65}}} \\ \varphi &= \frac{Q_n + Q_{ww} + Q_{fw}}{Q_c + Q_n + Q_{ww} + Q_{fw}} \end{aligned} \tag{2.4}$$

where $c(h)$ is the chord at height h, $\rho_w = 935\,\mathrm{kg/m^3}$, the exponent 4.65 is the Richardson–Zaki hindrance for creeping flow [18], Φ is the standard normal distribution and F is the share of the contacted water in drops smaller than d_c under a log-normal distribution of spread σ_g. At design the oil rises at 1.83 mm/s over $82.9\,\mathrm{m^2}$, the hindrance factor is 0.790, $d_c = 203\,\mu\mathrm{m}$ at a drop Reynolds number of 0.21, at the edge of the creeping-flow range in which both Stokes' law and the exponent apply, and $F = 0.0182$.

2.3.5 Where the Water and the Salt Go

Brine that the wash water never reached keeps its salt and its fine drops and leaves with the crude, while the contacted water settles except for the share F, as shown in Equation (2.5).

$$\begin{aligned} Q_{cont} &= \eta_c Q_n + Q_{ww} + Q_{fw} \\ S_{cont} &= \eta_c Q_n c_n + Q_{ww} c_{ww} + Q_{fw} c_{fw} \\ x_w^{*} &= F\,Q_{cont} + (1-\eta_c) Q_n \\ x_s^{*} &= F\,S_{cont} + (1-\eta_c) Q_n c_n \\ Q_{set} &= (1-F)\,Q_{cont}, \qquad S_{set} = (1-F)\,S_{cont} \end{aligned} \tag{2.5}$$

where Q_n is the crude's own brine at salinity c_n, 1.00 m³/h at 14 265 mg/L at design, and c_{ww} and c_{fw} are the salinities of the wash water and of tank-bottom free water. The salt removed is therefore $\eta_c(1-F) = 93.30\,\%$ at design. The salt and BS&W of the desalted crude follow from what it carries, $c_s = 1000\,x_s/Q_c$ in mg/L and BSW $= 100\,x_w/(Q_c + x_w)$ in percent by volume, a BS&W that counts water alone because the model carries no sediment. Because pounds per thousand barrels and milligrams per liter are both mass per standard volume, 1 PTB $= 453.59/158.99 = 2.853$ mg/L whatever the crude, and the density at 15 °C enters only a conversion to mg/kg.

The water and salt in the layers and in transit obey the balances of Equation (2.6).

$$\begin{aligned}
\frac{\mathrm{d}V_w}{\mathrm{d}t} &= Q_{set} - x_r R_f + x_r R_d - f_w Q_b \\
\frac{\mathrm{d}S_w}{\mathrm{d}t} &= S_{set} - x_r R_f c_{set} + x_r R_d c_r - f_w Q_b c_w \\
\frac{\mathrm{d}V_r}{\mathrm{d}t} &= R_f - R_d, \qquad \frac{\mathrm{d}S_r}{\mathrm{d}t} = x_r R_f c_{set} - x_r R_d c_r \\
\tau_o \frac{\mathrm{d}x_w}{\mathrm{d}t} &= x_w^* - x_w, \qquad \tau_o \frac{\mathrm{d}x_s}{\mathrm{d}t} = x_s^* - x_s
\end{aligned} \tag{2.6}$$

where $x_r = 0.5$ is the water share of the rag, $c_{set} = S_{set}/Q_{set}$, $c_r = S_r/(x_r V_r)$ and $c_w = S_w/V_w$ are the salinities of the settling water, the rag water and the layer, f_w is the water share of the brine draw from Equation (2.9), and $\tau_o = (V_{ves} - V_w - V_r)(1 - 115\beta)/Q_c$ is the residence time of the well-mixed oil zone, 18.2 min at design. Carried water and salt therefore reach the outlet with a first-order lag. At design 0.521 m³/h of water leaves in the crude, a BS&W of 0.104 %, with 1.913 mg/L or 0.670 PTB of salt.

2.3.6 The Rag Layer

The rag forms from the arriving emulsion in proportion to its stability and to the mixing, and it is broken by demulsifier and by its own slow resolution, both slower on a stable emulsion, as shown in Equation (2.7).

$$\begin{aligned}
R_f &= \min\left[k_r s_e \left(\frac{\Delta p}{1\,\mathrm{bar}}\right)^n \frac{Q_c}{Q_{c,0}},\ \frac{Q_{set}}{x_r}\right] \\
R_d &= \frac{k_d\,\theta_D + 1/\tau_r}{s_e} V_r
\end{aligned} \tag{2.7}$$

where the cap keeps the rag from taking more water than settles. Over cases 1 to 9 as shipped it binds only after the wash pump of case 5 trips, when nothing is contacted and nothing settles. The steady rag volume $R_f s_e/(k_d\theta_D + 1/\tau_r)$ therefore grows as s_e^2. At design $R_f = 2.48\,\mathrm{m}^3/\mathrm{h}$ and the rate constant is 0.375/h, which hold 6.61 m³, an 80.4 mm rag. On the heavier crude at the same 5 ppm the steady volume is 10.24 times larger, 3.2^2, and the rag 832 mm thick, 10.4 times the design thickness because the chord changes with height.

2.3.7 Grid Voltage, Current and Trip

The transformer drives the lower grid through a current-limiting reactor, and the load conductance rises with the crude's solids, with how wet the oil leaves and steeply as the rag top nears the grid, as shown in Equation (2.8).

$$\begin{aligned}
G &= g_o\,[1 + 0.2(s_e - 1)]\,[1 + 0.1\max(0, \omega - 1)] \\
&\quad + g_p\,\mathrm{e}^{-(h_g - h_t)/\lambda} \\
V &= \frac{a_g V_{tap}}{1 + R_s G}, \qquad I = GV
\end{aligned} \tag{2.8}$$

where ω is the carried water per unit of crude against its design value, V_{tap} is the open-circuit tap, an Advanced control at 27 kV, a_g is the supply's availability and $V = 0$ once the grid has tripped.

At design $G = 0.0427\,\mathrm{A/kV}$, of which 0.0027 A/kV is the path to the rag top, so $V = 24.87\,\mathrm{kV}$ and $I = 1.063\,\mathrm{A}$. On the dry design crude the current reaches the 3.0 A trip at a gap of 297 mm, 2.0 A at 375 mm and 1.5 A at 449 mm, so the last 150 mm carry most of the rise, and water at the grid would draw 10.84 A against the reactor's limit of $V_{tap}/R_s = 13.5\,\mathrm{A}$. The reactor also pulls the voltage down as the current rises, which weakens the field when the emulsion is wettest. A current at or above 3.0 A held for 12 s trips the grid, the trip latches, and the run stops.

> **TRY IT IN THE APP** **6 · Let the interface rise: the grid trip (Manual)**
>
> Load the case and press Run. Nothing has to be set. The brine valve is pinched at 30 min and the water begins to rise. Watch the grid current. It is 1.06 A at 30 min and only 1.21 A at 60 min, then passes its alarm at 100 min and trips the grid at 119.6 min, with the rag top 297 mm below the grid.

2.3.8 Brine and the Oil It Carries

The brine valve passes water while the layer stands above the draw nozzle, and the oil it carries rises as the water's residence falls, as rag nears the draw, and once oil is drawn straight through, as shown in Equation (2.9).

$$\begin{aligned} f_w &= \mathrm{clip}\Big(\frac{h_w}{h_b}, 0, 1\Big), \qquad Q_b = Q_{b,max}\,u_b \\ \tau_w &= \frac{V_w}{f_w Q_b}, \qquad \mathrm{OiW} = \mathrm{OiW}_0\sqrt{\frac{3\,\mathrm{h}}{\tau_w}} \\ &\quad + \min\left[\mathrm{OiW}_r e^{-(h_w-h_b)/\lambda_b},\ 2\times10^5\right] \\ &\quad + (1-f_w)\,8\times10^5\ \mathrm{mg/L} \end{aligned} \tag{2.9}$$

where $h_b = 150\,\mathrm{mm}$ is the draw nozzle and the brine valve is linear in flow. At design the water turns over in $\tau_w = 2.83\,\mathrm{h}$ and the brine carries 63.2 mg/L of oil against a 200 mg/L limit to the wastewater plant.

2.3.9 Actuators and Instruments

The valves and pumps follow their commands with first-order lags, and the mixing valve's pressure drop is quadratic in flow, as shown in Equation (2.10).

$$\begin{aligned} \tau_{bv}\frac{\mathrm{d}u_b}{\mathrm{d}t} &= u_b^{cmd} - u_b, \qquad \tau_{mv}\frac{\mathrm{d}u_m}{\mathrm{d}t} = u_m^{cmd} - u_m \\ \Delta p &= \Delta p_{open}\Big(\frac{Q_{tot}}{Q_0}\Big)^2 \frac{1}{\max(0.03, u_m)^2} \\ \tau_{ww}\frac{\mathrm{d}q_w}{\mathrm{d}t} &= a_w q_{w,SP} - q_w, \qquad \tau_p\frac{\mathrm{d}q_d}{\mathrm{d}t} = a_D q_{d,SP} - q_d \\ \tau_D\frac{\mathrm{d}D_{eff}}{\mathrm{d}t} &= q_d - D_{eff} \end{aligned} \tag{2.10}$$

where Q_{tot} is the crude with its brine, wash and free water, $Q_0 = Q_{c,0}(1 + w_0/100)$ is its design value, and a_w and a_D are 1 with the wash and dosing pumps running and 0 with either tripped. The demulsifier reaches the drop surfaces and the rag through $\tau_D = 45\,\mathrm{min}$, which makes it the slow handle. At design the mixing valve stands 50.1 % open for 1.00 bar. The charge is held whatever the mixing valve does, so the manual station stops at 15 %, where the drop at design flow is already 11.2 bar.

The interface instrument is the center of the chapter. The probe reads the water and part of the rag, with an offset, a lag and noise, as shown in Equation (2.11).

$$\begin{aligned} y &= h_w + \kappa t_r + b_I \\ \tau_{LT}\frac{\mathrm{d}y_f}{\mathrm{d}t} &= y - y_f, \qquad h_I = y_f + \sigma_I n_k \end{aligned} \tag{2.11}$$

where $\kappa = 0.5$ is the share of the rag the probe counts as water, $\tau_{LT} = 0.5\,\text{min}$, $\sigma_I = 4\,\text{mm}$ and n_k is a unit normal deviate drawn at each control execution. Water's permittivity is many times that of oil, so a capacitance probe responds to a water-continuous emulsion almost as to water and to an oil-continuous one only in part, and a reading partway into the rag stands closer to such a probe than a reading at its top. The value 0.5 is a teaching value chosen for margin in the challenge, where the shipped settings' current would peak at 3.80, 3.39 and 3.10 A with κ at 0.5, 0.6 and 0.7 if the trip were disabled. The BS&W meter has a 2 min lag and 0.003 % of noise, the mixing pressure drop 0.004 bar of noise, the oil-in-brine analyzer a 5 min lag. The salt analyzer draws a sample every 15 min and reports it 10 min later with 0.02 PTB of noise. A try-cock profile is drawn every 24 h, 8 h or 2 h, the first at $t = 0$, and reports the rag thickness and the water height 15 min later with 15 and 10 mm of noise. Flows, valve positions and the grid's voltage and current are read without lag or noise. The random draws of an execution do not depend on any setting, which keeps two runs that differ in one setting aligned noise for noise.

2.3.10 Parameters, Initialization and Numerics

Table 2.2 gives every parameter, marked adjustable (A) when it appears in the application's Parameters panel, hidden (H) when it does not, and fixed (F) for the constants written into the equations. None of the values was measured on a plant. They were chosen as typical of a single-stage desalter and so that the design point closes. A case begins lined out at its own settings. The rag takes its steady volume or the case's starting thickness, the water is placed by bisection so that the instrument reads the interface setpoint, the carried water and the grid voltage are iterated together thirty times, and the brine valve passes exactly the water that settles. The integrator is the classical fourth-order Runge–Kutta method with a step of 2 s. Halving it moves no quantity of the heavier-crude case at 6 h by more than 6.4×10^{-9} of its value. At the end of each of cases 1 to 9 the water ledger closes to 5.8×10^{-13} and the salt ledger to 2.2×10^{-13} of what entered. A simulated hour costs about 80 ms of computing in a JavaScript engine.

2.3.11 Assumptions

The vessel is one oil zone, one rag and one water layer, each well mixed, and the oil carries water to the outlet with a first-order lag. A real desalter has a distributor, a treating zone and a collector, so its outlet answers with more delay and less smoothing. The separation is a teaching correlation and not a population balance. It gives the right directions and plausible sizes, and it is not a design method. Brine the wash water never reached keeps its salt and leaves, so nothing is removed with the wash water lost, while a real desalter would still dehydrate some native brine. The rag kinetics, the grid conductance, the oil-in-brine law and the probe's rag share κ are synthetic correlations chosen for an 80 mm design rag, a current near 1 A at 25 kV and a trip near a 300 mm gap, and none is a prediction for a vessel, crude or chemical. The demulsifier helps monotonically up to the 60 ppm limit, whereas an overdose can hurt a real emulsion [11]. Temperature is fixed at 130 °C and the vessel is liquid full, so heat, pressure and the preheat train are absent, and the heads carry no volume. The grid trip, the alarm, the oil-in-brine limit and the salt and BS&W specifications are teaching values. The real ones come from the transformer rating and the vendor, the refinery's wastewater permit and its crude-unit specifications [2].

2.4 An Integrator, an Optimum and a Drifting Reading

2.4.1 The Water Layer Integrates

The water layer holds the difference between the water that settles and the brine drawn off, and nothing in the vessel pushes back until the water reaches the draw nozzle, so the layer is an integrating process. One percent of brine valve passes 0.6 m^3/h, which moves the interface 7.29 mm/h, the integrating gain. In case 1 the brine valve opens from 42.5 to 54.5 % at 20 min and the true water falls 89.5 mm/h, reaching 663 mm at 360 min with 235 mg/L of oil in the brine. Setting the valve to

Table 2.2 Model Parameters, Marked A When Adjustable in the Parameters Panel, H When Hidden and F When Fixed in the Equations

Symbol	Meaning	Value	Unit	
D, L	vessel diameter, tangent length	3.6, 24	m	A
h_g, h_b	lower grid, brine draw height	1950, 150	mm	A
$Q_{c,0}$	design crude charge (15 °C)	500	m³/h	A
β	crude expansion coefficient	7.5×10^{-4}	1/K	A
ρ_w	water density at 130 °C	935	kg/m³	A
w_0	design wash rate	5	%	H
Δp_{open}	mixing valve drop, wide open	0.25	bar	A
τ_{mv}, τ_{bv}	mixing, brine valve lags	0.1, 0.1	min	A
b, c	contact coefficient, wash exponent	3.0, 0.7	bar, —	A
d_0, a	drop median at 1 bar, exponent	15, 0.6	µm, —	A
G_0	coalescence growth at design	46.5	—	A
m, q	exponents on φ and s_e	1.0, 0.45	—	A
α, K_D	demulsifier effect, half dose	6, 25	—, ppm	A
D_0, V_0	reference dose, voltage	5, 25	ppm, kV	H
σ_g	drop-size spread	1.8	—	A
4.65	hindered-settling exponent	4.65	—	F
τ_D	demulsifier action lag	45	min	A
k_r, n	rag formation, exponent on Δp	2.48, 1.0	m³/h, —	A
k_d, τ_r	rag breaking, resolution time	1.5, 8	1/h, h	A
x_r	water share of the rag	0.5	—	H
κ	share of the rag read as water	0.5	—	A
$t_{r,0}$	starting rag (0 = steady)	0	mm	A

Symbol	Meaning	Value	Unit	
V_{tap}	transformer tap (open circuit)	27	kV	H
R_s	current-limiting reactor	2.0	kV/A	A
g_o, g_p	emulsion, contact conductance	0.040, 2.0	A/kV	A
λ	decay length of the path	100	mm	A
0.2, 0.1	conductance gains on s_e, ω	0.2, 0.1	—	F
I_{trip}, I_{HA}	grid trip (12 s), alarm	3.0, 2.0	A	A
$Q_{b,max}$	brine valve capacity	60	m³/h	A
τ_{ww}, τ_p	wash loop, dosing pump	0.5, 1	min	H
c_{ww}, c_{fw}	wash, free-water salinity	0, 1000	mg/L	A
OiW_0, OiW_r	oil in brine, base and rag	60, 10 000	mg/L	A
λ_b	decay length of rag entrainment	120	mm	A
τ_{bl}	line mixing of a crude switch	10	min	A
salt, BS&W	specifications	1.0, 0.2	PTB, %	A
OiW limit, rag alarm	brine and rag limits	200, 300	mg/L, mm	A
τ_{LT}, σ_l	probe lag, noise	0.5, 4	min, mm	F
—	LIC input filter	1	min	F
—	salt analyzer cycle, delay	15, 10	min	F
—	profile delay / noise (t_r, h_w)	15 / 15, 10	min / mm	F
Assays (design, heavier)				
ρ_{15} / μ	density at 15 °C / viscosity at 130 °C	865, 898 / 1.4, 1.9	kg/m³ / mPa s	F
salt / brine	raw salt / own water	10, 10 / 0.2, 0.5	PTB / %	F
solids / asph.	filterable solids / asphaltenes	5, 16 / 1.0, 3.2	PTB / wt%	F

pass the water arriving, 42.4 %, stops the fall wherever the interface happens to be, and the rag and the reading ride down with the water.

2.4.2 An Interior Optimum in the Mixing Valve

The steady salt and BS&W against the mixing pressure drop, computed with the model's own functions at three states the cases visit, are shown in Figure 2.2. On the design crude the salt is 1.69, 0.67, 0.54 and 0.89 PTB at 0.6, 1.0, 1.29 and 2.0 bar, with its least value of 0.540 PTB at 1.292 bar. Below the optimum too little brine meets wash water, and above it the drops leave the field too small to settle. The BS&W reaches the 0.2 % specification at 1.356 bar, just past the salt optimum, so on the design crude the optimum is feasible with a margin of 0.064 bar. The BS&W is not monotonic

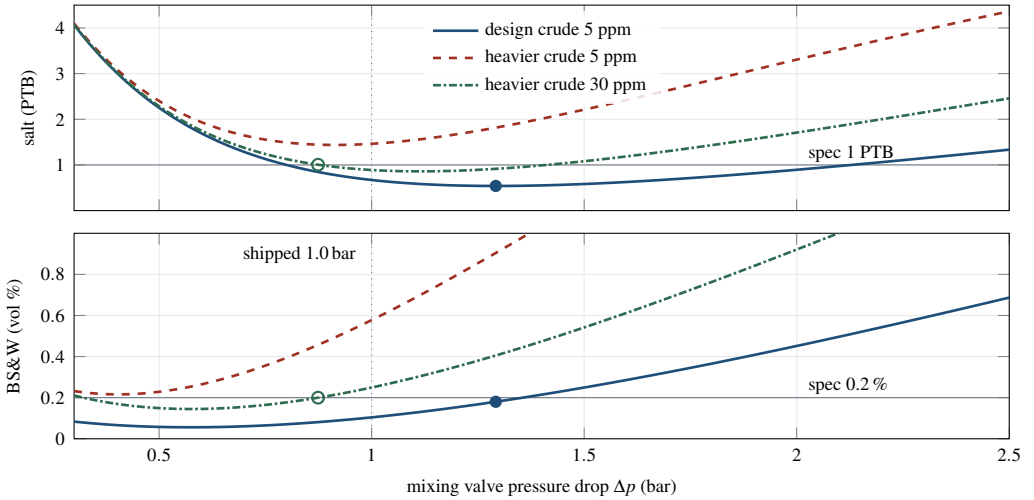


Figure 2.2 Steady salt and BS&W of the desalted crude against the mixing pressure drop, from the application's own map. Design crude at 5 ppm (case 3 at $t = 0$), heavier crude at 5 ppm (case 4 at 240 min, rag 292 mm, dashed) and at 30 ppm (case 4 with 30 ppm from 120 min, at 720 min, rag 320 mm, dash-dotted). The design crude is the solid line. Thin gray lines mark the specifications, and the dotted vertical line the shipped pressure drop. The filled dot is the design crude's least salt, 0.540 PTB at 1.29 bar. The ring is the least salt meeting 0.2 % on the heavier crude at 30 ppm, 1.005 PTB at 0.87 bar. At 5 ppm no pressure drop meets the BS&W specification on the heavier crude.

either. It falls from 0.084 % at 0.3 bar to a minimum of 0.056 % at 0.57 bar and then rises, because below that point the uncontacted brine, which keeps its fine drops, outweighs the settling that gentler mixing allows.

🕮 **Pause.** Try Exercise 2.1 before reading on.

A crude switch moves the whole map. On the heavier crude at 5 ppm, at 240 min in case 4, three hours after the switch, the least BS&W anywhere is 0.216 % at 0.40 bar, where the salt is 3.06 PTB, so no setting of the fast handle meets the specification. Raising the demulsifier to 30 ppm restores feasibility. The least salt with the BS&W at 0.2 % or less is then 1.005 PTB at 0.87 bar, while the unconstrained least salt, 0.860 PTB at 1.11 bar, carries 0.304 % of water. The fast handle chooses where on a curve to sit, and the slow handle decides which curve exists. With the wash water lost the salt is 10.0 PTB, the raw crude's, at every pressure drop.

🕮 **Pause.** Try Exercise 2.3 before reading on.

2.4.3 A Reading That Drifts With the Rag

From Equation (2.11), a loop that holds h_I at its setpoint holds $h_w = h_{I,SP} - \kappa t_r - b_I$, so every millimeter the rag grows lowers the true water by κ millimeters with no change in the reading, no error in the loop and no move of the valve that an operator would notice. The water then has less time to release its oil and the rag comes closer to the draw. The same reading also lifts the rag top as the rag grows, $h_t = h_{I,SP} + (1 - \kappa)t_r - b_I$, toward the grid. A profile sample measures the bias κt_r directly, but only as often as it is drawn, and it ages between samples while the rag keeps changing. Section 1.4.6 places this misread interface among the refinery's measurement problems.

🕮 **Pause.** Try Exercise 2.2 before reading on.

> **TRY IT IN THE APP 7 · The rag fools the interface instrument**
>
> Open the app, pick this case from the list and press Run. Nothing has to be set. Watch the labels at the right of the vessel. The reading stays near its 1000 mm setpoint for eight hours while the true water sinks from 750 to 632 mm. At 192 min the brine passes its limit for oil.

2.5 The Control System as Implemented

The controllers execute every $\Delta t_c = 6\,\text{s}$, three physics steps, and the historian logs a row every 30 s. The order of the work within a step is that of the shared engine (Section 1.6.1). Apart from extremum seeking, which reads less, the controllers see the plant only through a packet of the instrument readings of Section 2.3 and the declared assay of the tank on charge, and one of the application's tests checks that no other key ever appears in that packet.

2.5.1 Interface and Pressure-Drop Loops

Both valve loops use the velocity-form PID of Equations (1.16) to (1.19), executed at Δt_c with no derivative action, a tracking time equal to τ_I and the limit on the output alone. Both are direct acting in the instrument sense ($a = -1$, so the error is PV $-$ SP), a high interface opening the brine valve and a high pressure drop opening the mixing valve. An integral time of 600 min or more removes integral action. The interface controller (LIC) reads h_I through a first-order filter of 1 min, which keeps the probe's noise out of the brine valve, and its defaults are $K_c = 0.3\,\%$ of valve per millimeter and $\tau_I = 60\,\text{min}$ on a 0 to 100 % output. With the integrating gain $K = 7.29\,\text{mm/h}$ per percent, proportional action alone gives a closed-loop time constant $1/(KK_c) = 27.4\,\text{min}$, and a PI loop on an integrating process is critically damped at $\tau_I = 4/(KK_c)$, 110 min here, so the default 60 min trades some overshoot for speed. The pressure-drop controller (PDIC) has $K_c = 20\,\%$ per bar and $\tau_I = 0.5\,\text{min}$ on a 3 to 100 % output.

The header's Manual and Auto switch both valves. In Manual each valve sits at its operator station and both PIs track it as Section 1.6.2 describes, so the transfer back to Auto is bumpless. The ratio stations keep running in Manual, because a wash pump and a dosing pump are always in ratio, while the trim and the optimizer hold.

2.5.2 Ratio Stations and the Rag Trim

The wash water setpoint is $q_{w,SP} = (w/100)\,Q_{c,m}$ on the metered crude flow. The demulsifier setpoint is the operator's ratio D, or the output of a rag trim when the trim is on, much as a slow trim on a ratio dose works in Volume 3, Chapter 6. The trim is the same PI with $K_c = 0.6\,\text{ppm}$ per millimeter, $\tau_I = 120\,\text{min}$ and a target of 80 mm, executed once per reported profile sample over the time since its last execution, as shown in Equation (2.12).

$$q_{d,SP} = \text{clip}\big(D + \text{PI}(t_{r,sample} - t_{r,SP}),\; D,\; 60\,\text{ppm}\big) \tag{2.12}$$

where the floor is the operator's ratio, so the trim only adds. Between samples the dose holds. An earlier version of the trim, replaced when the application was reviewed, tracked the dose on its floor while it was off or at the moment it was switched on. It so stored the gap between a thin rag and its target as integral action the floor hid, 21.7 ppm with the rag at 44 mm against 80 mm, which the first reported sample then released as a dose increase although the rag was still below target. A trim on its floor now holds no positive integral action, the state its own back-calculation reaches there, and with the rag above target the transfer is bumpless.

2.5.3 Correcting the Reading by Profile

With Interface reading set to Corrected by profile, each reported sample updates a bias $b = h_I(t_s) - h_{w,sample}$, where $h_I(t_s)$ is the reading logged when the sample was drawn, and the LIC holds $h_I - b$ instead of h_I. The bias is the update of Equation (1.14) with a gain of 1, so it takes each new sample's

value, and a bound of 1.5 m. The correction removes the rag bias, and in doing so it moves the water, and the rag on it, by the bias at once.

2.5.4 The Steady-State Optimizer

With Loops + optimizer selected, a steady-state optimizer writes the PDIC setpoint, as shown in Equation (2.13).

$$\begin{aligned} \Delta p_{SP} &= \arg \min_{0.3 \le \Delta p \le 2.0} \hat{c}_s(\Delta p)\, \mathrm{e}^{b_s} \\ \text{s.t.} \quad & \widehat{\mathrm{BSW}}(\Delta p)\, \mathrm{e}^{b_w} \le \mathrm{BSW}_{max} \\ & \hat{I}(\Delta p) \le 2.5\,\mathrm{A} \end{aligned} \tag{2.13}$$

where the hats are the steady state of Equations (2.3) to (2.8) computed with the parameters the case was commissioned with, from the metered crude and wash flows, the dosing pump's delivery through the nominal 45 min lag, the declared assay, the tap, and a rag top estimated as $h_I + (1-\kappa)t_{r,sample}$. Free water from a slug is not measured and enters the model as zero. The search runs on a 0.01 bar grid every 30 min and at once when the declared assay changes, the setpoint moves at most 0.25 bar per execution, and BSW_{max} defaults to 0.18 %, a margin under the specification. The biases b_s and b_w compare the logarithms of the salt analyzer and the BS&W meter with the model at the present pressure drop, through a filter gain of 0.3, and only when the plant is steady, which requires the setpoint to have stayed put for 30 min and both the meter and the model's own prediction to lie within 3 % of their values 20 min earlier. A new assay clears them. With no feasible point the optimizer reports infeasible and takes the least value of the salt plus 10 PTB per percent of BS&W or ampere of current over its limit. The optimizer is a one-dimensional search that ran on a model and on measurements, so it is real-time optimization in the narrow sense, and nothing in the application is a model predictive controller.

2.5.5 Model-Free Extremum Seeking

Extremum-seeking control (ESC) perturbs a command to infer an improving direction from measured performance [19, 20]. Here it writes only the PDIC setpoint and uses no input–quality map, declared assay, hidden emulsion state or plant predictor. At sample time s_n, the quality cost is

$$J_n = c_{s,m}(s_n) + 10\max\{B_m(s_n) - B_{max}, 0\}, \tag{2.14}$$

where $c_{s,m}$ is reported salt in PTB, B_m is measured BS&W in percentage points and $B_{max} = 0.18$ is the selected quality limit. The coefficient is 10 PTB per percentage point, so J_n has units PTB. The analyzer draws every 15 min and reports 10 min later. Each distinct fresh timestamp is consumed once and paired with the historical meter reading at s_n, rather than its value when the report arrives.

With period $P = 360$ min, amplitude $a_p = 0.10$ bar and $\tau_f = P/4 = 90$ min, the executed recurrences are

$$\begin{aligned} \alpha_n &= 1 - \exp(-\Delta s_n/\tau_f), \quad h_n = J_n - b_{n-1} \\ b_n &= b_{n-1} + \alpha_n h_n \\ \hat{g}_n &= (1-\alpha_n)\hat{g}_{n-1} + \alpha_n \frac{2h_n}{a_p}\sin\theta_n \\ z_n &= \Pi[z_{n-1} - \Delta s_n\, \mathrm{clip}(k_g \hat{g}_n, -v_c, v_c)] \\ \Delta p_{SP}(t) &= \Pi\left[z_n + a_p \sin\frac{2\pi(t-t_0)}{P}\right], \end{aligned} \tag{2.15}$$

where $\Delta s_n = s_n - s_{n-1}$, $\theta_n = 2\pi(s_n - t_0)/P$, b_n is the baseline, h_n its washout output, $\hat{g}_n$ the filtered gradient estimate in PTB/bar, z_n the center and t_0 the activation or recovery time. Projection Π clips to 0.3–2.5 bar, $k_g = 0.001$ bar^2/(PTB min), and $v_c = 0.001$ bar/min. The first accepted report initializes b to J with $h = 0$, $\hat{g} = 0$ and a 15 min interval. Center adaptation is disabled until $s_n - t_0 \ge \tau_f$. Thus

Table 2.3 The Ten Cases

#	Case, preset and events	Teaching point
1	Manual, brine valve +12 points at 20 min	the integrating layer
2	P-only LIC, SP +100 mm at 60, wash 5 to 7 % at 240 min	P and PI on an integrator
3	Salt map view, optimizer available	the interior optimum
4	Heavier crude from 60 min	rag, grid current, demulsifier
5	Wash pump trips at 360 min	wash ratio, brine load
6	Manual, brine valve pinched to 8 % at 30 min	the grid trip
7	Heavier crude, 500 mm rag, SP 1000 mm, 2 h samples	the rag bias
8	350 mm rag, LIC 0.1/120 min, slug at 60 min	averaging against tight
E	Extremum seeking, 0.9 bar, 30 ppm, crude switch at 360 min	learning from delayed quality
9	Challenge: heavier crude at 3 h, slug at 8 h	the whole problem

the first ordinary adaptation is at clock 100 min, on the sample drawn at 90 min. The center changes only on fresh reports, by at most 0.015 bar for an ordinary 15 min interval. It holds between reports while the probe changes every 6 s. Tuning ranges are 240–720 min for P, 0.02–0.20 bar for a_p and 0–0.005 bar^2/(PTB min) for k_g. The displayed probe is the clipped request minus z_n. A positive gradient lowers the center. For a static smooth cost, $J(z + a_p \sin\theta) = J(z) + a_p J'(z) \sin\theta + O(a_p^2)$ explains the sign and the factor $2/a_p$, without making $\hat{g}_n$ an exact plant derivative.

The raw measured interface reaching $h_g - 0.30$ m is a hold of the ESC alone, not an interface alarm or a rag-top estimate. Measured current at I_{HA}, reported rag at 300 mm, a grid trip, Manual, nonfinite data or salt timestamp age above 45 min also stop probing and adaptation, preserving the current bounded request. Recovery starts acquisition again and discards pre-recovery draws. Numerical failure causes this same hold, without a transfer to the optimizer. The LIC, PDIC, ratio stations, rag trim and physical trip keep their functions. ESC entry has zero setpoint and valve jump in the transfer tests. On ESC exit to the map optimizer, the new target jumps 0.247 08 bar, while PDIC tracking keeps the valve command continuous. These tests do not establish quality feasibility.

2.5.6 Protection and the Three Models

The grid trip of Section 2.3 reads the simulator's truth and is a teaching guardrail, not certified protection. The application keeps three models apart. The plant is the whole of Section 2.3. The identified model is whatever slope a learner fits to the falling interface of case 1, 7.29 mm/h per percent of valve in the plant. The controller's model is the optimizer's copy of the steady-state map with the commissioned parameters, which knows neither the slug's free water, nor the probe's offset, nor a parameter changed after the case started, and which the biases correct only on steady data.

2.6 Ten Cases, Measured

The ten cases appear in Table 2.3. All start from the defaults of Section 2.5 with 1.0 bar of mixing, 5 % of wash water, 5 ppm of demulsifier, the trim off, daily samples and the design crude, and depart from them only as their rows say. The numbers that follow were produced by stepping the engine of Section 1.6.1 headless at the default seed, in the order the browser steps it, and a setting named for a case is entered before the run unless a time is given. A case gives the same trajectory whether it is the first one opened, picked from the list after a different case, or restarted with Reset. The measured outcomes are collected in Table 2.4. The extremum-seeking case carries no number in the app, which lists it just before the challenge, and this chapter calls it case E. A requirement of a case is met only by a combination of something the learner did and a state the vessel reached (Section 1.6.4), and the chapter's exercises are built on the cases.

2.6.1 Proportional and Integral Action on the Layer

Case 2 ships a proportional-only LIC. The 100 mm setpoint step at 60 min settles with no offset, −0.2 mm over 200 to 240 min, because at the old brine flow the layer is already balanced. After the wash water rises from 5 to 7 % at 240 min, 10 m^3/h more water has to leave for good, the brine

Table 2.4 Measured Outcomes by Case, Default Seed

#	Setting	Outcome
1	as shipped	water −89.5 mm/h, 663 mm and 235 mg/L at 360 min
2	P-only	SP step −0.2 mm, wash step +55.3 mm offset
	τ_I 60, 30 min	offset −1.0 mm, overshoot 28, 39 mm
3	0.6, 1.0, 2.0 bar	1.69, 0.67, 0.89 PTB, BS&W up to 0.452 %
	optimizer	1.25 then 1.29 bar, 0.539 PTB, 0.179 %
4	as shipped	BS&W 0.526 % at 120, alarm 140.5 min, rag 619 mm, 2.72 A
	30 ppm at 120 min	rag 320 mm, 0.250 %, 0.89 PTB at 720 min
	trim, 2 h samples	60 ppm at 261 min, rag 261 mm, 0.197 %
5	wash 3, 5, 7, 8 %	1.48, 0.67, 0.36, 0.28 PTB, brine 15.5 to 40.5 m^3/h
	pump trip	5 PTB at 373 min, 10.0 PTB by 600 min
6	as shipped	alarm 100 min, trip 119.6 min at 3.01 A, gap 297 mm
7	as shipped	water 750 to 632 mm, oily at 192 min, bias 372 mm
	corrected, SP 1000	water 688 to 986 mm, trip 259.9 min
	corrected, SP 850	water 853 mm, 106 mg/L, peak 2.65 A
8	averaging (shipped)	brine 42.4 m^3/h, rag top 1613 mm, 2.45 A, 75 mg/L
	tight, 1.0/30 min	brine 60 m^3/h, rag top 1487 mm, 1.48 A, 99 mg/L

valve must stay 17.4 points further open (42.1 to 59.5 %), and the loop holds it there with an error of 55.3 mm at 0.3 % per millimeter, against 55.6 mm predicted from 10 m^3/h at 0.6 m^3/h per percent of valve and the gain. A 60 min integral time removes the offset to −1.0 mm over the last hour with 28 mm of overshoot on the setpoint step, and 30 min overshoots by 39 mm.

2.6.2 The Optimizer on the Salt Map

In case 3 the optimizer moves the setpoint from 1.00 to 1.25 and then to 1.29 bar and holds the salt at 0.539 PTB with the BS&W at 0.179 %, its own limit, after four hours. Its biases stay small, 0.002 on the logarithm of the BS&W and 0.030 on that of the salt, because its map is the plant's own correlation and only the instruments and its estimate of the rag top separate the two. Steps by hand to 0.6 and 2.0 bar settle at 1.69 PTB and at 0.89 PTB with 0.450 % of water.

2.6.3 A Heavier Crude, Wash Water and the Grid

The heavier crude of case 4 hardens the emulsion at once, and within an hour of the switch the BS&W rises from 0.104 to 0.526 % and the salt from 0.67 to 1.37 PTB. The rag grows more slowly, from 80 to 619 mm by 720 min. Because the probe reads half of it as water, the loop lets the true water fall to 940 mm and lifts the rag top to 1559 mm, 391 mm below the grid, and the current passes its 2.0 A alarm at 140.5 min and reaches 2.72 A. Demulsifier at 30 ppm from 120 min holds the rag at 320 mm, and the trim on 2 h samples, switched on at 120 min, takes the dose to its 60 ppm limit by 261 min and holds the rag at 261 mm with the BS&W at 0.197 %.

In case 5 the salt at 3, 5, 7 and 8 % of wash is 1.48, 0.67, 0.36 and 0.28 PTB while the BS&W stays between 0.104 and 0.115 %, and the brine rises from 15.5 to 40.5 m^3/h with its oil from 50 to 79 mg/L. After the wash pump trips at 360 min the salt passes 5 PTB at 373 min and reaches the raw 10.0 PTB by 600 min, a target no tuning can recover. In case 6, with the brine valve pinched to 8 % at 30 min, the current is 1.06 A at 30 min and only 1.21 A at 60 min while the rag top is far from the grid, then passes the alarm at 100 min and trips the grid at 119.6 min with 3.01 A and the rag top 297 mm below it. Opening the brine valve to 42.5 % at the alarm prevents the trip with a peak of 2.01 A, and so does a wide-open valve as late as 115 min, with 2.70 A.

2.6.4 The Rag Fools the Instrument

Case 7 starts on the heavier crude with a 500 mm rag and the interface setpoint at 1000 mm, so the probe reads 1000 mm with the water at 750 mm. The loop holds its reading for eight hours while the rag grows to 735 mm and the water falls to 632 mm. The brine passes 200 mg/L of oil at 192 min, by which time the water turns over in 1.34 h instead of 1.50 h at the start, and by 480 min the reading

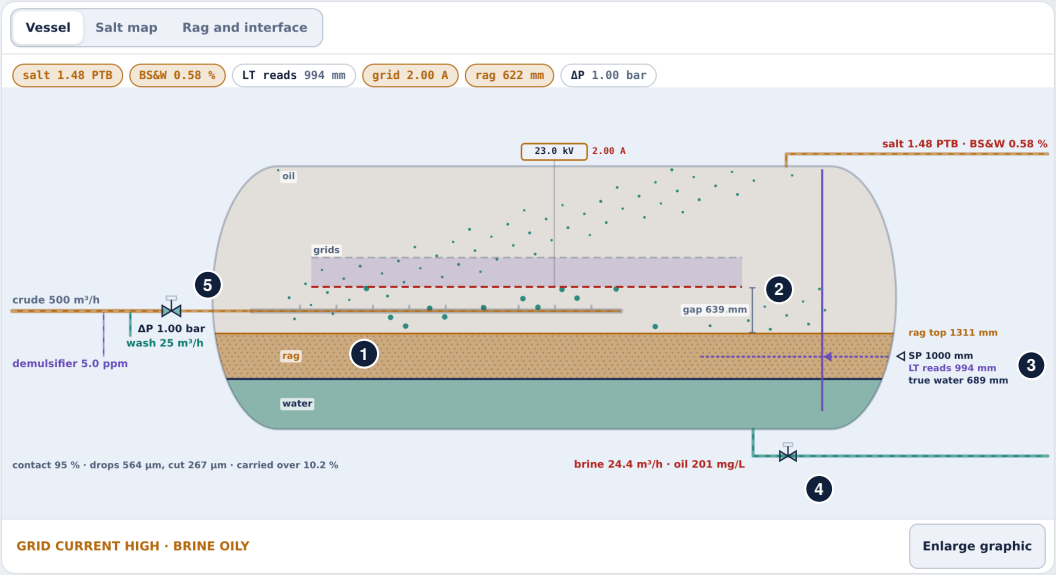


Figure 2.3 The rag misleading the interface instrument in case 7 at $t = 195$ min, Vessel view, as shipped, cropped to the plant view with numbered callouts. On the heavier crude the rag (1) has grown to 622 mm and its top stands at 1311 mm, 639 mm below the energized grid (2). The probe reads 994 mm (dotted) against its 1000 mm setpoint while the true water is at 689 mm (3). The brine valve (4) at 40.6 % passes 24.4 m³/h carrying 201 mg/L of oil, the grid draws 2.00 A at 23.0 kV, and the status strip reads GRID CURRENT HIGH and BRINE OILY. The mixing valve (5) holds 1.00 bar. Under the vessel, contact is 95 %, the drops leave the field at 564 µm against a 267 µm cut and 10.2 % of the contacted water is carried over, so the analyzer reports 1.48 PTB and the meter 0.579 % of water, with 85.3 % of the salt removed. In the full window, outside this crop, the trends over a two-hour window show the reading on setpoint with the true water and rag top drawn beside it, and the progress panel shows the first of the case's four completion steps met.

sits 372 mm above the water. The state at 195 min is shown in Figure 2.3. Correcting the reading at 200 min with the setpoint left at 1000 mm makes the loop close a bias of 282 mm, lifting the water from 688 to 986 mm and the rag with it until the grid trips at 259.9 min. Corrected with the setpoint at 850 mm, the water holds at 853 mm, the brine clears to 106 mg/L and the current peaks at 2.65 A. A thick rag leaves only a narrow band for the water between the draw and the grid. Corrected at 200 min with a setpoint of 900 mm, the water meets the case's one-hour hold at 301.8 min and the grid still trips at 464.3 min.

> ✦ **DID YOU KNOW? Correcting the Instrument Tripped the Grid**
>
> From the simulator (Table 2.4, case 7). With the reading corrected by profile and the setpoint left at 1000 mm, the true water rose from 688 to 986 mm and the grid tripped at 259.9 min. The same correction with the setpoint at 850 mm held the water at 853 mm with a peak of 2.65 A.

2.6.5 Averaging Against Tight Control Through a Slug

The slug of case 8 brings 50 m³/h of free water for 30 min, 25 m³ in all, onto a rag whose top stands 566 mm below the grid when it arrives. The brine valve passes at most 60 m³/h and normally runs at 26, so the layer must store part of the slug under any tuning. With the shipped averaging tuning, 0.1 % per millimeter and 120 min, the brine rises only to 42.4 m³/h, the rag top from 1384 to 1613 mm and

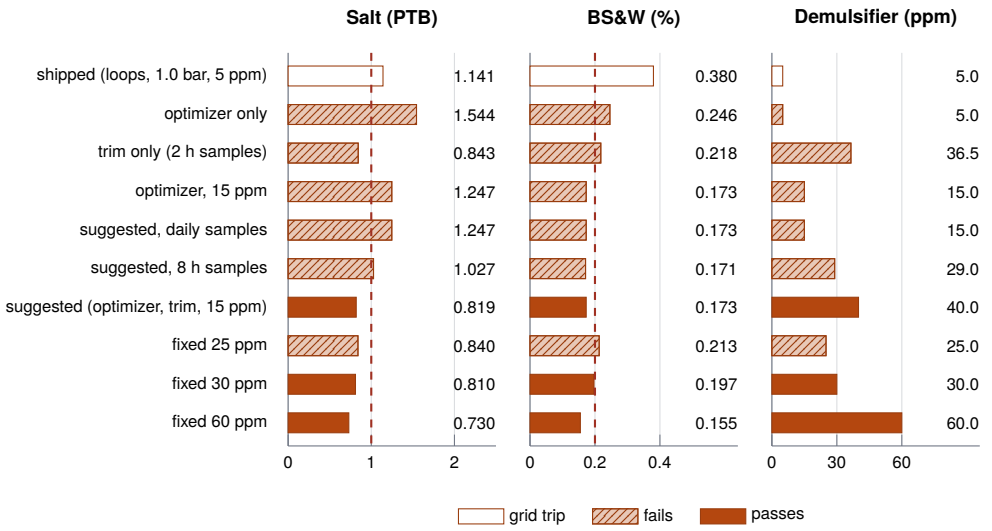


Figure 2.4 The ten challenge runs of Table 2.5 as bars for the average salt, the average BS&W and the average demulsifier over 12 h. The fill carries the verdict, solid for a pass, hatched for a miss and white for the run that tripped the grid, and the dashed lines are the targets of 1 PTB and 0.2 %. Notice that a run passes only with both of its first two bars left of the dashed lines, that the optimizer alone leaves both bars past their lines and the trim alone leaves one, and that the fixed 30 ppm passes with a shorter demulsifier bar than the suggested settings. The worst-sample column of the table is not drawn.

Table 2.5 The Challenge Over 12 h. Salt in PTB, Worst Analyzer Sample, BS&W Average and Demulsifier Average

Settings	Salt	Worst	BS&W (%)	ppm	Verdict
shipped (loops, 1.0 bar, 5 ppm)	1.141	1.471	0.380	5.0	trip 493.3 min
optimizer only	1.544	2.054	0.246	5.0	fails
trim only (2 h samples)	0.843	1.360	0.218	36.5	fails
optimizer, 15 ppm	1.247	1.783	0.173	15.0	fails
suggested, daily samples	1.247	1.783	0.173	15.0	fails
suggested, 8 h samples	1.027	1.744	0.171	29.0	fails
suggested (optimizer, trim, 15 ppm)	0.819	1.374	0.173	40.0	passes
fixed 25 ppm	0.840	1.428	0.213	25.0	fails
fixed 30 ppm	0.810	1.303	0.197	30.0	passes
fixed 60 ppm	0.730	0.985	0.155	60.0	passes

the current to 2.45 A, past its alarm. With tight tuning, 1.0 % per millimeter and 30 min, the valve goes wide open, the rag top peaks at 1487 mm and the current at 1.48 A, but the slug goes straight to the wastewater plant and the oil in the brine peaks at 99 mg/L instead of 75. An offshore separator makes the same trade between averaging and tight interface tuning in Volume 10, Chapter 6.

2.6.6 The Challenge and What It Does Not Price

The challenge runs 12 h with the heavier crude from 3 h and the slug at 8 h. Its target is the true salt at 1 PTB or less on average and 2 PTB or less at every analyzer sample, the BS&W at 0.2 % or less on average, the oil in the brine never above 200 mg/L, and no grid trip. The BS&W is judged as an average because even the suggested run exceeds 0.2 % for 125 min, with a peak of 0.306 %, during the switch and the slug. The variants appear in Table 2.5, and Figure 2.4 draws the same rows as bars.

As shipped, the unit runs the heavier crude on the design crude's settings, its BS&W averages 0.57 % from 240 to 480 min, its rag passes 400 mm by the time the slug arrives, and the grid trips at

Table 2.6 Paired Mean Quality and Time Above 0.2% Water Over 0–720 min, Seed 20261012

Case	Law	Salt (PTB)	BS&W (%)	$T_{B>0.2}$ (min)	ESC reports	ESC hold (min)	Trip (min)
Salt map	Fixed	0.6704	0.1041	0.0	0	0.0	—
	RTO	0.5425	0.1761	0.0	0	0.0	—
	ESC	0.6006	0.1317	0.0	48	0.0	—
Heavier crude	Fixed	1.3825	0.5290	646.0	0	0.0	—
	RTO	1.8305	0.2538	651.2	0	0.0	—
	ESC	1.4504	0.6137	648.8	7	609.7	715.8
Water slug	Fixed	0.6678	0.1082	0.0	0	0.0	—
	RTO	0.5704	0.1794	44.4	0	0.0	—
	ESC	0.6678	0.1082	0.0	0	720.0	—
Challenge, shipped chemistry	Fixed	4.6441	2.0285	526.0	0	0.0	493.3
	RTO	1.5441	0.2456	531.6	0	0.0	—
	ESC	4.6349	2.0332	526.9	18	451.2	493.2
Case E	Fixed	0.8304	0.1146	287.9	0	0.0	—
	RTO	0.6718	0.1375	62.8	0	0.0	—
	ESC	0.7842	0.1295	195.0	48	0.0	—

493.3 min, 13.3 min into the slug, with the rag top 394 mm below the grid. With the trip disabled the current would have reached 3.80 A. The suggested settings pass with 0.819 PTB on average, a worst sample of 1.374 PTB, 0.173 % of water, at most 94 mg/L of oil in the brine and a peak of 2.20 A. The optimizer holds 1.44 bar on the design crude at 15 ppm, and from 240 min on it keeps the setpoint between 0.69 and 0.94 bar. The trim adds nothing at the samples reported at 15 and 135 min, which find the rag at 57 and 49 mm, below its target. The sample drawn at 4 h and reported at 255 min finds 115 mm and takes the dose to 36.1 ppm, and the next, at 375 min, finds 181 mm and takes it to the 60 ppm limit. The rag then levels at no more than 237 mm, and the slug lifts its top to 1533 mm, 417 mm below the grid. The suggested settings pass at seeds 1 to 5 as well, with 0.778 to 0.852 PTB.

That passing run depends on the sample at 4 h. On daily samples the trim acts once, at 15 min, on a sample drawn before the switch, adds nothing, and the run reproduces the optimizer at a fixed 15 ppm to the digit. On 8 h samples the first useful sample comes too late and the run misses at 1.027 PTB. With the earlier version of the trim, both variants passed at the default seed, because the stored integral action raised the dose on the first sample, and they missed at most other seeds. The target also sets no chemical budget. Of the fixed ratios tried, 10, 15, 20, 25, 30, 40 and 60 ppm, the smallest that passes is 30 ppm, with 0.810 PTB and 0.197 % at every seed tried, while 25 ppm misses only on the BS&W, 0.213 %. The suggested settings average 40.0 ppm, a third more chemical than that fixed ratio, because the trim keeps adding while the heavier crude's rag stays above its target. The challenge therefore rewards holding the rag thin and the specifications met, and it does not reward the optimizer or the trim for saving chemical.

2.6.7 Measured Learning After a Crude Switch

Case E starts at 0.9 bar and 30 ppm, switches crude at 360 min, and runs to 720 min. Fixed pressure, the commissioned map optimizer (RTO) and ESC use the same chemistry, seed 20261012, instrument noise and events. The first ESC tuning candidate is frozen for every case. Other cases retain their shipped fixed pressure and dose. Each comparison reloads the case before selecting its law, retaining the same initial pressure and dose. Table 2.6 uses full-duration true outlet quality integrals, not analyzer averages. Post-trip field-off operation is included. The unchanged challenge verdict instead stops its accumulation at the trip, explaining its shipped 1.141 PTB in Table 2.5 versus 4.6441 PTB over the complete comparison window.

In case E, ESC reduces mean salt from fixed pressure's 0.8304 to 0.7842 PTB, but RTO achieves 0.6718 PTB. Time above 0.2 % BS&W is 287.9, 195.0 and 62.8 min, respectively, as shown in Figure 2.5. Mixing-valve travel is 131.30, 140.95 and 171.69 percentage points on the existing logged-sample metric. All three have zero valve saturation, current-alarm time and trips. With $k_g = 0$, probing alone gives 0.8461 PTB, so perturbation has a measurable cost. A hindsight fixed 1.2 bar

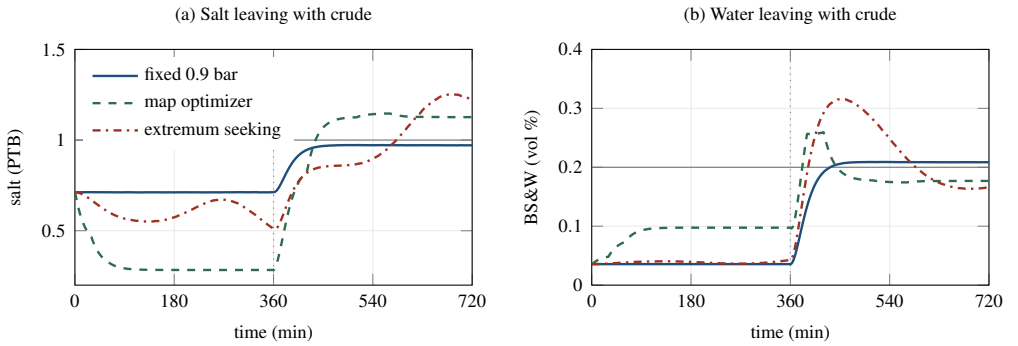


Figure 2.5 Outlet quality through the crude switch in case E. Both panels show simulator truth for evaluation, not ESC information. Fixed pressure is the solid line, the map optimizer dashed and extremum seeking dash-dotted. Thin gray horizontal lines mark the 1 PTB salt and 0.2% water specifications, and the dotted vertical line marks the 360 min switch. All three laws share the full 720 min window, 30 ppm dose and initial 0.9 bar.

gives 0.5978 PTB, but exceeds the water specification for 333.2 min. Neither this pressure nor extra chemical is credited as learned information. Seeds 1–3 retain ESC's salt loss to RTO. The center reaches 1.074 13 bar at the switch and 0.765 23 bar at the end, still moving 0.120 bar over the last 120 min. Two probing cycles establish no optimum-acquisition time. At 720 min, instantaneous ESC salt is 1.2223 PTB, versus fixed pressure at 0.9712 PTB and RTO at 1.1265 PTB.

ESC loses both quality measures to fixed pressure on the heavier-crude case, holds for 609.7 min, and trips at 715.8 min. The water-slug case's initial reported 350 mm rag holds it throughout. Matching the suggested challenge chemistry gives ESC 0.8126 PTB and 0.2762 %, versus RTO's 0.8191 PTB and 0.1731 %. ESC misses the water target while the suggested run with the optimizer still passes. The policies differ. ESC uses a soft water penalty and 0.3–2.5 bar projection, while RTO prioritizes feasible predictions on its 0.3–2.0 bar grid and knows the declared assay. The comparison therefore tests the implemented information and constraint policies rather than equal mathematical optimizations.

Independent state integration verifies all 15 primary rows. Halving the physics step preserves their salt rankings, with a largest change of 0.001 12 PTB. Replaying 7201 controller packets through independent filter recurrences differs by at most 1.78×10^{-15}. A quadratic fixture checks finite-filter demodulation to 4.44×10^{-16}, not plant-gradient accuracy or convergence. Computation includes every ESC call and history update. One sample in the Chrome browser at 1 min/s measures 9.091 µs per call over 22 calls and one fresh report, using 0.009 08 % of a core.

2.7 Technology in Practice

The model of Section 2.3 gives each loop a named instrument and an ideal valve. Figure 2.6 draws the same loops with the hardware a refinery installs in their place.

What Measures It

The interface is read by a capacitance probe through the top of the vessel. The probe answers to the electrical character of whatever wets it, between a value for oil and a value for water, and with a rag present it reports an approximate, averaged position [21]. That averaged position is the share κ of Equation (2.11), and the lag τ_{LT} and noise σ_I of the same equation are the model's whole account of the transmitter. The mixing pressure drop is read by a differential pressure transmitter with one line to each side of the valve.

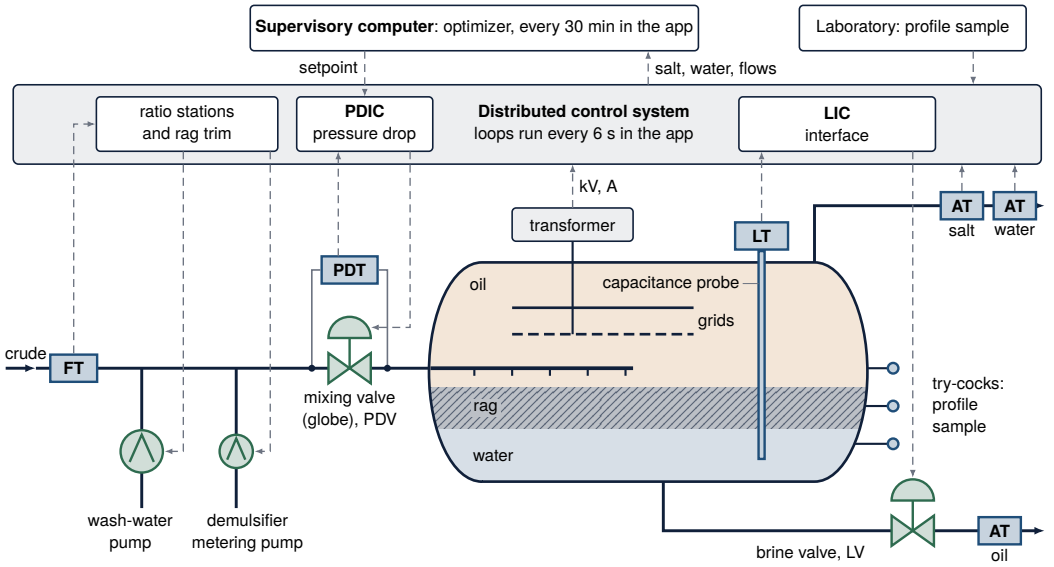


Figure 2.6 The loops of the desalter with their hardware. Lettered boxes are sensing devices, namely the flow transmitter (FT), the differential pressure transmitter (PDT), the level transmitter (LT) on the capacitance probe and the analyzers (AT). Valve and pump symbols are final elements. Dashed arrows are signals. The regulatory loops run in the distributed control system, and the optimizer above it writes only the pressure-drop setpoint. Notice that the level transmitter is the single path from the vessel to the brine valve, so its error passes to the water unseen. Not to scale, and not a piping or safety diagram.

> **TECHNOLOGY** **Sensor: Capacitance Interface Probe**
>
> A capacitance probe measures the electrical character of the fluid that touches it. As the interface rises along the rod, the measured capacitance moves from its value in oil, 0% of the range, to its value in water, 100% [21]. The contrast is large, a relative permittivity of about 2.1 for oil against 80 for water [22]. A probe with a single electrode needs temperature compensation and cannot locate more than one interface [22]. It fails in three ways. A new feedstock or a rag changes the electrical properties, and the change reads as a move of the level. A coating of solids can disable the rod. Any repair needs a shutdown, because the rod sits in the process [21]. On a trend the first failure is a steady reading over a moving interface, as in case 7.

What Moves It

The mixing valve is commonly a globe valve, and spray nozzles or static mixers may premix the water ahead of it [2]. The model's brine valve is linear in flow and its mixing valve follows the square law of Equation (2.10), each behind a short lag. The wash water and the demulsifier are moved by pumps, which the ratio stations set from the metered crude flow.

> **TECHNOLOGY** **Final Element: Globe Valve With a Pneumatic Actuator**
>
> A globe valve regulates flow with a movable plug and a fixed ring seat, and the shape of the plug sets how flow follows lift, as quick opening, linear or equal percentage [23]. The usual actuator is pneumatic. Air in a sealed flexible chamber works against a pre-compressed spring and moves the stem. The spring also decides whether the valve fails open or closed when the air is lost, and that choice is made by asking which position is safer for the process [23]. A pneumatic actuator answers with a delay [23], which the model carries as a lag of 0.1 min on each valve (Table 2.2). The mixing valve cannot be throttled far. At 15 % open its drop at design flow is already 11.2 bar, and the manual station stops there.

Where the Controller Runs

A refinery runs loops of this kind in a distributed control system. In such a system a supervisory level sends set points to distributed controllers, and these act on the process through its actuators from the feedback of its sensors [24]. Loops at the two levels cycle in times from milliseconds to minutes [24]. The app executes its controllers every 6 s and its optimizer every 30 min, inside that range. The interface controller, the pressure-drop controller, the ratio stations and the rag trim would sit in the distributed controllers. The optimizer and extremum seeking would sit in the supervisory level, since each only writes the pressure-drop setpoint.

> **IN PRACTICE** **The Salt Number, the Heat and the Interface**
>
> The salt figure an operator works to comes from a laboratory test, ASTM D3230. It infers chlorides from the electrical conductivity of the crude over 1.0 to 150 PTB, and it is used to judge both whether a crude needs desalting and how well the desalter works [25]. In service the crude is heated to between 150 and 350 °F, as far as its vapor pressure allows [3]. At the interface, an industry question-and-answer panel recorded in 2016 that capacitance probes and floats give a general position but little about the width of the interface or its growth, and that knowing both lets operators adjust the demulsifier ahead of an upset [26].

What Is Modernizing It

Three developments bear on this chapter. The first is the profile instrument, which replaces one reading with a measurement at many heights. One kind is a vertical array of low-energy gamma sources facing arrays of detectors, which gives the density at each height. Another is a capacitance sensor of 64 segmented electrodes [22]. A profile read in real time [21] would show the bias κt_r of Section 2.4 continuously, where the simulated unit sees it only at each profile sample. The second is the electrode. Recent electrocoalescers all use insulated electrodes to prevent the short circuit that the grid trip guards against [8]. The third is the network. In many modern systems the control system is interfaced with the corporate network to give the business a view of production [24], which makes the security of the control network part of the design of every loop that runs on it.

2.8 Limitations

The sizes and correlations describe a typical single-stage desalter on a 500 m^3/h crude unit, chosen so that the design point closes, and none was fitted to a vessel. The separation, rag, grid and oil-in-brine laws are synthetic, so where the salt optimum lies, how thick the rag grows and at what gap the grid trips are reasonable sizes for teaching and not forecasts for a particular vessel, crude or chemical. The model has one oil zone, one rag and one water layer, a fixed temperature of 130 °C, no second stage and no handle for temperature, so a truly heavy crude cannot be brought to specification here. Given the heavier assay's solids and asphaltenes at API 22.3 (919 kg/m^3 and 2.9 mPa s) and dosed at 60 ppm, no pressure drop brings the BS&W to 0.2 % and the least salt anywhere on the map is 1.40 PTB, where a refiner would raise the desalter's temperature. The demulsifier never overdoses, and the probe's rag share $\kappa = 0.5$ is a constant where a real probe's response depends on the emulsion it sits in.

> **OPEN CHALLENGE** **Seeing Both Faces of the Rag**
>
> A 2006 review called the accurate measurement of several interfaces in one vessel a challenge, and reported that for lack of it separators were run by trial and error [22]. Each established instrument misses the emulsion in its own way. Displacers and differential pressure cannot detect it, ultrasound has difficulty with it, and a nucleonic gauge finds the midpoint of a diffuse band but not its depth [21, 22]. Ten years later some refiners had moved to nuclear gauges on desalters and were still evaluating them [26]. Progress would be a profile reliable enough for a loop to hold the true water and the rag top apart, which case 7 does by hand from a two-hour sample.

The optimizer's model is the plant's own correlation with the commissioned parameters, so it meets model error only through unmeasured free water, the probe and a changed parameter, and its biases learn only on steady data. The trim acts on samples alone, and the challenge prices neither chemical nor wastewater, so a fixed 30 ppm meets the target that the suggested architecture was designed for. The grid trip reads the truth and freezes the run rather than model a restart, and the disturbances arrive as scheduled steps and pulses.

ESC's sampled washout and low-pass filters do not remove process lag, crude-change transients or clipping bias. The known analysis delay is aligned by timestamp, but no dynamic phase compensation is fitted. Protective holds can consume all 720 min, and bounded commands do not guarantee a feasible product or prevent a later trip. The measured objective prices water excess in PTB equivalents rather than dollars.

2.9 Conclusion

A crude desalter is a level loop under a hazard, wrapped around a mixing valve whose best setting moves with every crude, and this simulator computes the salt map, the rag, the probe's reading and the grid current together from a twenty-state model whose ledgers close to round-off. On the design crude the salt is least at 1.29 bar, 0.54 PTB, the BS&W reaches its specification at 1.36 bar and has a minimum of its own at 0.57 bar, and on the heavier crude at 5 ppm no pressure drop meets the specification, so demulsifier decides which curve exists and the mixing valve where to sit on it. A probe that reads half the rag lets a loop on its setpoint drop the water from 750 to 632 mm while the brine turns oily, and correcting the reading without lowering the setpoint lifts the rag into the grid. The challenge passes with the optimizer, a rag trim acting from the sample drawn at 4 h and 15 ppm, at 0.819 PTB, and equally with a fixed 30 ppm, because it sets no chemical budget. Model-free extremum seeking lowers the mean salt of case E from fixed pressure's 0.8304 to 0.7842 PTB, while the map optimizer remains stronger at 0.6718 PTB, and two probing cycles establish no time to find the optimum. The correlations are synthetic and the temperature fixed. A fuller model would add desalter temperature as a handle, a second stage with recycled water, a demulsifier that can overdose and a chemical budget in the challenge. The simulator shows that a loop on its setpoint is only as right as the instrument it reads.

KEY POINTS

- Mixing harder washes out more salt only up to a point. Past it the water drops are too small to settle, so the best setting lies in between and moves with every crude.
- The demulsifier decides which salt curve exists, and the mixing valve where to sit on it.
- A probe that reads half the rag lets a loop on its setpoint lower the true water while the brine turns oily.
- Correcting the reading without lowering the setpoint lifts the rag into the grid.
- A loop on its setpoint is only as right as the instrument it reads.

Where else this applies. The same kind of interior optimum is held in Chapter 13, the trade between averaging and tight interface tuning returns in Volume 10, Chapter 6, and a sensor that misleads its loop in Volume 5, Chapter 12.

References

[1] M. R. Gray, P. E. Eaton, and T. Le, "Kinetics of hydrolysis of chloride salts in model crude oil," *Petroleum Science and Technology*, vol. 26, no. 16, pp. 1924–1933, 2008.

[2] J. Pereira, I. Velasquez, R. Blanco, M. Sanchez, C. Pernalete, and C. Canelón, "Crude oil desalting process," in *Advances in Petrochemicals*, V. Patel, Ed. InTech, 2015.

[3] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[4] U.S. Energy Information Administration, “Refinery capacity report 2026, table 1: Number and capacity of operable petroleum refineries by PAD district and state as of January 1, 2026,” U.S. Department of Energy, Washington, DC, Tech. Rep., 2026.

[5] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson, “RTO: An overview and assessment of current practice,” *Journal of Process Control*, vol. 21, no. 6, pp. 874–884, 2011.

[6] S. Skogestad, “Simple analytic rules for model reduction and PID controller tuning,” *Journal of Process Control*, vol. 13, no. 4, pp. 291–309, 2003.

[7] D. E. Seborg, T. F. Edgar, D. A. Mellichamp, and F. J. Doyle, III, *Process Dynamics and Control*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2016.

[8] S. Mhatre, V. Vivacqua, M. Ghadiri, A. M. Abdullah, M. J. Al-Marri, A. Hassanpour, B. Hewakandamby, B. Azzopardi, and B. Kermani, “Electrostatic phase separation: A review,” *Chemical Engineering Research and Design*, vol. 96, pp. 177–195, 2015.

[9] S. Less and R. Vilagines, “The electrocoalescers’ technology: Advances, strengths and limitations for crude oil separation,” *Journal of Petroleum Science and Engineering*, vol. 81, pp. 57–63, 2012.

[10] E. Aryafard, M. Farsi, and M. R. Rahimpour, “Modeling and simulation of crude oil desalting in an industrial plant considering mixing valve and electrostatic drum,” *Chemical Engineering and Processing: Process Intensification*, vol. 95, pp. 383–389, 2015.

[11] M. Al-Otaibi, A. Elkamel, T. Al-Sahhaf, and A. S. Ahmed, “Experimental investigation of crude oil desalting and dehydration,” *Chemical Engineering Communications*, vol. 190, no. 1, pp. 65–82, 2003.

[12] F. G. Cottrell and J. B. Speed, “Separating and collecting particles of one liquid suspended in another liquid,” U.S. Patent 987 115, Mar. 21, 1911.

[13] W. S. Barnickel, “Process for treating crude oil,” U.S. Patent 1 093 098, Apr. 14, 1914.

[14] G. Egloff, E. F. Nelson, C. D. Maxutov, and C. Wirth, “Desalting crude oils,” *Transactions of the AIME*, vol. 127, no. 1, pp. 48–67, 1938.

[15] J. D. McLean and P. K. Kilpatrick, “Effects of asphaltene solvency on stability of water-in-crude-oil emulsions,” *Journal of Colloid and Interface Science*, vol. 189, no. 2, pp. 242–253, 1997.

[16] H. D. Beggs and J. R. Robinson, “Estimating the viscosity of crude oil systems,” *Journal of Petroleum Technology*, vol. 27, no. 9, pp. 1140–1141, 1975.

[17] J. O. Hinze, “Fundamentals of the hydrodynamic mechanism of splitting in dispersion processes,” *AIChE Journal*, vol. 1, no. 3, pp. 289–295, 1955.

[18] J. F. Richardson and W. N. Zaki, “Sedimentation and fluidisation: Part I,” *Chemical Engineering Research and Design*, vol. 75, pp. S82–S100, 1997.

[19] M. Krstić and H.-H. Wang, “Stability of extremum seeking feedback for general nonlinear dynamic systems,” *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.

[20] M. Guay and D. Dochain, “A time-varying extremum-seeking control approach,” *Automatica*, vol. 51, pp. 356–363, 2015.

[21] P. Hewitt, “Oil/water interface control for desalters,” *Petroleum Technology Quarterly*, Oct. 2007, article 1001472 at digitalrefining.com.

[22] S. F. A. Bukhari and W. Yang, “Multi-interface level sensors and new development in monitoring and control of oil separators,” *Sensors*, vol. 6, no. 4, pp. 380–389, 2006.

[23] D. Katzman, J. Moreno, J. Noelanders, and M. Winston-Galant, “Valves: Types, kinds, and selection,” in *Chemical Process Dynamics and Controls*, P. Woolf, Ed. University of Michigan open textbook, LibreTexts, 2023, ch. 3.9.

[24] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, “Guide to operational technology (OT) security,” National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[25] ASTM International, “Standard test method for salts in crude oil (electrometric method),” ASTM D3230-19, West Conshohocken, PA, 2019.

[26] American Fuel & Petrochemical Manufacturers, “Question 57: Desalter level control equipment and its effectiveness at detecting and controlling rag layer, oil undercarry and water carryover,” Question and answer record, Washington, DC, 2016, afpm.org, node 42221.

2.10 Exercises

2.1 Concept check · 2 min

On the design crude at 5 ppm of demulsifier, the salt left in the desalted crude falls from 1.69 PTB at 0.6 bar of mixing pressure drop to 0.54 PTB at 1.29 bar and rises again to 0.89 PTB at 2.0 bar. What makes the salt rise again above the optimum?

(a) The analyzer's delay makes the reading lag the mixing, so the rise is an artifact of sampling

(b) Harder mixing contacts more brine but shears the water into drops too small for the field to recover

(c) Harder mixing heats the crude by valve friction and thins the oil past the best viscosity

(d) Harder mixing throttles the wash water pump, so less wash water reaches the crude's brine

2.2 Concept check · 2 min

In case 7 the interface loop holds the probe reading on its 1000 mm setpoint while the rag grows by 100 mm. The probe counts half of the rag as water ($\kappa = 0.5$) and has no offset. What happens in the vessel?

(a) The true water rises 50 mm and the rag top rises 150 mm

(b) The true water stays put and the rag top rises 100 mm, because the loop holds the water

(c) The true water falls 50 mm and the rag top rises 50 mm, with no error on the loop

(d) The true water falls 100 mm and the rag top stays put, because the loop holds the rag top

2.3 Concept check · 2 min

Three hours after the switch to the heavier crude at 5 ppm, the least BS&W anywhere on the salt map is 0.216 %, at 0.40 bar. Which change can bring the desalted crude within the 0.2 % specification?

(a) Lowering the mixing pressure drop to the 0.3 bar end of its range

(b) Raising the mixing pressure drop to 1.5 bar to contact more of the brine

(c) Selecting Loops + optimizer with its BS&W limit at 0.18 %

(d) Raising the demulsifier ratio from 5 to 30 ppm and waiting for it to act

2.4 Theory and calculation · 7 min

The vessel is a horizontal cylinder 3.6 m across and 24 m long, the interface stands near 1.25 m, and one percent of brine valve passes 0.6 m^3/h. The crude charge is 500 m^3/h.

(a) Find the width of the interface, its area and the integrating gain of the water layer in millimeters per hour per percent of valve. *Checkpoint:* 3.43 m, 82.3 m^2 and 7.29 mm/h per percent.

(b) With a proportional-only interface loop at $K_c = 0.3$ % per millimeter, find the closed-loop time constant, and the integral time at which a PI loop with this gain is critically damped. *Checkpoint:* 27.4 min and 110 min.

(c) The wash water rises from 5 to 7 % of the crude. Find how far the brine valve must open for good and the offset the proportional-only loop holds. *Checkpoint:* 16.7 points and 55.6 mm.

2.5 Theory and calculation · 6 min

In case 7 the probe reads $h_I = h_w + \kappa t_r$ with $\kappa = 0.5$ and no offset, the interface setpoint is 1000 mm and the lower grid is at $h_g = 1950$ mm. The rag starts 500 mm thick.

(a) With the reading on its setpoint, find the true water, the rag top and the gap from the rag top to the lower grid. *Checkpoint:* 750 mm, 1250 mm and 700 mm.

(b) The rag grows to 735 mm with the reading still on its setpoint. Find the same three heights. *Checkpoint:* 632.5 mm, 1367.5 mm and 582.5 mm.

(c) A profile sample finds the reading 282 mm above the water, and the reading is corrected by it with the setpoint left at 1000 mm. Find the rag thickness the bias implies and the rag top once the loop holds the water itself at 1000 mm. What setpoint keeps the rag top where it was before the correction? *Checkpoint:* 564 mm, 1564 mm (a 386 mm gap) and about 718 mm.

2.6 Theory and calculation · 7 min

The transformer drives the lower grid with an open-circuit tap $V_{tap} = 27\,\mathrm{kV}$ through a reactor $R_s = 2.0\,\mathrm{kV/A}$, and $V = V_{tap}/(1 + R_s G)$, $I = GV$ (Equation (2.8) with full supply). On the dry design crude the conductance is $G = 0.040 + 2.0\,\mathrm{e}^{-(h_g - h_t)/\lambda}$ in amperes per kilovolt with $\lambda = 100\,\mathrm{mm}$.

(a) At design the rag top is 660 mm below the grid. Find G, the grid voltage and the current. *Checkpoint:* 0.0427 A/kV, 24.87 kV and 1.063 A.

(b) Find the gap at which the current reaches the 3.0 A trip. *Checkpoint:* $G = 0.1429\,\mathrm{A/kV}$, a gap of 297 mm.

(c) Find the voltage and the current with water at the grid (a gap of zero), and compare the current with the reactor's limit. *Checkpoint:* 5.31 kV, 10.84 A against 13.5 A.

2.7 App challenge · 9 min

Open case 2, "The interface loop on an integrating level", at the default seed, and run each part to 480 min at 60 min per second (about 8 s a run). Export the CSV file after each run. The setpoint steps from 1250 to 1350 mm at 60 min, and the wash water from 5 to 7 % at 240 min.

(a) Run as shipped, proportional only. Average `interface_meas_m` minus `interface_sp_m` over 200 to 240 min and over 420 to 480 min, and average `brine_valve_cmd_pct` over the same two windows. *Checkpoint:* −0.2 mm and 55.3 mm (within 2 mm), valve 42.4 and 59.2 % (within 0.3 points).

(b) Reload the case, set Interface integral time to 60 min before Run, and run again. Give the average error over the last hour and the overshoot of the reading above 1350 mm between 60 and 240 min. *Checkpoint:* −1.0 mm and 28 mm (within 3 mm).

(c) Repeat with 30 min. Give the overshoot, and say why the setpoint step left no offset in part (a) while the wash step did. *Checkpoint:* 39 mm (within 3 mm).

2.8 App challenge · 9 min

Open case 7, "The rag fools the interface instrument", at the default seed and the shipped settings (heavier crude, a 500 mm rag, interface setpoint 1000 mm, profile samples every 2 h). Export the CSV file after each run.

(a) Run to 480 min at 20 min per second (24 s). In the CSV file, read the true water (`truth_water_interface_m`) at 0 and 480 min, and the rag (`truth_rag_m`) and the reading (`interface_meas_m`) at 480 min. Find the first row in which the oil in the brine (`oil_in_brine_mgL`) passes 200 mg/L. *Checkpoint:* Water 750 to 632 mm, rag 735 mm, reading within the probe's noise of 1000 mm, oil in brine past 200 mg/L at 192 min.

(b) Reload the case, run at 5 min per second, and pause between 195 and 210 min. Set Interface reading to Corrected by profile, leave the setpoint at 1000 mm, and run on at 20 min per second. Read `interface_correction_m` just after the switch, and the time the grid trips. *Checkpoint:* A bias of 282 mm, and a trip between 255 and 268 min.

(c) Reload, pause in the same window, correct the reading and set the setpoint to 850 mm. Run to 480 min and read the true water, the oil in the brine and the peak grid current. *Checkpoint:* Water 852 mm (within 3 mm), 106 mg/L, peak 2.65 A and no trip.

2.9 App challenge · 8 min

Open the challenge, case 9, at the default seed. Each run lasts 12 h of plant time, about 2.4 s at 300 min per second, and stops at 720 min with its verdict.

(a) Run as shipped. When does the grid trip? *Checkpoint:* At 493.3 min, during the slug.

(b) Reload, set the Demulsifier ratio to 25 ppm before Run, and run. Read the verdict's average salt and average BS&W. *Checkpoint:* 0.84 PTB and 0.213 %: missed on the BS&W alone.

(c) Reload and repeat at 30 ppm. Then reload, reveal the suggested settings (optimizer, rag trim, 2 h samples, 15 ppm) and run. Compare the two passing runs on the average of `demulsifier_ppm` in the CSV file. *Checkpoint:* 0.81 PTB and 0.197 % at 30.0 ppm, 0.82 PTB and 0.173 % at 40.0 ppm.

2.10 Application area · 3 min

The raw crude of this chapter carries 10 PTB of salt. A current account of the process puts the refinery's target near 1.5 PTB and quotes about 90 % removal for one stage and 99 % for two.

(a) Find the salt left by one stage and by two at those removals, and compare them with the simulated unit's 93.3 %. *Checkpoint:* 1.0 PTB, 0.1 PTB and 0.67 PTB.

(b) Name two ways in which salt left in the crude harms the units downstream. *Checkpoint:* Hydrochloric acid in the overhead, and scale or poisoned catalyst.

2.11 Application area · 3 min

A slug of 50 m^3/h of free water arrives for 30 min, as in case 8. The brine valve passes at most 60 m^3/h and normally runs near 26 m^3/h.

(a) At least how much of the slug must the layer store, even with the valve wide open at once? *Checkpoint:* About 8 m^3.

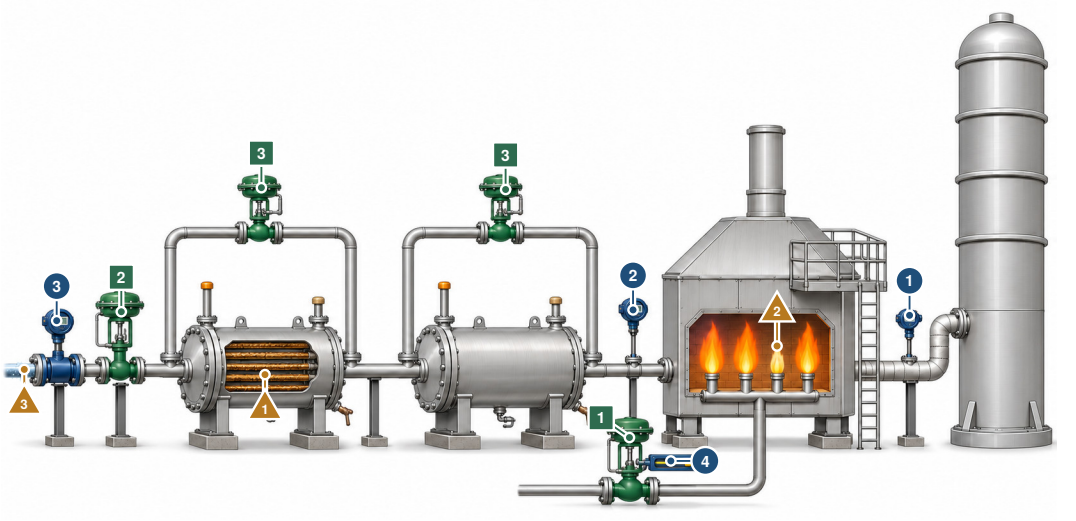
(b) The averaging tuning sends at most 42.4 m^3/h of brine with 75 mg/L of oil and lets the current reach 2.45 A. The tight tuning sends 60 m^3/h with 99 mg/L and holds the current at 1.48 A. Which would you choose when the grid alarm matters more than the wastewater plant, and what does it cost? *Checkpoint:* The tight tuning, at the cost of the brine surge.

2.12 Review the agent · 5 min

An AI assistant reviewing case E writes: "Extremum seeking needs no model, so it cannot be misled by a wrong salt map. After the crude switch at 360 min it will find the new optimum faster than the map optimizer, and it should replace the optimizer in the challenge." Find the flaw and say how you would check the claim with the app.

CHAPTER 3

Fouling Crude Preheat Train Simulator for Override Control on Throughput



The controller moves 1 fuel gas valve of the furnace 2 crude charge valve 3 cleaning bypass of each exchanger

The controller reads 1 column inlet temperature 2 furnace inlet temperature 3 crude charge flow 4 fuel valve position

What disturbs it 1 fouling inside the tubes 2 leaner fuel gas 3 colder crude from the tanks

Once fouling has driven the fuel gas valve under the furnace wide open, only the charge valve at the far left can still hold the temperature read at the column inlet on the right. Illustration, not to scale.

IN PLAIN TERMS Slowing the Plant to Hold a Temperature

The main idea. Crude oil must be hot before a refinery can split it into fuels. Most of the heat comes free from hot products in a row of heat exchangers, and a gas furnace adds the rest.

The challenge. Dirt builds up inside the exchangers, so the furnace burns more gas until its fuel valve is fully open. Then the only way to keep the oil hot is to send less oil through, like a driver who slows on a steep hill.

Why it matters. Every tonne of oil held back is product the refinery does not make, so the cut must be kept small.

THREE WAYS TO READ THIS CHAPTER

Explore	9 min	This page, the Did you know? boxes in Sections 3.1 and 3.5.3, the History boxes in Sections 3.2 and 3.4, the Try it box in Section 3.5.2, and Key points in Section 3.9. No equations.
Study	65 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Sections 3.5.1 to 3.5.3 with Figure 3.3, Section 3.5.6, Section 3.6, Section 3.7 and the Open challenge box in Section 3.8.

The control question Fouling and cleaning outages take the fuel valve to its stop. What does holding the column inlet then cost in throughput, and how should an override that cuts the charge be set?

Learning objectives

- Explain why a loop held on setpoint through weeks of fouling hides the lost margin.
- Compute the fouled coefficient and the time to the cleaning threshold at two charge rates.
- Compute the heat-balance firing and the duty a cleaning outage removes.
- Choose a valve position target from both temperature and crude processed.
- Explain why overlapping cleanings cost far more than separate ones.

Prerequisites Sections 1.4.2, 1.4.5, 1.6.2, 1.6.3 and 1.6.4.

The app https://apmonitor.com/apps/crude-preheat/

See also Volume 4, Chapter 9 for valve position control, Volume 4, Chapter 4 for throughput as the override's handle, and Volume 3, Chapter 4 for a heater's feedforward and limits.

3.1 Introduction

The usual evidence that a control loop is working is a measurement sitting on its setpoint. On the unit of this chapter that evidence is incomplete. Three weeks of exchanger fouling leave the column inlet temperature unchanged, hold the integrated absolute error at 0.0 °C h over 504 h, and raise no alarm, while the fuel gas valve opens from 81.3 % to 97.5 % and the share of the crude duty recovered from product streams falls from 60.1 % to 52.1 %. Three weeks of fuel gas costs 76.7 TJ, about $460k at the $6/GJ carried in the parameter list, and roughly a tenth of that is the fouling. A performance report limited to control error would rate the three weeks as flawless. The only variable that records what happened is the position of the valve, which is why valve position is a monitored quantity on every constrained unit in a refinery.

> **DID YOU KNOW? The Largest Energy User in the Refinery**
>
> A 2005 guide for refinery energy managers calls crude distillation the largest energy consuming process of a refinery, and its preheater especially susceptible to fouling. It quotes a federal estimate that the cost penalty for fouling could be as much as $2 billion a year in material and energy [1].

The arrangement ends when the valve reaches its limit. When one exchanger is taken out of service for a sixteen hour cleaning at full charge rate, the fired heater saturates, sits on its 40.6 MW limit for 96.5 % of the outage, and delivers 340.3 °C to a column with a 360 °C setpoint. No feedback tuning recovers duty the furnace cannot supply. One manipulated variable remains, the crude charge rate. A lower charge rate requires less furnace duty and gains more temperature rise per tonne across the same 1620 m^2 of surface, so a small rate reduction produces a large temperature recovery. Reducing the charge rate also reduces refinery throughput directly, and that trade is the subject of this chapter. A valve position controller allowed to write a charge rate holds 360.0 °C through the identical outage at a cost of 555 t of crude out of 19 200 t.

Two general classes of problem sit inside this one plant. The first is override control, in which a second controller on a second manipulated variable is selected only when the primary actuator runs out of authority. Valve position control is its most common industrial form, and the same structure holds the most-open valve in a parallel network near full open, keeps a pump on its curve, and protects a compressor from surge [2, 3]. Volume 4, Chapter 2 treats the surge case, and Volume 4, Chapter 4 uses throughput as the override's handle, as here, with the ground speed of a harvester cut against engine and rotor limits. Constraint switching of exactly this kind has been given a systematic design procedure in terms of the classical structures that implement it [4], and the multivariable version of the same decision, in which throughput is pushed against whichever unit is currently the

bottleneck, is the standard justification for a plantwide coordinating controller [5]. The second class is the slowly moving process gain with a maintenance outage attached. Reverse osmosis membranes, boiler economizers, air cooler finning, pasteurizers and deactivating catalyst all degrade over weeks and are all managed by a schedule rather than by a controller, and the engineering question in every case is when to stop and clean.

However, a first-order transfer function identified from a step test represents neither class. Fouling does not appear in a transfer function because the loop absorbs it, and a firing limit does not appear because a linear model has no saturation. A trend sitting on setpoint establishes only that one loop still has actuator authority in reserve. This chapter works with a crude preheat train and its fired heater, described in enough detail to be rebuilt. It quantifies the throughput cost of the override when it is selected, measures what the valve position target buys on either side of the point where the constraint binds, and shows why two cleaning cycles of unequal period produce their largest excursion when they coincide. The analysis rests on two measured outcomes. Holding temperature through a cleaning outage costs 555 t of crude for the upstream exchanger and about three times that for the downstream one. Moving one of two colliding cleanings 32 h later divides the integrated error by almost twelve while changing the crude processed by 17 t in 199 000. The plant has twelve states on two time scales separated by a factor of five thousand, four control architectures and fourteen scenarios, each with its own completion requirement. The performance report keeps temperature error and throughput separate so that a trade between them is visible. The fouling law, the exchanger model and the override arithmetic are all established practice.

The control ideas are override and valve position control, feedforward design from a steady-state energy balance, actuator saturation and anti-windup, and the time-scale separation between a regulatory loop and a slowly varying process gain, all of which are standard material in process dynamics and control texts [6, 7]. Section 1.4.5 introduces the same three ideas on the running example R-1: the preheat train, the fouling that the outlet loop hides and the valve position override of Equation (1.10). The model itself rests on energy conservation over lumped thermal capacities, a log-mean temperature difference driving force, and a fouling resistance in series with the clean film resistances, and the closed-loop behavior follows from the feedback fundamentals collected in [8]. The behavior matters industrially because a crude unit operates at its most profitable point only when throughput is pushed against the currently active constraint [5], and the engineer who cannot distinguish a capacity limit from a tuning problem retunes a loop that has nothing left to give.

Section 3.2 reviews how a train of recovery exchangers and a cleaning schedule became the front end of every crude unit. Section 3.3 gives the twelve-state model, with fouling as a series resistance, the exchanger and furnace balances, the cleaning event, the parameters and the assumptions. Section 3.4 sets out the temperature loop, the heat-balance feedforward, the valve position override and the predictive option. Section 3.5 measures three weeks of fouling (Section 3.5.1), the cleaning outages (Section 3.5.2), two colliding cleaning cycles (Section 3.5.3), three disturbances that resemble fouling (Section 3.5.4), the graded challenge (Section 3.5.5) and the predictive controller against the override (Section 3.5.6). Section 3.6 sets the loops beside the hardware of a real unit, and Sections 3.7 to 3.9 state the limitations, discuss the scheduling argument and conclude.

3.2 Background

The preheat train exists because fired heater duty is expensive. Crude arriving at tank temperature has to reach the flash zone of the atmospheric column, and every degree supplied in a fired heater is fuel gas burned at a thermal efficiency below unity, while every degree recovered from a hot product stream leaving the same column carries no fuel cost. Refinery energy accounting puts crude distillation among the largest single consumers on the site, and the fired heaters that serve it among the largest emitters [9]. The engineering answer was to place as much surface as the economics allow between tankage and the heater, and the question of how much surface, where, and against which hot stream became a design discipline when heat exchanger network synthesis was posed as a problem

with a thermodynamic target rather than as a matter of judgment [10]. A modern crude unit therefore carries twenty or more shells in series and parallel, with a recovery fraction that falls below the design value in operation.

The shortfall is caused by fouling, and the History box below tells how fouling became a rate equation, a deposition term less a removal term, with an asymptote where the two balance. That model, its limitations and the fifty years of work that followed it are reviewed by Müller-Steinhagen [11], and the mechanisms, velocity dependence and cleaning practice are collected by Bott [12]. The application uses that law directly, because it is the smallest description that reproduces the two behaviors an operator recognizes, a fouling rate that is fastest at the start of a run and a resistance that eventually stops growing.

> ⧖ HISTORY **1959 and 1995: Fouling Becomes a Rate Equation**
>
> Fouling became a quantity that could be modeled, where before it was a design margin, when Kern and Seaton wrote its net rate in 1959 as deposition minus removal. Removal in their model is proportional to the thickness of the deposit, so the resistance approaches an asymptote [13]. Crude oil then acquired a feature of its own. In 1995 Ebert and Panchal reanalyzed tests on crude drawn from a refinery preheat exchanger. Fouling was significant at film temperatures of 370 to 400 °C with velocities of 1.2 and 2.5 m/s, and none was detected at similar wall temperatures with 3.8 and 5.2 m/s. Their rate law sets formation, which rises with film temperature, against removal by wall shear. Where removal exceeds formation, negligible fouling is expected, and the line between is the fouling threshold [13]. The model of this chapter keeps the first idea and leaves out the second.

For crude oil the box also describes a threshold surface condition below which a given exchanger does not foul at all, and the useful design and operating question is which shells in the train sit above it. The threshold picture was reduced to a correlation fitted against laboratory and plant data [14] and then used to decide retrofits, since moving an exchanger to a different place in the train changes both its wall temperature and its velocity [15]. Distributed dynamic models that carry the deposit layer along the tube followed, and they show how far a lumped resistance is from the real spatial picture [16]. The economic weight of the problem is large. Fouling in process heat exchangers is a recurring energy penalty with combustion emissions attached to it, large enough to be studied as an environmental problem of its own [17], and crude preheat carries the largest share of it because it carries the largest surface [18].

Because the resistance grows monotonically, the response is a schedule rather than a controller. Optimal cleaning policies for a fouling exchanger network were posed as a mixed integer problem with the cleaning times as decisions [19], and applied to refinery preheat trains where the objective is the fuel saved against the cost of the outage [20]. A later addition to that literature is the term this application is built around. A fouled train does not merely burn more fuel. It eventually takes the fired heater to its limit, at which point the unit loses throughput continuously without an operating decision, and cleaning schedules computed with throughput in the objective are not the same schedules as those computed on fuel alone [18].

The control side of the same problem developed separately and much earlier, as the History box in Section 3.4 relates. A temperature controller that has saturated its actuator has no authority left, and the practice that answers it is to give the constraint a controller of its own, acting on a second manipulated variable, selected by a limit or a signal selector so that the loop reverts to normal service the moment the constraint releases. Shinskey set out override and valve position control in this form, including the rule that the constraint controller is tuned slower than the loop it protects [2], and the same structures with feedforward designed from a steady-state balance rather than from an identified model are standard textbook material [3]. Two developments since have changed what surrounds them rather than what they are. Constraint switching has been given a systematic design procedure, so the choice of which controller takes which variable in which region is no longer left to heuristics [4]. Model predictive control absorbed the same decision into an optimization

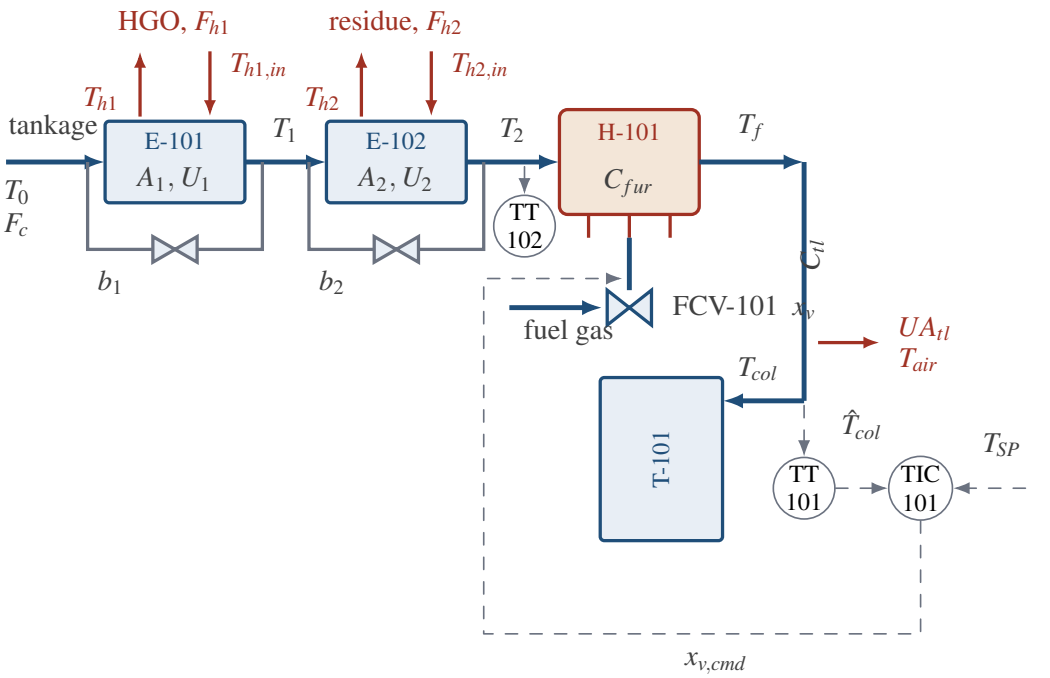


Figure 3.1 Crude preheat train E-101 and E-102, fired heater H-101 and fractionator T-101. Crude enters at tankage temperature T_0 and charge rate F_c, gains heat from heavy gas oil in E-101 and from atmospheric residue in E-102, and leaves the furnace coil at T_f into a transfer line of capacitance C_{tl} that loses $UA_{tl}(T_{col} - T_{air})$ to ambient. TT-101 supplies the measured column inlet to TIC-101, which drives fuel gas valve FCV-101 to position x_v, and TT-102 supplies the furnace inlet T_2 to the feedforward heat balance. Bypass fractions b_1 and b_2 take an exchanger out of service for cleaning, scaling its crude flow, its hot flow and its area in service together. The valve position controller that writes F_c appears in Figure 3.2.

with the constraints written in, which is how a modern crude unit handles it, and the survey of industrial practice records constraint handling rather than trajectory tracking as the primary reason for adoption [21]. The regulatory structure remains in service underneath, because a valve position controller requires one function block and no model.

3.3 Process Model

Crude leaves tankage at 25 °C and a design rate of 400 t/h, passes through two recovery exchangers in series, is fired the rest of the way in heater H-101, and arrives at the flash zone of the atmospheric column at 360 °C. That last temperature is the controlled variable, and the overflash of the column and the yield of every product drawn off it depend on it, so the deviation alarm band is a narrow 3 °C. E-101 is 620 m^2 of crude against heavy gas oil supplied at 285 °C. E-102 is 1000 m^2 of crude against atmospheric residue supplied at 345 °C, and its outlet is the furnace inlet. Each exchanger carries a bypass that takes it out of service for cleaning. The arrangement, the instrument tags and the symbols used below are all in Figure 3.1.

At the design condition with the fouling the application opens on, E-101 recovers 22.87 MW and lifts the crude to 118.6 °C, E-102 recovers 26.79 MW and takes it to 228.1 °C, and H-101 supplies the remaining 32.98 MW with its fuel valve at 81.3 %. The crude absorbs 82.63 MW in all, so 60.1 % of the duty arrives free. The important property of the fuel valve is that it stops: at 100 % the furnace absorbs 40.56 MW and there is no more.

Table 3.1 Controlled, Manipulated and Disturbance Variables With the Twelve Integrated States. Nominal Values Are the Design Condition Every Scenario Opens On.

Symbol	Role	Nominal	Range
T_{col} (TT-101)	CV, measured	360.0 °C	330–375 °C SP
x_v (FCV-101)	MV	81.3 %	0–100 %
$F_{c,cmd}$	MV, override or MPC	400 t/h	240–F_{dem}
F_{dem}	DV, measured	400 t/h	200–440 t/h
T_0 tankage	DV, measured	25 °C	0–45 °C
HV heating value	DV, unmeasured	100 %	80–115 %
$T_{h2,in}$ residue	DV, measured	345 °C	300–360 °C
F_{h2} residue flow	DV, measured	230 t/h	140–300 t/h
b transmitter bias	DV, unmeasured	0 °C	±10 °C
T_{c1}	state	118.6 °C	—
T_{h1}	state	155.8 °C	—
T_{c2}	state	228.1 °C	—
T_{h2}	state	183.8 °C	—
T_f	state	363.0 °C	$\leq$ 400 °C
T_{col}	state	360.0 °C	—
$R_{f,1}$	state	0.320	≥ 0
$R_{f,2}$	state	0.050	≥ 0
F_c	state	400 t/h	—
x_v	state	81.3 %	0–100 %
b_1, b_2	state	0	0–1

Table 3.1 lists the controlled, manipulated and disturbance variables together with the integrated states. Three annunciators bound normal operation, a column inlet deviation outside 3 °C, the fuel valve at or above 99.5 %, and a coil outlet above the 400 °C coking limit. All three annunciate without stopping the run, so a severe excursion can be observed to its end.

3.3.1 Working Units

Time is in hours, because fouling develops over three weeks while the temperature loop settles in minutes. The rest of the unit set follows refinery practice: mass in tonnes, flow in t/h, energy in megajoules, duty in MJ/h displayed as megawatts, heat capacity in MJ/t/K, UA in MJ/h/K, area in m^2, U in MJ/h/m^2/K and R_f in hm^{2}K/MJ. The set is self-consistent, so $Q = F c_p \Delta T$ carries no conversion factor anywhere. Two conversions relate the working units to familiar values. One megawatt is 3600 MJ/h, so the 40.56 MW furnace is 146 000 MJ/h, and 1 W/m^2/K is 3.6×10^{-3} in these units, so the 350 W/m^2/K clean coefficient of E-101 is 1.26.

3.3.2 Fouling as a Series Resistance

A deposit layer is a conduction resistance in series with the two clean film resistances, so the coefficients add reciprocally, as shown in Equation (3.1).

$$\frac{1}{U_i} = \frac{1}{U_{clean,i}} + R_{f,i} \qquad \Longleftrightarrow \qquad U_i = \frac{U_{clean,i}}{1 + U_{clean,i} R_{f,i}} \tag{3.1}$$

where $U_{clean,i}$ is the clean overall coefficient of exchanger i and $R_{f,i}$ is its fouling resistance, floored at zero. Nothing else about the exchanger changes. The area, the two flows and the inlet driving force are unchanged, and the duty falls because UA falls. At its cleaning threshold of 0.42 hm^{2}K/MJ the E-101 coefficient is down from 1.26 to 0.824, a loss of 34.6 %, and at its own threshold of 0.60 hm^{2}K/MJ E-102 is down from 1.152 to 0.681, a loss of 40.9 %, both in MJ/h/m^2/K. Every result in this chapter is a consequence of (3.1) and of the fact that R_f grows monotonically between cleanings.

The resistance follows a deposition term minus a removal term, both dependent on crude velocity, as shown in Equation (3.2).

$$\frac{\mathrm{d}R_{f,i}}{\mathrm{d}t} = (1-b_i)\left[\frac{R_{f\infty,i}}{\tau_{f,i}}\,v^{-n_d} - \frac{v^{n_r}}{\tau_{f,i}}R_{f,i}\right], \qquad v = \frac{F_c}{F_{c,D}} \tag{3.2}$$

where $R_{f\infty,i}$ is the asymptotic resistance at design velocity, $\tau_{f,i}$ is the fouling time constant at design velocity, n_d and n_r are the deposition and removal velocity exponents, v is velocity as a fraction of design and b_i is the bypass fraction, which stops the clock while an exchanger is isolated. Both sides carry units of $\mathrm{m^2K/MJ}$. The steady solution of (3.2) is $R_{f\infty,i}\,v^{-(n_d+n_r)}$ reached with time constant $\tau_{f,i}/v^{n_r}$, so with $n_d = 0.4$ and $n_r = 1.0$ the asymptote scales as $v^{-1.4}$ and operation at reduced rate raises the asymptote while lengthening the approach to it. At 300 t/h the E-101 asymptote rises from 0.55 to $0.55 \times 0.75^{-1.4} = 0.823$ and the time constant stretches from 420 to 560 h, so the threshold is reached at hour 124 instead of hour 240. Deposition falls with velocity because a faster stream spends less time at the wall, while shear removal rises faster than the deposition falls.

3.3.3 Exchanger Energy Balances

Each exchanger is two lumped thermal capacities with a log-mean driving force between them, as shown in Equations (3.3) to (3.5).

$$C_{ic}\frac{\mathrm{d}T_{ci}}{\mathrm{d}t} = (1-b_i)F_c c_{p,c}\,(T_{c,in,i} - T_{ci}) + Q_i \tag{3.3}$$

$$C_{ih}\frac{\mathrm{d}T_{hi}}{\mathrm{d}t} = (1-b_i)F_{hi} c_{p,hi}\,(T_{h,in,i} - T_{hi}) - Q_i \tag{3.4}$$

$$Q_i = (1-b_i)\,U_i A_i\,\mathrm{LMTD}\big(T_{h,in,i} - T_{ci},\; T_{hi} - T_{c,in,i}\big) \tag{3.5}$$

where C_{ic} and C_{ih} are the cold and hot side capacitances in MJ/K, T_{ci} and T_{hi} are the two outlet temperatures, $T_{c,in,i}$ and $T_{h,in,i}$ are the two inlets, and LMTD is the log mean of the two end temperature differences. Where either end is non-positive, which a transient produces while a bypassed exchanger is coming back on line, the arithmetic mean takes over so that the right-hand side stays finite. The crude inlet of E-101 is the tankage temperature and the crude inlet of E-102 is the E-101 station outlet. At steady state (3.3) to (3.5) reproduce the counterflow design equation $Q = UA \cdot \mathrm{LMTD} = \varepsilon C_{min}\Delta T_{max}$ exactly, which one of the application's own verification tests checks to one part in ten thousand.

The bypass fraction b_i appears three times, scaling the crude flow through the tubes, the hot flow through the shell and the area in service by the same factor. Scaling all three together is what a bypass around a fixed exchanger does, and it has two useful consequences. The exchanger's own steady outlet temperature does not depend on b_i, while the station duty falls linearly with it, as shown in Equation (3.6).

$$\begin{aligned} T_{station,i} &= (1-b_i)T_{ci} + b_i T_{c,in,i} \\ Q_{s,i} &= F_c c_{p,c}\,(T_{station,i} - T_{c,in,i}) \end{aligned} \tag{3.6}$$

At $b_i = 1$ every derivative inside that exchanger is zero, so an isolated exchanger holds still instead of drifting toward an unphysical stagnation temperature, and its fouling clock stops by the same factor in (3.2).

3.3.4 Fired Heater and Transfer Line

The furnace adds a duty proportional to valve position and to fuel gas heating value, and the transfer line carries the crude to the column while losing heat to ambient, as shown in Equations (3.7) and (3.8).

$$C_{fur}\frac{\mathrm{d}T_f}{\mathrm{d}t} = F_c c_{p,c}(T_2 - T_f) + \frac{x_v}{100}\,Q_{fur,max}\,\frac{HV}{HV_0} \tag{3.7}$$

$$C_{tl}\frac{\mathrm{d}T_{col}}{\mathrm{d}t} = F_c c_{p,c}(T_f - T_{col}) - UA_{tl}(T_{col} - T_{air}) \tag{3.8}$$

where T_2 is the furnace inlet, $Q_{fur,max}$ is the absorbed duty at full firing, HV/HV_0 is the heating value as a fraction of design, C_{fur} is the coil and firebox capacitance and C_{tl} is the transfer line capacitance. Equation (3.7) is where the constraint of the whole application lives, because x_v is clamped to 100 % and $Q_{fur,max} = 146\,000\,\mathrm{MJ/h}$ is a constant. The loss term in (3.8) is about 0.75 MW at the design condition, which is why the coil has to leave about 3 °C hotter than the column wants and why UA_{tl} appears explicitly in the feedforward of Section 3.4. The heating value ratio multiplies the duty and appears nowhere in any controller, which makes it the disturbance the feedforward is structurally blind to.

The four actuators are first-order lags on commands the controller has already clamped, as shown in Equation (3.9).

$$\begin{aligned} \tau_F \frac{\mathrm{d}F_c}{\mathrm{d}t} &= F_{c,cmd} - F_c, \quad \tau_{fire} \frac{\mathrm{d}x_v}{\mathrm{d}t} = x_{v,cmd} - x_v \\ \tau_{byp} \frac{\mathrm{d}b_i}{\mathrm{d}t} &= b_{i,cmd} - b_i \end{aligned} \tag{3.9}$$

No actuator can therefore walk past its stop and none needs a blocking test at the limit. A bypass commanded shut is snapped onto its command once the remaining travel falls below 10^{-5}, because a first-order response otherwise approaches the command only asymptotically.

3.3.5 Cleaning as a Discrete Event

A cleaning request opens the bypass of one exchanger, holds it open for $t_{clean} = 16\,\mathrm{h}$, then sets that exchanger's resistance to zero in one instant and commands the bypass shut. The sixteen hours represent the outage rather than the cleaning alone, covering isolation, steam out, hydroblasting, box up and line out, and the outage duration is what determines the throughput cost. Nothing else about the exchanger is changed, so a cleaning is exactly a reset of one state in (3.2), and the fouling that follows starts again from zero with a run length set by the charge rate.

3.3.6 Parameters, Initialization and the Instrument Chain

Table 3.2 lists the constants. Every parameter value is exposed and adjustable in the application, so every result below can be reproduced or perturbed.

The initial state is solved in closed form rather than by simulation to steady state, because with both bypasses shut the train is a forward chain: the counterflow effectiveness relation fixes E-101 from the crude and gas oil inlets, then E-102 from the E-101 outlet, and then the firing that puts the column inlet exactly on setpoint, as shown in Equation (3.10).

$$\begin{aligned} T_f^* &= T_{SP} + \frac{UA_{tl}(T_{SP} - T_{air})}{F_c c_{p,c}} \\ x_v(0) &= \frac{100\,F_c c_{p,c}(T_f^* - T_2)}{Q_{fur,max}} \end{aligned} \tag{3.10}$$

At the design condition (3.10) returns $T_f^* = 363.05\,^\circ\mathrm{C}$ and $x_v(0) = 81.32\,\%$, and manual operation provides the same heat balance calculation on request. Where the firing that (3.10) asks for exceeds 100 %, the initializer takes the steady state the clamped valve actually produces rather than the one that was requested.

Only two measurements reach the controller. The column inlet transmitter adds a constant bias and a noise sequence, built by filtering Gaussian draws and rescaling the result to a specified standard deviation, then passes the sum through a thermowell lag of $\tau_{TT} = 0.01\,\mathrm{h}$ (36 s). The furnace inlet transmitter carries no bias, a noise sequence of 60 % of the same amplitude and a faster lag of $\tau_{FT} = 0.004\,\mathrm{h}$, faster because the value of the feedforward depends on the quality of that measurement. Each filter is advanced by $\hat{T}_k = \alpha \hat{T}_{k-1} + (1-\alpha)T_k$ with $\alpha = \exp(-\Delta t_c/\tau)$, the exact solution over one controller period, so no choice of period can destabilize it. There is no lag and no noise on the charge flow indication, the duties or the fouling resistances, which are reported exactly.

Table 3.2 Model Parameters at the Design Condition

Symbol	Meaning	Value	Unit
$F_{c,D}$	design charge rate	400	t/h
$c_{p,c}$	crude heat capacity	2.2	MJ/t/K
$F_{c,min}$	minimum stable charge rate	240	t/h
A_1	E-101 area	620	m^2
$U_{clean,1}$	E-101 clean coefficient	1.26	MJ/h/m^2/K
F_{h1}	heavy gas oil flow	260	t/h
$c_{p,h1}$	heavy gas oil heat capacity	2.45	MJ/t/K
$T_{h1,in}$	heavy gas oil supply	285	°C
C_{1c}, C_{1h}	E-101 capacitances	90, 70	MJ/K
A_2	E-102 area	1000	m^2
$U_{clean,2}$	E-102 clean coefficient	1.152	MJ/h/m^2/K
$c_{p,h2}$	residue heat capacity	2.60	MJ/t/K
C_{2c}, C_{2h}	E-102 capacitances	250, 170	MJ/K
$R_{f\infty,1}$	E-101 asymptote	0.55	hm^2K/MJ
$R_{f\infty,2}$	E-102 asymptote	0.85	hm^2K/MJ
$\tau_{f,1}$, $\tau_{f,2}$	fouling time constants	420, 340	h
$R_{f,clean,1}$	E-101 cleaning threshold	0.42	hm^2K/MJ
$R_{f,clean,2}$	E-102 cleaning threshold	0.60	hm^2K/MJ
n_d, n_r	velocity exponents	0.4, 1.0	—
t_{clean}	cleaning outage	16	h
$Q_{fur,max}$	absorbed duty at full firing	146000	MJ/h
η	furnace thermal efficiency	0.88	—
C_{fur}	coil and firebox capacitance	100	MJ/K
C_{tl}	transfer line capacitance	85	MJ/K
UA_{tl}	transfer line heat loss	8	MJ/h/K
T_{air}	site ambient	25	°C
$T_{f,max}$	coil outlet coking limit	400	°C
—	fuel gas price	6.0	$/GJ
τ_{fire}	firing response lag	0.05	h
τ_F	charge rate response lag	0.05	h
τ_{byp}	bypass swing time	0.20	h
τ_{TT}	column inlet thermowell lag	0.01	h
τ_{FT}	furnace inlet thermowell lag	0.004	h
T_{band}	column inlet alarm band	3	°C

3.3.7 Time Scales and Numerical Integration

The simulation holds a temperature loop that settles in minutes and a fouling cycle that develops over three weeks inside one integration step and one session, so the step size is chosen against the fast end while the session length is chosen against the slow one. The fastest thermal mode in the plant is the transfer line, whose time constant is given by Equation (3.11).

$$\tau_{tl} = \frac{C_{tl}}{F_c c_{p,c} + UA_{tl}} = \frac{85}{880 + 8} = 0.0957\,\text{h} \tag{3.11}$$

which is 5.74 min. The step, the controller period and the log interval below are the three clocks of Section 1.6.1. The fixed integration step is $\Delta t = 0.01$ h (36 s), so (3.11) is resolved 9.6 times over, which is comfortable for classical fourth-order Runge-Kutta on a smooth right-hand side. The controller period is $\Delta t_c = 0.02$ h (72 s), a realistic scan for a furnace outlet loop, and the loop it closes is dominated by the coil at $C_{fur}/(F_c c_{p,c}) = 100/880 = 0.114$ h, which gives 5.7 controller executions per time constant. Neither ratio is generous. Halving either would double the cost of the three-week run for no visible change in the trajectory.

The slow end fixes the session. Three weeks is 504 h, or 50 400 integration steps, each one four evaluations of a twelve-state right-hand side. Measured on this model at 2.4 µs per step, a full three-week run takes 122 ms of processor time, so at the top acceleration of 100× the whole three weeks replays in about five seconds. That replay time is what keeps the fouling scenarios observable within one session, because the lesson of those scenarios is a trajectory shape that develops over the full three weeks. The recorded history holds one sample every third integration step, 0.03 h (1.8 min), or 16 801 rows for a three-week run.

3.3.8 Assumptions

Model detail is concentrated on the constraint. Each simplification below is stated together with the conditions under which it fails.

Two lumped capacities per exchanger with a log-mean driving force. The formulation reproduces the counterflow design equation at steady state and gives the right dominant time constant, but it carries no spatial temperature profile, so a temperature cross, shell-side bypassing and the difference between one shell and two in series cannot appear. A distributed model that tracks the deposit along the tube shows how much of the real behavior that costs [16].

Fouling is one uniform resistance per exchanger driven only by velocity. Real crude fouling depends strongly on wall temperature and on the asphaltene stability of the particular blend, is worse at the hot end of the train than the cold, and is not uniform over the surface. The threshold behavior that decides where in a real train deposit forms at all is absent, and that behavior is the whole basis of fouling-aware design [14, 15]. The time scale is also compressed for teaching, because a real crude preheat exchanger runs six months to two years between cleanings, so no interval reported here predicts a real cleaning schedule.

Cleaning is perfect and instantaneous at the end of a fixed duration. Mechanical cleaning typically recovers 85 to 95 % of the clean coefficient, and the residue left behind makes each successive run shorter.

The hot streams are boundary conditions. In the real unit the heavy gas oil and the atmospheric residue come off the same column the crude is being fed to, so a change in preheat changes the column, which changes the hot streams, which changes the preheat. That loop is cut here, and it is the largest structural omission in the model. Chapter 4 ties its preheat to the pumparound duties that recover the tower's heat.

Furnace absorbed duty is proportional to valve position with a fixed efficiency. There is no draft control, no excess oxygen, no pass flow balancing and no radiant and convective split, so nothing can be said about furnace efficiency optimization or pass balancing, which is where much of the real work on a fired heater goes. Volume 3, Chapter 4 treats the air side of such a heater, with oxygen trim and draft control.

Nothing vaporizes and crude properties are constant. A real crude heater has 30 to 50 % vaporization at the coil outlet, so the coil is two-phase over its last third. Between the tank farm and the flash zone the specific heat rises about 15 % and the viscosity falls by two orders of magnitude, which moves both the duty and the film coefficients by more than the fouling does over a short run.

The column is a temperature target and nothing else. There is no flash calculation, no overflash, no product draw and no tray hydraulics, and the minimum stable charge rate of 240 t/h stands in for the whole of the column's own constraint set.

Perfect repeatability is a teaching convenience. A fixed step and a fixed scan give one trajectory everywhere, which no real control system does.

3.4 Controller Design

Three control architectures of increasing capability are provided, as shown in Figure 3.2, with a model predictive controller as a fourth option beside them (Section 3.4.4). Feedback alone is the first, feedback plus a furnace heat balance is the second, and the third adds the valve position controller

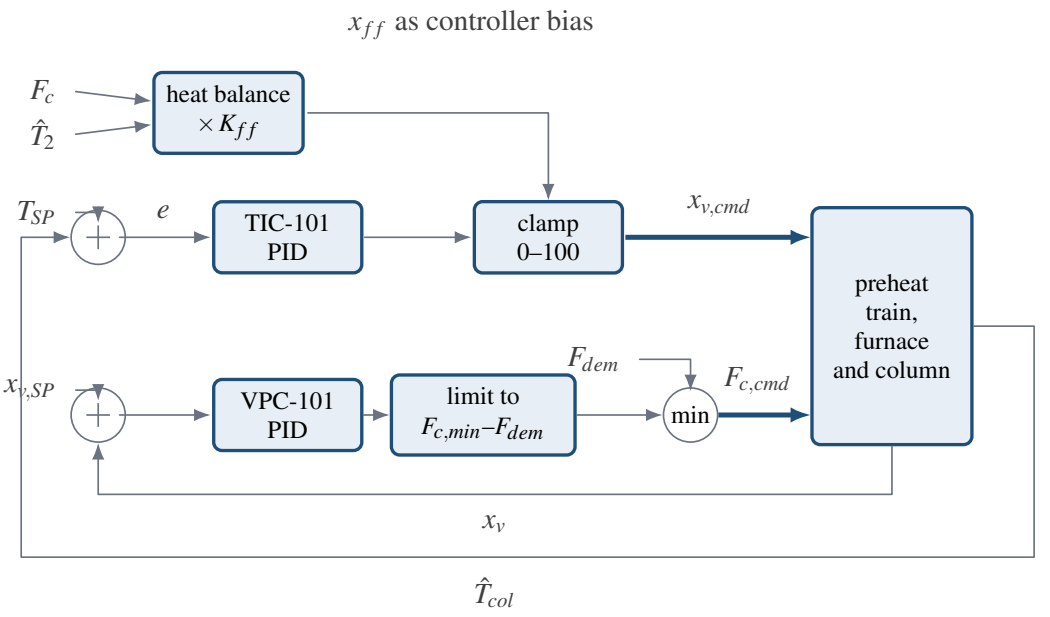


Figure 3.2 The three control architectures. TIC-101 alone is the *Single* architecture. Adding the furnace heat balance, scaled by K_{ff} and entered as the controller bias, gives + *Feedforward*. Adding VPC-101, whose process variable is the fuel valve position itself and whose output is a charge rate low-selected against the operator demand, gives *Override*. The upper output limit of VPC-101 is F_{dem}, so an override that is not selected rests on that clamp and cannot wind up.

that is the subject of this chapter. Switching between them is a bumpless transfer rather than a step, because moving from the first to the second puts the whole heat balance into the controller bias and would otherwise arrive as a demanded step of about eighty points on the fuel valve, the 81.3 % the balance asks for at the design condition. For one execution after any architecture change both controllers are made to track the actuator positions the plant is already holding, which is the same mechanism the manual to automatic transfer of Section 1.6.2 uses, and the switch leaves no mark on the trend.

3.4.1 Temperature Controller TIC-101

The column inlet controller is the velocity-form algorithm of Section 1.6.2, written in the loop's own symbols in Equations (3.12) to (3.14). The derivative acts on the measurement, so a setpoint move on

its own puts no kick into the fuel valve.

$$e_k = T_{SP,k} - \hat{T}_{col,k} \tag{3.12}$$

$$\Delta u_k^P = K_c\,(e_k - e_{k-1})$$

$$\Delta u_k^I = \frac{K_c}{\tau_I}\,e_k\,\Delta t_c$$

$$\Delta u_k^D = -\,K_c\,\tau_D\,\frac{\hat{T}^f_{col,k} - 2\hat{T}^f_{col,k-1} + \hat{T}^f_{col,k-2}}{\Delta t_c} \tag{3.13}$$

$$s_k = s_{k-1} + \Delta u_k^P + \Delta u_k^D$$

$$c_k = x_{ff,k} + s_k + I_{k-1}$$

$$u_k = \mathrm{sat}(c_k, 0, 100)$$

$$I_k = I_{k-1} + \Delta u_k^I + \frac{\Delta t_c}{T_t}\,(u_k - c_k) \tag{3.14}$$

Here s_k is the running sum S_k of Equation (1.17), c_k is the valve position the algorithm demands, the v_k of Equation (1.18) with the feedforward x_{ff} as its bias, and the integral used at scan k is the one accumulated through scan $k-1$. The action is $a = +1$ in Equation (1.16), reverse acting in the ISA convention, because more fuel valve raises the column inlet temperature. The derivative increment is a second difference on a measurement first passed through a filter of time constant τ_f defaulting to 0.02 h, and τ_D defaults to zero because derivative action against thermowell measurement noise is rarely justified on a loop of this speed. Reset is limited by back-calculation, with a tracking time T_t that the application passes explicitly as $T_t = \tau_I$, so at the default tuning each scan removes $\Delta t_c/\tau_I = 5.7\,\%$ of the clipped excess. With $T_t = \tau_I$ the integral comes to rest where $x_{ff} + I$ equals the limit, so the proportional term alone carries the demand past it and the valve leaves its stop as soon as the error changes sign. An integral time at its upper bound of 4 h is treated as infinite, and the integral update, back-calculation included, is then skipped. Output limits are 0 and 100 % with no rate limit on top, since the firing lag τ_{fire} of (3.9) is already the rate limit. In manual the untouched valve station is initialized to the solved starting valve position (81.3 % at design conditions), and selecting Manual before Run holds that lineup. A valve position the learner sets survives initialization and Reset. The integral state of (3.14) is seeded so that u_k equals the manual output, which is why no dedicated transfer logic exists. The adjustable ranges are $0 \le K_c \le 6\,\%/^\circ\mathrm{C}$, $0.05\,\mathrm{h} \le \tau_I \le 4\,\mathrm{h}$ and $0 \le \tau_D \le 0.5\,\mathrm{h}$, and the default tuning is $K_c = 1.5\,\%/^\circ\mathrm{C}$, $\tau_I = 0.35\,\mathrm{h}$, $\tau_D = 0$.

The gain limit follows from the process gain and the total lag. The process gain from valve to column inlet follows from (3.7) and (3.8) at steady state as $(Q_{fur,max}/100)/(F_c c_{p,c} + UA_{tl}) = 1460/888 = 1.64\,^\circ\mathrm{C}$ per percent of valve, and the lags between the valve and the thermowell total 0.269 h once the firing response (0.05 h), the coil (0.114 h), the transfer line (0.096 h) and the thermowell (0.01 h) are added. A gain of 6 therefore puts a loop gain near ten around that chain, which is why the measured tuning sweep in Section 3.5 degrades above 1.5 rather than continuing to improve.

3.4.2 Heat Balance Feedforward

The second architecture computes the firing the furnace needs directly from a heat balance on H-101 and the transfer line, as shown in Equation (3.15), the balance of Equation (1.2) with the transfer line loss added.

$$x_{ff} = \frac{100\,K_{ff}}{Q_{fur,max}}\,F_c c_{p,c}\left(T_{SP} + \frac{UA_{tl}(T_{SP} - T_{air})}{F_c c_{p,c}} - \hat{T}_2\right) \tag{3.15}$$

where $\hat{T}_2$ is the measured furnace inlet temperature and K_{ff} is a detuning gain with a default of 1.0. Both inputs are instruments a crude unit already has, the charge flow transmitter and the thermocouple at the train outlet, and there is no step test and no model fit behind (3.15) because the heat balance is the model. At $K_{ff} = 1$ it is algebraically the same expression as the steady-state firing requirement

of (3.10), so the feedback trim on a settled plant is zero to the resolution of the trend, and measured in scenario 3 just before its rate cut at 2 h the feedforward asks for 81.45 % of valve where the plant holds 81.45 %, a trim of 0.004 points. The result is clamped to 0–100 % and entered as the controller bias x_{ff} in (3.14), which is the whole of the implementation. Because the bias passes through the clamp, the back-calculation and the manual tracker, no term had to be added anywhere else. Equation (3.15) is written at the design heating value, because the heating value is unmeasured, and that omission is why the fuel gas scenario defeats the feedforward. Chapter 14 feeds the Wobbe index of the gas forward to its burners instead.

Pause. Try Exercise 3.2 before reading on.

3.4.3 Valve Position Override VPC-101

The third architecture adds a second PID whose process variable is the fuel valve position, the override of Equation (1.10), and whose output is a crude charge rate in t/h, combined with the operator demand as shown in Equation (3.16).

$$\begin{aligned} F_{c,cmd} &= \min\left(F_{dem},\ \mathrm{PID}_{VPC}(x_{v,SP}, x_v)\right) \\ u_{max}^{VPC} &= F_{dem} \end{aligned} \tag{3.16}$$

Three properties of (3.16) make an override inexpensive to implement. It is a high limit rather than a selector chain, since the override's own upper output limit is the operator's demand, so whenever the furnace has margin the controller rests on that clamp and the operator demand sets the charge rate. Its anti-windup is therefore free, because the saturated output that the back-calculation in (3.14) drives the integral toward is the demand itself, and an override that is not selected cannot wind up. The tracking time is again the integral time, so the integral comes to rest on F_{dem} and the proportional term holds the unclamped output $K_{c,VPC}(x_{v,SP} - x_v)$ above it, which means the override is selected only once the valve crosses its target. A recursion that absorbed the whole clipped excess in one scan would leave the integral at $F_{dem} - K_{c,VPC}(x_{v,SP} - x_v)$ and start cutting as soon as the valve began to rise, and in a counterfactual run of that form the E-102 outage of Section 3.5 dips to 326.9 °C instead of 317.5 °C. While VPC-101 is not selected, under another architecture, in manual, or on the execution that switches architecture (the first execution of every run among them), it tracks the demand, and with the valve below its target its integral is set to F_{dem}, the state in which back-calculation holds it. A target lowered before the valve reaches it therefore cuts nothing. It is deliberately slow, with a default integral time of 0.5 h against the 0.35 h of the temperature loop, following the rule that the constraint controller is tuned slower than the loop it protects [2]. Reset is retained all the same, because an override with proportional action alone would settle a fixed distance below its valve position target and surrender throughput the furnace could carry. The default gain is 4 t/h of charge cut per percent of valve above target, the valve position target is 95 %, and the lower output limit is the minimum stable charge rate of 240 t/h (or the demand, when that is lower), which is a fractionator constraint rather than a furnace one.

The order inside one control execution never varies. The tankage temperature and its diurnal term are evaluated first, then both transmitters with their bias, noise and filters, then the feedforward, then the tuning is re-applied, then TIC-101 or its manual tracker, then VPC-101 or its tracker, then the low select of (3.16). The two temperatures the controllers use therefore come through transmitters with their bias, noise and lag, while the charge flow in the feedforward and the valve position that VPC-101 controls are read exactly.

> **HISTORY 1945: Limit Control Enters the Textbooks**
>
> Process control grew up with the large continuous plants, refineries among them, that appeared around 1920. They were first run by hand, with one operator for each valve [22]. Proportional control came in the 1920s, automatic reset followed, and cascade and ratio control were in wide use in the 1930s. Selective

control, also called limit or override control, came into use to handle changes of constraint, and a 1945 book on industrial process control already described it beside ratio, cascade and split range control. By the 1960s the process industry worked with a standard set of such elements, valve position control among them [22]. Predictive control became common in refining at the end of the 1980s, and the review that gives these dates finds the older elements still thriving in practice beside it [22].

3.4.4 Model Predictive Control

The fourth architecture hands the fuel valve and the charge rate to one receding-horizon quadratic program [23] in automatic mode, and the limit VPC-101 exists for becomes a hard bound on x_v inside the optimization. Only scenario 13 opens on it, the three architectures above remain the application's default, and the comparison of Section 3.5.6 treats the predictive controller as an alternative to them rather than as a further rung. The option follows the general form of Section 1.6.3, with a filtered output bias in place of a Kalman filter and the plan's own objective below.

Controller model. The steady state is the feedforward balance of (3.15) solved for the column inlet on the measured furnace inlet, as shown in Equation (3.17).

$$T_{ss} = \frac{F_c c_{p,c} \hat{T}_2 + (x_v/100)\, Q_{fur,max} + UA_{tl} T_{air}}{F_c c_{p,c} + UA_{tl}} \tag{3.17}$$

Its four constants are read where the feedforward reads them, so a learner's change of furnace capacity reaches both structures or neither, and the heating value is the design value, as in (3.15). The slope K_v is 1.644 °C per percent of valve at 400 t/h and 2.14 at 306.6 t/h. The dynamics were identified from valve and charge steps at the design condition with noise off. Three equal lags of $\tau_c = 0.0921$ h from the valve command to the measured column inlet leave a residual of 0.54 % of the response, against 1.42 % for a first-order lag with dead time. At fixed $\hat{T}_2$ the balance carries only 50.7 % of the charge's steady effect, −0.327 of −0.644 °C per t/h, because a slower crude also leaves E-102 hotter. The remainder is the identified furnace inlet response $\delta T_2 = g_{T2}\, \delta F_c / (\tau_{T2} s + 1)$ with $g_{T2} = -0.3238$ °C per t/h and $\tau_{T2} = 0.3783$ h, with which a charge step on the column inlet is predicted within 1.06 % of the response. The duty part of (3.17) enters at the first lag, and the furnace inlet part enters 41 % at the first lag and the rest at the second, a split that keeps the model within 0.54 °C of the measurement while an E-102 bypass with the valve held swings $\hat{T}_2$ by 112 °C. An output bias, the measured column inlet less the model output filtered over 0.06 h, makes the prediction offset-free [24]. Fouling therefore reaches the plan through the furnace inlet thermocouple, while a lean fuel gas, a biased thermocouple and every model error reach it through the bias. The four-state model is relinearized at the applied commands at every decision.

Objective and constraints. Every $T_s = 0.06$ h, three controller scans, the plan covers $N = 33$ steps (1.98 h) in ten move blocks of 1, 1, 1, 2, 2, 3, 4, 5, 6 and 8 steps, 20 decision variables and one slack, applies the first move and holds it, minimizing Equation (3.18).

$$\begin{aligned} J = \sum_{k=1}^{N} \left[q_T \left(\hat{T}_{col,k} - T_{SP} \right)^2 - p F_{c,k} T_s \right] + \rho \left(\varepsilon^2 + \varepsilon \right) \\ + \sum_{j} \left[S_v \Delta x_{v,j}^2 + S_F \Delta F_{c,j}^2 \right] \end{aligned} \tag{3.18}$$

Here $\hat{T}_{col,k}$ is the predicted column inlet, $q_T = 10$ per °C^2 and step, $p = 1$ per tonne prices the charge as a linear cost, $S_v = 0.05$ per %^{2} and $S_F = 0.005$ per (t/h)^2 price the moves Δ at each block boundary, and $\rho = 10^3$. The valve is held hard in $0 \le x_v \le x_{v,max}$, with the ceiling $x_{v,max} = 100$ % as shipped, and the charge hard in $\min(F_{c,min}, F_{dem}) \le F_c \le F_{dem}$, the range VPC-101 clamps to. The one slack $\varepsilon \ge 0$ softens a floor $\hat{T}_{col,k} \ge T_{SP} - T_{band}$ over the whole horizon. The price keeps the charge at the demand while the valve has travel, and once the valve sits on its bound it holds the cut where the full furnace carries the setpoint, 306.612 t/h at hour 10 of scenario 13 against 306.616 t/h from the closed-form

Table 3.3 The Fourteen Scenarios, With the Mechanism Each Is Built Around

#	Scenario	What it exercises
1	Fire the furnace by hand	charge rate as the dominant disturbance
2	Temperature loop against a rate change	feedback distance, the gain limit
3	Feedforward from the heat balance	a model written rather than identified
4	Three weeks of fouling, nothing else	a perfect loop hiding a lost margin
5	Clean E-101 at full rate on feedforward	a saturated actuator ends the loop
6	The same clean under override	throughput as the second MV
7	Clean E-102, the expensive one	where the exchanger sits in the train
8	Two cleaning cycles out of phase	a schedule as a control decision
9	The same two cleanings colliding	two constraints binding at once
10	Fuel gas heating value to 88 %	the disturbance feedforward cannot see
11	Thermocouple reading 6 °C high	feedforward as a soft sensor
12	Three weeks at 300 t/h	velocity dependence of fouling
13	MPC with the whole valve as a hard limit	a gain traced to its sources
14	Graded challenge over three weeks	architecture before tuning

balance. Per degree of column inlet a charge move is priced below a valve move, $S_F/K_F^2 = 0.012$ against $S_v/K_v^2 = 0.019$ per °C^2 with K_F the charge's steady gain, so a fast upset also dips the charge while valve remains. The weights were set once on scenario 13 and are not retuned for any other case.

Acceptance, fallback and cost. A plan is applied only if the solver reports it optimal, every planned command lies inside its box, the plan's first command is the move sent, and its cost, slack and predictions are finite. Otherwise *Override* takes both handles, its TIC-101 and VPC-101 having tracked the applied commands exactly, a fallback mark goes into the historian, and the first accepted solve after it hands both handles back to the plan. Forced at six instants of scenario 13 with noise off, the hand-over moves the valve by at most 0.003 % and the charge by 0.05 t/h, except during the restore at hour 18.3, where the override's own first moves are 0.99 % and 0.48 t/h. The shipped rule that rests the VPC-101 integral at F_{dem} is suspended under the plan, because at hour 6 it would turn a fallback into a charge step of 73 t/h. In Chrome a plan takes 183 to 191 µs on average over the 800 plans of scenario 13, about 0.3 % of one core at a playback rate of one plant hour per second.

Three models kept apart. The plant is the twelve-state model with its true fouling, heating value and thermocouple bias, the identified model is the three lags and the furnace inlet response, and the controller's model is (3.17) with those dynamics and the bias, fed only the two transmitters, the setpoint, the applied commands, the demand and the floor.

3.5 Results

Fourteen scenarios are arranged in instructional order, listed in Table 3.3. Each one fixes the opening mode, the architecture, the three tuning constants, the disturbance settings, any parameter overrides and any timed events, and supplies a briefing, a prediction question, a procedure, one or two hints, a worked explanation and a completion requirement that the app judges from the run (Section 1.6.4). The app lists scenario 14 last, as the challenge, without a number.

3.5.1 Fouling Under Closed-Loop Control

Scenario 4 runs three weeks with no other disturbance. The results are in the first column of Table 3.4, and the principal observation is that the temperature never moves. The integrated absolute error over the whole 504 h is 0.0 °C h and the column inlet sits on 360.0 °C, while the entire disturbance goes into the fuel valve, which opens from 81.3 % to 97.5 % without a step and without an alarm. Figure 1.4 draws the same hidden loss on R-1. The E-101 duty falls from 22.9 to 21.1 MW and E-102 from 26.8 to 21.9 MW, the furnace makes up all of it by going from 33.0 to 39.5 MW, and the recovered share of the duty falls from 60.1 % to 52.1 %.

Table 3.4 Three Weeks (504 h) of Fouling With No Other Disturbance, With $K_c = 1.5$ and $\tau_I = 0.35$ h. Every Figure Is Read at 504 h. The First Two Columns Run on + *Feedforward*. The Third Runs on *Override* and Cleans Each Exchanger at Its Own Threshold, at Hours 240 and 396.

	400 t/h no clean	300 t/h no clean	400 t/h two cleans
IAE (°C h)	0.0	—	57
Fuel valve, final (%)	97.5	—	—
$R_{f,1}$ final	0.481	0.618	—
$R_{f,2}$ final	0.668	0.870	—
E-101 threshold reached (h)	239.6	124.2	240
E-102 threshold reached (h)	395.5	—	396
Recovery, final (%)	52.1	—	—
Crude processed (kt)	201.6	151.2	199.1
Fuel fired (TJ)	76.7	—	72.8

Table 3.5 A Sixteen-Hour Cleaning Outage Started at $t = 2$ h, Every Figure Read at $t = 48$ h. Columns 1 to 3 Take E-101 Out, Column 4 Takes E-102 Out. All Four Carry the Same Feedback Tuning and $K_{ff} = 1.0$. The Worst Column Inlet Is Read From $t = 2$ h and the Lined-Out Value at the End of the Outage.

	feedfwd only	override $x_{v,SP} = 95$	override $x_{v,SP} = 85$	override E-102
Worst column inlet (°C)	340.3	358.6	358.8	317.5
Column inlet lined out (°C)	340.3	360.0	360.0	360.0
IAE over 48 h (°C h)	287	1.3	0.9	59.4
Minimum charge rate (t/h)	400	363	340	295
Average charge rate (%)	100.0	97.1	95.2	91.7
Crude processed (kt)	19.200	18.645	18.279	17.616
Time at firing limit (% of outage)	96.5	0	0	13.3

Pause. Try Exercise 3.1 before reading on.

The remaining margin is the concern. The valve stops at 97.5 % with 2.5 % left, while every measure a temperature loop reports indicates success. Raising the charge demand to 420 t/h puts the same three weeks on the stop at hour 215, and at 440 t/h at hour 32. Fuel over the three weeks comes to 76.7 TJ, about $460k at $6/GJ, and the fouling accounts for roughly a tenth of it.

The second column of Table 3.4 is a counterintuitive result. Turning the unit down to 300 t/h fouls it *faster*. After three weeks the E-101 resistance is 0.618 against 0.481 and E-102 is 0.870 against 0.668, which in threshold units is 1.47× and 1.45× against 1.14× and 1.11×, and E-101 crosses its cleaning threshold at hour 124 instead of hour 240. The mechanism is in (3.2). Deposition falls as $v^{-0.4}$ because a faster stream spends less time in contact with the wall, removal rises as $v^{1.0}$ because shear is what scrubs a soft deposit off, and the asymptote is the ratio, which goes as $v^{-1.4}$. The cost appears on two accounts. The turned-down month processed 151.2 kt rather than 201.6 kt, which was the intended reduction, and it also leaves the next run starting from a higher resistance, which was not.

3.5.2 Cleaning Outages Under Override Control

Scenario 5 opens the E-101 bypass at full charge rate on the feedforward architecture, and scenario 6 repeats it with the override selected and nothing else changed. Table 3.5 carries both, together with the same event on E-102 and the effect of moving the valve position target. Figure 3.3 draws three of its rows as a chart.

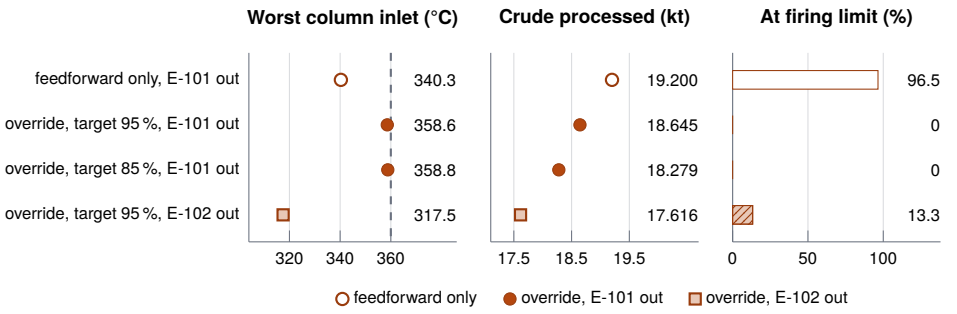


Figure 3.3 Three rows of Table 3.5 as a chart, the last as a share of the outage. The dashed line is the setpoint. The open circle and white bar are feedforward only, filled circles and solid bars the override with E-101 out, the square and hatched bar the override with E-102 out. Notice that the 85 % target gives up more crude than the 95 % target for the same temperature, and that the E-102 outage costs most on all three.

The duty accounting explains the first column. E-101 goes from 22.9 MW to zero, but E-102 goes *up*, from 26.8 to 37.6 MW at hour 6, because its crude inlet has just fallen from 118.6 °C to 25 °C and its driving force has risen by the same amount. The train therefore loses 12.0 MW rather than 22.9 MW. The furnace covers 7.6 of that by going from 33.0 MW to its 40.56 MW ceiling, and the remaining 4.4 MW is the 18 °C the column does not get at hour 6, widening to 19.7 °C by hour 18 as E-102 warms through and gives part of the gain back. The column inlet parks at 340.3 °C and recovery falls from 60.0 % to 48.1 %. The valve reaches its stop 0.57 h after the bypass starts to swing and stays there until 0.54 h after E-101 returns, 96.5 % of the outage, and for that time the temperature loop has no authority at all, because back-calculation holds the integral where it reproduces the saturated output and more gain only raises a demand the valve cannot follow.

The second column is the same event with VPC-101 allowed to write the charge rate. The column inlet lines out on 360.0 °C with a worst deviation of 1.4 °C while the bypass is swinging, the integrated error over the 48 h window falls from 287 to 1.3 °C h, and the fuel valve never reaches its stop. The recovery is obtained by cutting the charge rate to 363 t/h, 90.8 % of design, for the sixteen hours the exchanger is away, and the price is 555 t of crude out of 19 200 t, or 2.9 %. Two effects combine in that cut. A lower rate requires less furnace duty, and the crude also gains more temperature rise per tonne crossing the same 1620 m^2 of surface, which is why a 9 % rate reduction recovers about twenty degrees.

The third column shows why the valve position target is a design decision rather than a default, because the relationship is not monotone. At 100 % the override has nothing to act on once the valve is on its stop, and the run repeats scenario 5 exactly, 19.200 kt with a worst column inlet of 340.3 °C. At 95 % the temperature holds and 18.645 kt goes through. At 85 % the worst column inlet during the outage is no better, 358.8 °C against 358.6 °C, while the charge falls to 340 t/h and 18.279 kt goes through. Below the point at which the constraint actually binds, a tighter target buys no temperature and costs tonnes, and the temperature trend does not show the loss.

🕮 **Pause.** Try Exercise 3.3 before reading on.

The fourth column is the asymmetry between the two exchangers, and it is structural rather than a matter of size. When E-101 is bypassed, E-102 sits downstream of the gap and picks up more than half of the loss by itself. When E-102 is bypassed there is nothing downstream of it except the furnace, so the crude arrives at the coil at 124.6 °C instead of 211.3 °C and the furnace is asked for 58.3 MW, which is 144 % of what it has. The override cuts to 295 t/h, 74 % of design, and the temperature still dips to 317.5 °C because the override cannot cut faster than the duty disappears with a bypass swinging over twelve minutes. Setting the demand to 300 t/h before the run, so that the unit

Table 3.6 Two Cleaning Outages Over Three Weeks, Read at 504 h. The Third Column Moves the Second Clean of Scenario 9 to Hour 280 in a Counterfactual Run. The Fourth Shortens Both Outages to 8 h, the Comparison Scenario 9 Asks For.

	separated (scen. 8)	8 h apart (scen. 9)	moved to 280 h	8 h outages
IAE (°C h)	57	559	48	14
Worst column inlet (°C)	319.8	262.6	326.6	344.9
Minimum charge rate (t/h)	298	240	307	290
Crude processed (kt)	199.061	199.056	199.039	200.315

is already lined out at the lower rate *before* the bypass opens at hour 2, which is what an operator does, turns that dip into 354.4 °C and the integrated error over the window from 59.4 into 6.3 °C h, and it processes 14.334 kt against the 17.616 kt the override gave, so the planned cut processes 3.3 kt less crude than the override. A planned reduction trades throughput for temperature margin, and the trade has a measurable price.

> **TRY IT IN THE APP** **7 · Clean E-102, the expensive one**
>
> Load the scenario and press Run. Nothing has to be set. At hour 2 the bypass opens and E-102 leaves for sixteen hours. Watch the crude charge rate. The override cuts it to 295 t/h, 74 % of design, and the column inlet still dips to 317.5 °C, because the duty disappears faster than the override can cut.

3.5.3 Interaction of Two Cleaning Cycles

Scenario 8 cleans each exchanger when it reaches its own threshold, E-101 at hour 240 and E-102 at hour 396, so only one is ever away at a time. The third column of Table 3.4 carries the outcome: 199.1 kt at 98.7 % of design, an integrated error of 57 °C h over three weeks, the charge rate cut for 6.5 % of the run, and fuel of 72.8 TJ against 76.7 TJ for the same three weeks with no cleaning at all. Two cleans therefore cost 2.5 kt of throughput and buy back 3.9 TJ of fuel gas, about $23k at $6/GJ, plus the whole of the valve margin. Whether that is a good trade requires a crude margin the application does not carry, which is the form the scheduling argument takes in practice.

The periods are different by construction. Started from clean at the design rate, E-101 takes 606 h to reach its threshold and E-102 takes 416 h, and the application opens with E-101 0.76 of the way through its run and E-102 nearly clean, so the two thresholds arrive 156 h apart. Two cycles of unequal length drift in and out of phase forever, and scenario 9 is the in-phase half. Raising the E-102 starting resistance from 0.05 to 0.36 brings E-102 to its threshold at hour 229 and E-101 to its own at hour 239, and the schedule cleans them at hours 240 and 248, so the two sixteen-hour outages overlap for eight hours. The comparison in Table 3.6 is the strongest result in the application.

The reason the tonnes did not move is the reason the temperature did. The override cuts to 240 t/h, exactly the minimum stable charge rate, and then stops, because below that the fractionator trays weep. With the override on its own floor and the fuel valve on its stop, both manipulated variables are gone at once, and the column inlet goes wherever the heat balance puts it, which is 262.6 °C. Moving the second clean to hour 280 costs 17 t out of 199 000 and changes the integrated error by a factor of almost twelve, from 559 to 48 °C h. That run is a counterfactual, because the scripted clean at hour 248 always fires in the application, so scenario 9 poses the same comparison by shortening both outages to 8 h, which lets E-102 leave as E-101 returns. The worst column inlet is then 344.9 °C, the override reaches only 290 t/h and the integrated error is 14 °C h, while 10 h outages that still overlap by two hours fall to the same 262.6 °C and reach 184 °C h. The overlap rather than the length of the outages is what the column cannot carry. The consequence of two constraints binding together is not the sum of the two events, because the constraint set is not additive, and separating two outages that could collide costs essentially nothing.

Table 3.7 Feedforward Against Two Disturbances. The Charge Rate Case Is a 400–340–400 t/h Sequence Read at 16 h, the Fuel Gas Case a Step to 88 % Heating Value Read at 12 h. Identical Feedback Tuning in Every Column.

	charge rate		fuel gas	
	single	+ ff	single	+ ff
Peak deviation (°C)	+10.3	+3.0	−7.14	−7.12
IAE (°C h)	10.7	2.5	2.80	2.61

+ DID YOU KNOW? The Tonnes Stayed, the Temperature Did Not

From the simulator (Table 3.6). Two cleaning outages that overlap for eight hours process 199.056 kt of crude in three weeks, and the same two kept apart process 199.061 kt. The integrated temperature error is 559 °C h against 57 °C h, and the worst column inlet 262.6 °C against 319.8 °C.

3.5.4 Disturbances That Resemble Fouling

Anything that slowly loses duty resembles fouling in the duty accounting, and the application carries three such disturbances. Scenario 10 drops the fuel gas heating value to 88 %, so the same valve position delivers less heat. Table 3.7 puts the two architectures side by side on that disturbance and on the charge rate change of scenarios 2 and 3.

Against the charge rate the feedforward removes 77 % of the accumulated error, cutting the peak deviation from 10.3 to 3.0 °C and the integrated error to 2.5 °C h from 10.7. On the fuel gas it removes almost nothing. The two architectures reach the same peak to within 0.02 °C and differ by 7 % in integrated error. Feedforward acts on measurements and there is no heating value measurement in this plant, so (3.15) returns an unchanged output and the entire correction comes from the feedback trim. The valve settles about twelve points higher at 93.4 %, close to the $1/0.88$ that lean gas costs, and the signature is a valve that moved while every reported duty held still. Gain cannot substitute for either. A sweep on the charge rate case gives 32.0 °C h at $K_c = 0.5$, 10.7 °C h at 1.5, then the wrong way to 82.0 °C h at 3.0 and 142.6 °C h at 6.0, where the loop rings continuously.

The second is a hot stream that has cooled. Dropping the residue supply from 345 to 315 °C takes 30 °C of driving force out of E-102 and produces the same duty loss as fouling. The two are told apart because U is reported beside the duty, and fouling lowers U while a cold supply lowers the duty at unchanged U.

The third is a biased instrument, the only failure a feedback loop cannot detect on its own. Scenario 11 puts 6 °C of bias on the column inlet thermocouple, together with 0.6 °C of noise, so averaged over hours 12 to 24 the true column inlet is 354.0 °C while the measurement reads 360.0 °C and the loop performs perfectly throughout. The evidence exists outside the loop. Over the same hours the fuel valve averages 78.91 % where the same run without the bias puts it at 82.56 %, and 3.65 % of valve at the 1.64 °C per percent process gain is exactly the 6.0 °C. More directly, the feedforward term of (3.15) averages 82.55 % in both runs because it is computed from instruments that have not drifted, while the feedback trim moves from +0.01 to −3.64 %. A feedforward built from independent measurements is also a soft sensor, and a persistent disagreement between it and the controller that trims it indicates a failed instrument. Measurement noise has a separate cost, paid entirely in valve travel. At 0.6 °C of transmitter noise the fuel valve command travels 645 % of full stroke in a day with no disturbance on the plant, and at 2 °C it travels 2149 %, as the scorecard's valve travel reports.

3.5.5 Graded Challenge

Scenario 14 applies a fuel gas step at hour 100, an E-101 cleaning at hour 180, a cold front at hour 250 and an E-102 cleaning at hour 290 over the full three weeks, and asks for the column inlet held within ±3 °C once both exchangers are back, at least 190 kt processed, and the fuel valve off its stop at

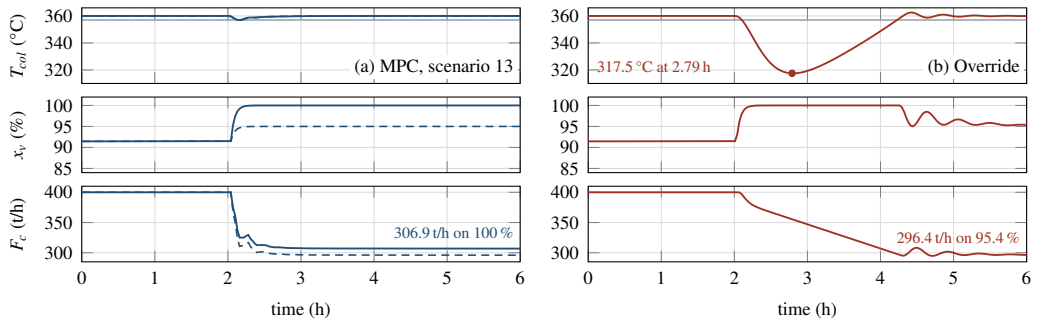


Figure 3.4 Scenario 13, the E-102 outage of scenario 7 beginning at hour 2, over its first six hours, the span the application's 12 h trend window shows at hour 6, under (a) the predictive controller, solid, and the same plan with its valve ceiling at 95 %, dashed, and (b) *Override* with its 95 % valve position target. Top, the column inlet against the plan's floor at 357 °C, the thin gray line. Middle, the fuel valve. Bottom, the charge rate. The plan's lowest column inlet is 357.0 °C at hour 2.16 and the override's 317.5 °C at hour 2.79.

the end. It opens on the single loop with a cautious tuning, and the result establishes that architecture rather than tuning decides this problem. On *Single* at the default gain the run processes 201.6 kt and spends 46.9 % of the three weeks with the fuel valve on its stop, with an integrated error of 3166 °C h. Retuning to $K_c = 1.5$ and $\tau_I = 0.35$ h gives 3091 °C h and adding feedforward gives 3085 °C h, so the two adjustments together are worth 2.5 %, and every one of these runs ends with the valve on its stop and fails. On the shipped settings the challenge's rule reads a worst deviation of 7.7 °C after the cleans and 201.6 kt processed, so the run fails two of its three criteria. Selecting *Override* from the shipped settings and changing nothing else gives 3141 °C h, the same as feedforward at that tuning, because the valve position target opens at 100 % and an override told to keep the valve below 100 % never acts. Moving that one number to 95 %, with the retuned loop, gives 86 °C h, a worst deviation after the cleans of 0.86 °C, reached as the E-102 bypass finishes closing, 2.1 °C inside the band, and 0.7 % of the run at the firing limit, for 194.4 kt against 201.6 kt, 4.4 kt above the floor, and an average charge rate of 96.4 %. The valve ends at 95.0 %, the margin the override was told to keep. With the shipped cautious tuning left in place the same target also passes, at 123 °C h and a worst deviation of 0.52 °C. Fuel falls as well, 69.9 TJ against 72.0 TJ, because a furnace held off its saturation limit does not overfire during recovery.

An integral of absolute error does not stop growing when the transient does, so every IAE above is quoted with the window it was read over, and the comparison carries the information rather than the absolute value. The 1.3 °C h of scenario 6 and the 287 °C h of scenario 5 are both read at 48 h on the identical event, and the three-week figures are all read at 504 h.

3.5.6 The Predictive Controller Against the Override

Table 3.8 runs scenarios 4 to 11 and the challenge, scenario 14, under the shipped structures and the predictive controller with each scenario's own seed, duration, instruments and noise and no preview, the plan at scenario 13's settings throughout, and on the challenge every row but the shipped one from the suggested settings. Three more rows locate the gain. The plan with $x_{v,max} = 95$ % keeps the override's margin. Override with its valve position target raised to 99 % is the shipped loop retuned once in the plan's direction. The same override with the balance of (3.15) as a feedforward on the charge, low-selected at the rate at which the target valve holds the setpoint on $\hat{T}_2$, is the like-for-like classical structure. It exists only in the measurement script, not in the application.

In scenario 13 the plan cuts the charge while the bypass is still swinging, so the lowest column inlet is 357.0 °C against the override's 317.5 °C, and by hour 6 it holds the valve on 100 % at 306.9 t/h

Table 3.8 The Predictive Controller Against the Shipped Structures, Whole Runs Measured on the Engine. Scenario 7's Plant Under the MPC Is Scenario 13.

Case	Structure	Not processed (t)	IAE (°C h)	Worst (°C)	Valve on stop (min)	Fuel (TJ)	Charge travel (t/h)	Plans, fallbacks
4 (504 h)	+ Feedforward, and both rows at 99 %	0	0.00	0.00	0	76.70	0	–
	Override	532	0.00	0.00	0	76.34	6	–
	MPC	0	0.38	0.00	0	76.70	0	8400, 0
	MPC, ceiling 95 %	533	1.05	0.01	0	76.34	6	8400, 0
5, 6 (48 h)	+ Feedforward (scenario 5)	0	286.7	19.7	959	6.82	0	–
	Override (scenario 6)	555	1.27	1.36	0	6.69	78	–
	MPC	386	1.55	0.77	929	6.82	54	800, 0
	MPC, ceiling 95 %	564	1.53	0.79	0	6.68	75	800, 0
	Override 99 % + charge feedforward	419	1.24	1.36	27	6.79	57	–
	Override 99 %	393	16.8	9.42	185	6.80	68	–
7, 13 (48 h)	Override	1584	59.4	42.5	128	6.81	271	–
	MPC	1497	2.57	3.02	954	6.92	240	800, 0
	MPC, ceiling 95 %	1676	2.53	3.14	0	6.79	269	800, 0
	Override 99 % + charge feedforward	1525	4.33	7.93	40	6.90	192	–
	Override 99 %	1003	427.2	65.1	707	6.92	204	–
8 (504 h)	Override	2539	57.4	40.2	184	72.83	421	–
	MPC	2280	4.51	3.21	1895	73.07	387	8400, 0
	MPC, ceiling 95 %	2642	4.42	3.19	0	72.80	445	8400, 0
	Override 99 % + charge feedforward	2340	5.92	7.86	77	73.02	297	–
	Override 99 %	1705	508.2	62.6	1082	73.04	320	–
9 (504 h)	Override	2544	559.1	97.4	572	73.73	816	–
	MPC	2427	485.3	62.6	1384	73.87	369	8400, 0
	MPC, ceiling 95 %	2623	595.2	76.2	0	73.66	367	8400, 0
	Override 99 % + charge feedforward	2443	485.4	62.6	574	73.85	321	–
	Override 99 %	1747	1098.7	133.4	997	73.85	522	–
10 (12 h)	+ Feedforward, and both rows at 99 %	0	2.61	7.11	0	1.63	0	–
	Override	0	2.61	7.11	0	1.63	2	–
	MPC, at either ceiling	3.6	3.00	7.63	0	1.62	28	200, 0
11 (24 h)	+ Feedforward, Override, both rows at 99 %	0	141.1	7.18	0	3.13	0	–
	MPC, at either ceiling	10.2	142.8	6.95	0	3.12	288	400, 0
14 (504 h)	Single, shipped	0	3165.7	97.6	14180	72.00	0	–
	Override, suggested	7194	86.4	52.2	205	69.94	629	–
	MPC	4131	13.6	8.13	14423	72.03	475	8400, 0
	MPC, ceiling 95 %	7312	13.8	8.13	0	69.91	512	8400, 0
	Override 99 % + charge feedforward	4168	426.4	25.0	1381	71.71	455	–
	Override 99 %	3677	827.6	82.5	1381	71.71	454	–

Not processed is $\int (F_{dem} - F_c)\, dt$. IAE is the scorecard's integral on the true column inlet and worst its largest deviation from setpoint. Valve on stop counts the minutes with the fuel valve at or above 99.5 %, the scorecard's firing-limit test. Charge travel sums the changes of the charge command over the logged rows. The MPC rows decide every 0.06 h at scenario 13's settings, and on scenarios 10 and 11 its two ceilings give identical runs.

and the override 95.4 % at 296.4 t/h, as shown in Figure 3.4. Over 48 h the plan processes 17.703 kt against 17.616 kt at an integrated error of 2.57 against 59.35 °C h.

The extra tonnes are the last five points of valve rather than the optimization. With its ceiling at 95 % the plan leaves 1676 t unprocessed on scenario 13 against the override's 1584 t, so those five points are worth 179 t over the outage, and at equal authority the plan gives 92 t back by cutting at once rather than letting the column run 42.5 °C cold. The three runs of the scenario's completion rule, read at hour 26, give 8.903 kt, 8.816 kt and 8.724 kt processed under the plan, the override and the plan at a 95 % ceiling, the same 87 t gain made of the same 179 t less 92 t. At that ceiling it processes less than the override on every case with a clean, 564 against 555 t unprocessed on scenarios 5 and 6, 2642 against 2539 on scenario 8, 2623 against 2544 on scenario 9 and 7312 against 7194 on the challenge. The shallow dip comes from acting on the measured furnace inlet. The override given the balance as a charge feedforward at its 95 % target dips 4.64 °C on scenario 13 against the plan's 3.02 °C, and at a 99 % target it comes within 28 t and 1.8 °C h of the plan. On the challenge that structure comes within 37 t but loses 413 °C h, because a balance written at the design heating value cannot see the gas that turned lean at hour 100, while the plan's output bias carries it. With that event removed from the challenge the gap is 59 t and 1.5 °C h.

Override at 99 % alone beats the plan on tonnes on scenarios 7, 8, 9 and the challenge, though not on scenarios 5 and 6 (393 against 386 t), with 1003 against 1497 t unprocessed on scenario 13 and 3677 against 4131 t on the challenge, by letting the column run as much as 65.1 °C cold at an

integrated error of 427 against 2.57 °C h. It even passes the challenge, with 197.9 kt, a worst deviation of 0.5 °C after the cleans and the valve at 99.0 %, because the temperature clause is read only once both exchangers are back. It is a third structure rather than a recommendation, its tonnes paid for in column inlet temperature that a single score would conceal.

The plan loses scenario 10, at 3.00 against 2.61 °C h and a worst deviation of 7.63 against 7.11 °C, and gives up 3.6 t by dipping the charge to 386.2 t/h with the valve at 86 %. An unmeasured heating value step reaches the plan only through the bias, filtered over 0.06 h and acted on every 0.06 h, where TIC-101 acts every 0.02 h. Counterfactual copies isolate the cause: a pinned charge gives 3.05 °C h, a decision every scan 2.35, a 0.02 h bias filter 2.28 and both together 1.63. On scenario 11 both structures hold the biased reading, at 142.8 against 141.1 °C h, and the plan gives up 10.2 t, of which 8.3 t answer the 0.6 °C noise with the charge (1.9 t without noise).

On the challenge the plan meets the temperature clause, at 1.0 °C worst after the cleans, and the tonnage clause, at 197.5 kt against the suggested override's 194.4 kt, but fails the requirement that the fuel valve finish off its stop, at 100.0 %, because the plan spends the whole valve by design. At a 95 % ceiling it passes with 194.3 kt, 0.9 °C and 95.0 %. The scorecard's firing-limit fraction reads 33.1 % of scenario 13 under the plan against 4.4 % under the override, and 47.7 % of the challenge against 0.7 %. Under the plan that time on the stop is chosen, not authority lost. None of the 71 600 plans in the runs behind Table 3.8 was rejected.

3.5.7 Implementation

Manual and automatic modes are provided, with operator helpers that start a cleaning outage (Clean E-101, Clean E-102), evaluate the heat balance of (3.10) (Balance the furnace), and solve for the highest feasible charge rate in the current exchanger condition (Feasible charge rate). Seeds, Reset and the exported file work as Sections 1.6.1 and 1.6.4 describe, so every run repeats exactly. The application also runs thirty-three internal verification checks, seventeen of them on the predictive controller, covering the design steady state, agreement between $UA \cdot \mathrm{LMTD}$ and effectiveness-NTU to one part in ten thousand, the closed-form fouling integral over 504 h, the $v^{-1.4}$ scaling, sign conventions and the cleaning state machine. Figure 3.5 shows the recovery after a cleaning outage under override control.

3.6 Technology in Practice

Figure 3.6 draws the override structure of Section 3.4 with the hardware a refinery installs in place of the model's ideal instruments.

What Measures It

Both temperatures that the controllers use come from thermocouples in thermowells. The lags τ_{TT} and τ_{FT} of Table 3.2 stand for the wells, and the bias of scenario 11 for an error in the measuring circuit. The process variable of VPC-101 is no process measurement. In the general scheme it is the input that the first controller writes, held near a resting value on a slow time scale [22], and the model reads the valve position exactly.

> **TECHNOLOGY** **Sensor: Thermocouple in a Thermowell**
>
> A thermocouple is two wires of unlike metals joined at the hot end. The junction produces a voltage roughly proportional to temperature, but the connection to the instrument forms a second junction that opposes it, so the receiver must compensate for the temperature at its own terminals [25]. The wrong kind of extension wire moves that second junction, and the compensation is then made for the wrong temperature [25]. The result is a standing offset, and scenario 11 shows what a loop does with one: it holds the wrong temperature exactly. The sensor sits in a thermowell, a pressure-tight sheath that lets it be replaced with the process under pressure. The mass of the well adds a first-order lag [25], which the model carries as 0.01 h on the column inlet (Table 3.2).

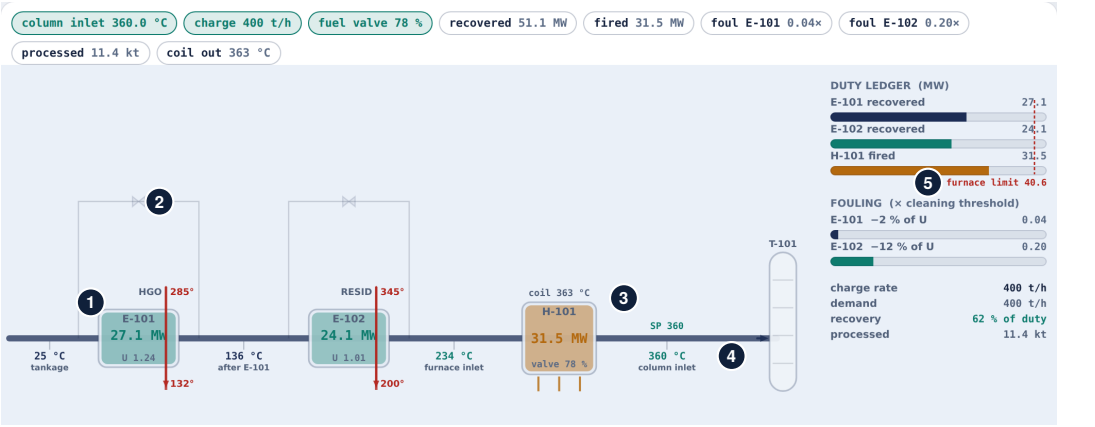


Figure 3.5 Scenario 6, recovery from a sixteen-hour E-101 cleaning outage under override control with the shipped settings, at $t = 30$ h and twelve hours after E-101 returned clean, cropped to the plant view with numbered callouts. E-101 (1) is back in service with its bypass (2) shut, the fired heater (3) supplies the rest of the duty, the column inlet (4) is on its setpoint, and the duty ledger (5) marks the furnace limit. In the full window, outside this crop, the 48 h trends show the charge rate cut to 363 t/h while the exchanger is out of service and back at 400 t/h by hour 18.6, the column inlet within 1.4 °C of 360 °C throughout, and the fuel valve settling at 77.7 % against the 81.3 % it opened on because the cleaning recovered 3.6 points of valve.

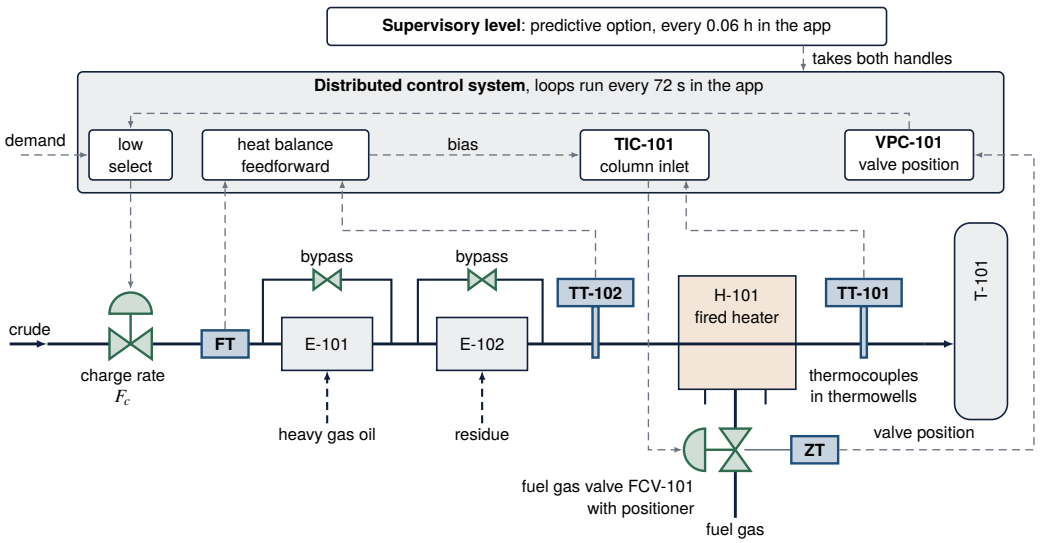


Figure 3.6 The loops of the override structure with their hardware. Lettered boxes are sensing devices, valve symbols are final elements and dashed arrows are signals. Notice that the measurement of VPC-101 is the fuel valve itself, read by a position transmitter (ZT), and that its output reaches the charge valve only through a low select against the operator's demand. Not to scale, and not a piping or safety diagram.

What Moves It

Three kinds of final element carry the chapter, the fuel gas valve FCV-101, the charge rate actuator and the two exchanger bypasses. The model gives each a first-order lag and no dead band or friction,

and its one nonlinearity is the one that matters here, a valve that stops at 100 %. Over 60 % of all fuel used in a refinery is burned in furnaces and boilers, and the average thermal efficiency of furnaces is estimated at 75 to 90 % [1], a range that holds the 0.88 of Table 3.2.

> **TECHNOLOGY** **Final Element: Control Valve With a Positioner**
>
> In a pneumatic control valve, air pressure on a diaphragm or piston works against a spring. Packing friction and the force of the fluid on the plug keep the stem from following the air pressure exactly, so a positioner is added, a device that compares stem position with the command and adjusts the air until they agree [25]. The spring sets the position on loss of air. An air-to-open valve fails closed, and the choice is to be made from the risk of the process, not from habit [25]. A burner management system has shutoff valves of its own to stop the fuel [25]. An electronic positioner also reports the stem position it measures and the total stem travel, like an odometer [25]. A valve position controller is a different thing. It uses a second input to keep the first from saturating [22].

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level sends set points to distributed controllers that act through the actuators from the feedback of the sensors, in cycles from milliseconds to minutes [26]. A survey of industrial practice puts basic control at every second and constraint control at every minute, and builds the conventional constraint layer from PID algorithms, lead-lag blocks and high and low select logic [21]. The structure of this chapter is of that kind, executed in the app every 0.02 h. A controller with integral action that a selector has de-selected tends to wind up, so an override needs limits on its integral [25]. Equation (3.16) gets them by making the operator's demand the upper output limit of VPC-101.

> **IN PRACTICE** **What the Operator Watches: Fouling Resistance and Furnace Inlet**
>
> The model reports its fouling resistances exactly. A refinery has to infer them from temperatures and flows, and only key exchangers carry all four stream temperatures [27]. The most popular indicator is, by far, the fouling resistance of Equation (3.1), computed through a lumped log-mean model like the one of this chapter, and typical measurement errors can put errors of the order of 20 % into it [27]. For the whole train the main indicators are the furnace inlet temperature and the furnace duty, the first often normalized to standard feed conditions. Cleaning needs crews, equipment and cranes, usually contracted at long notice, so operators ask for a stable schedule [28].

What Is Modernizing It

The first of three developments is predictive control, which replaces the blocks of the constraint layer with one controller [21]. Applications typically carry the valve positions of the loops below as extra controlled variables held inside their saturation limits [21], which is valve position control written as a constraint, the form of Section 3.4.4. The second is monitoring with detailed models. A fouling resistance cannot tell the growth of a deposit from a change in its properties, and dynamic thermal and hydraulic models of the exchangers have been proposed to separate them [27]. The third is the electronic positioner, whose sensors and digital communication give the maintenance staff and the control system diagnostic data on the valve [25].

3.7 Limitations

No result here transfers to a real preheat train. Two exchangers stand in for twenty, the fouling law depends only on velocity, nothing vaporizes, the column is a temperature target, and there is no crude margin, so a cleaning interval, an energy saving or the cost of one dirty exchanger are all outside what the model can state, and the fouling literature it draws on exists precisely because those questions are hard [14–16]. The economics are a single fuel gas price, so a decision can be reported in tonnes and gigajoules but the application cannot state which way the decision should go. No limit trips. The

coking limit, the firing limit and the deviation alarm all annunciate while the run continues, which is what lets a severe excursion be read to its end.

One temperature loop, one feedforward term and one constraint controller are the entire shipped control system, and the predictive option beside them plans the same two handles on one heat balance. A modern crude unit runs a multivariable controller with a real-time optimizer above it, and the constraint this application resolves with a valve position controller is exactly the kind that a predictive controller handles by writing it into the optimization, which is the reason that technology was adopted in refining in the first place [21]. Real constraint control on this unit manages firing, flooding, pump capacity, product specifications and tube skin temperatures at once, and the plantwide version pushes throughput against whichever unit is the current bottleneck rather than against one furnace [5]. What survives the simplification is the part a controller has to be designed around, and there are four such features: a process gain that moves with charge rate, an actuator with a hard end, a second manipulated variable that is expensive rather than merely available, and a disturbance that is entirely predictable and therefore answered by a schedule. Section 3.3.8 states where each assumption stops holding.

The documented model also extends beyond the implemented PID, feedforward and override structure to advanced control methods. The predictive option of Section 3.4.4 coordinates the fuel valve and the charge rate as one constrained problem instead of a low select, with the firing limit written in as a bound, and a controller built on the full twelve-state model could add the coking limit as a constraint [29]. The two fouling resistances are unmeasured states that moving horizon estimation or Kalman filtering can reconstruct from the measured temperatures and duties, which turns the soft-sensing argument of Section 3.5 into a formal estimator. An economic objective carrying a crude margin beside the fuel price would let the same controller weigh throughput against fuel continuously, which is the cleaning schedule argument posed as an optimization. The model equations transfer directly to the APMonitor MPC and MHE framework [29] and to the Gekko Optimization Suite in Python [30].

The predictive controller’s model leaves out the heating value, the change in the train’s charge response with fouling or a bypassed exchanger, the slowing of the coil and transfer line at lower charge, the thermocouple bias and the coking limit. The output bias absorbs the first three, and its 0.06 h filter with the 0.06 h decision is what loses scenario 10, at 3.00 against 2.61 °C h. One slack covers the whole horizon, so the floor is not a per-step guarantee. Its advantage over the override is not the optimization as such, because a valve position controller given the same valve range and the same balance as a charge feedforward comes within 28 t and 1.8 °C h on scenario 13, a result that leaves valve position control, whose design and tuning are established in industrial practice [31], competitive for this one constraint.

3.8 Discussion

The quantities the application reports are the ones the real scheduling argument is conducted in, and there are three of them rather than two. The fuel argument favors frequent cleaning, because a fouled exchanger burns gas continuously and carries the emissions that go with it [17], and the two cleanings of scenario 8 recover 3.9 TJ over three weeks. The throughput argument favors rare cleaning, because margin on crude is worth far more per tonne than gas is per gigajoule, and those same two cleanings cost 2.5 kt. Neither argument settles the schedule, because the deciding term is the constraint. An exchanger left too long takes the fuel valve to its stop, and from that point the throughput loss occurs without an operating decision: scenario 4 is that condition approaching with 2.5 % of valve left, and scenario 14 shows the condition after it arrives, at 46.9 % of three weeks on the stop. Schedules computed with throughput in the objective are not the schedules fuel alone produces [18, 20].

⚠ OPEN CHALLENGE Scheduling the Cleaning and Running the Loops Together

The cleaning schedule and the flow control of a preheat train have usually been worked out apart, by heuristics. The operation then reaches a limit such as the firing limit of the furnace, and an emergency cleaning or a cut in production follows [28]. Joining the two is hard. The time scales run from weeks to years, crude blends change within hours or days, and a schedule that changes too often cannot be resourced [28]. A 2020 study closed both loops with predictive control on a model validated against 1240 days of refinery data, but its plant was that model [28]. Progress would be the same loop on a running unit, with a schedule stable enough for the crews to follow.

What the control engineer brings to that decision is a set of measured quantities rather than a controller: the duty the exchanger is still recovering, the fuel valve position, the tonnes the override gave up the last time it was selected, and the length of the next run at the rate the unit is actually being asked to hold. The last of those carries a penalty the application measures, because a unit turned down to 75 % of design fouls toward an asymptotic resistance 1.50 times higher and reaches its threshold in 124 h instead of 240 h, so throughput given up in one month raises the fouling cost of the next.

The scorecard is built around the same argument. It reports the column inlet error, the crude processed, the average charge rate, the fuel fired with its cost, the heat recovered over the run and at the instant, the fraction of the run with the charge rate cut and the fraction at the firing limit, and it never combines them into one number. A run that scores well on temperature and badly on average charge rate has abandoned the problem rather than solved it, and the only way to make that visible is to report both halves and refuse to weight them. What a controller spends on a constraint is invisible unless something counts it, which is the reasoning behind reporting throughput beside the thermal objective in cleaning schedule work [18]. The exported file holds exactly what the trends were drawn from, so a claim about tonnes can be checked in a spreadsheet.

3.9 Conclusion

This chapter studied a crude preheat train and its fired heater as a laboratory for override control, where the manipulated variable of last resort is the plant's own throughput. The simulator integrates twelve states, four exchanger temperatures, the coil outlet, the column inlet, two fouling resistances and four actuators, with fouling entering as a series resistance whose asymptote scales as crude velocity to the -1.4 power. At the design point E-101 recovers 22.87 MW, E-102 recovers 26.79 MW and the furnace supplies 32.98 MW of a 40.56 MW capacity, so 60.1 % of the 82.63 MW duty is recovered from product streams and the fuel valve sits at 81.3 %. Three weeks of fouling never move the column inlet off 360.0 °C and leave the integrated error at 0.0 °C h, while the valve opens to 97.5 % and recovery falls to 52.1 %, which is the case for a performance report that carries more than error. Taking E-101 out for cleaning at full rate delivers 340.3 °C with the valve on its stop for 96.5 % of the outage, while the valve position override holds 360.0 °C for 555 t of crude out of 19 200 t, and losing E-102 instead costs about three times as much because nothing downstream of it picks up the loss. Moving one of two colliding cleaning outages 32 h later changes the integrated error over three weeks from 559 to 48 °C h and the crude processed by 17 t in 199 000, and shortening both outages to 8 h so that they no longer overlap brings it to 14 °C h, which is more than any tuning in the application can offer. A predictive controller offered as a fourth option processes 87 t more than the override on scenario 13 at an integrated error of 2.57 against 59.35 °C h, yet at the override's 95 % valve ceiling it processes 92 t less, so its throughput is the valve's last five points and its temperature the measured furnace inlet acting on the charge. A fuller model would add a column model that closes the loop between preheat and the hot streams, a wall-temperature-dependent fouling law, a crude margin beside the fuel price, and the coking limit as a constraint of the predictive controller. The simulator remains a place to study constraint control where the second manipulated variable is the plant throughput.

KEY POINTS

- A temperature that sits on its setpoint for three weeks can hide a furnace that is running out of room. Only the position of the fuel valve records it.
- With the fuel valve on its stop, no feedback tuning recovers duty the furnace cannot supply. One manipulated variable remains, the crude charge rate.
- Below the point at which the constraint binds, a tighter valve position target buys no temperature and costs tonnes.
- Separating two cleaning outages that could collide costs essentially nothing. The overlap is what the column cannot carry.
- The predictive controller's extra tonnes are the last five points of valve, not the optimization.

Where else this applies. Valve position control returns in Volume 4, Chapter 9, an override that protects a compressor from surge in Volume 4, Chapter 2, throughput as the override's handle in Volume 4, Chapter 4, and a heater's feedforward and limits in Volume 3, Chapter 4.

References

[1] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[2] F. G. Shinskey, *Process Control Systems: Application, Design, and Tuning*, 4th ed. New York: McGraw-Hill, 1996.

[3] T. E. Marlin, *Process Control: Designing Processes and Control Systems for Dynamic Performance*, 2nd ed. New York: McGraw-Hill, 2000.

[4] A. Reyes-Lúa and S. Skogestad, "Systematic design of active constraint switching using classical advanced control structures," *Industrial & Engineering Chemistry Research*, vol. 59, no. 6, pp. 2229–2241, 2020.

[5] E. M. B. Aske, S. Strand, and S. Skogestad, "Coordinator MPC for maximizing plant throughput," *Computers & Chemical Engineering*, vol. 32, no. 1-2, pp. 195–204, 2008.

[6] D. E. Seborg, T. F. Edgar, D. A. Mellichamp, and F. J. Doyle, III, *Process Dynamics and Control*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2016.

[7] B. W. Bequette, *Process Control: Modeling, Design, and Simulation*. Upper Saddle River, NJ: Prentice Hall, 2003.

[8] K. J. Åström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*, 2nd ed. Princeton, NJ: Princeton University Press, 2021.

[9] A. Szklo and R. Schaeffer, "Fuel specification, energy consumption and CO2 emission in oil refineries," *Energy*, vol. 32, no. 7, pp. 1075–1092, 2007.

[10] B. Linnhoff and J. R. Flower, "Synthesis of heat exchanger networks: I. systematic generation of energy optimal networks," *AIChE Journal*, vol. 24, no. 4, pp. 633–642, 1978.

[11] H. Müller-Steinhagen, "Heat transfer fouling: 50 years after the Kern and Seaton model," *Heat Transfer Engineering*, vol. 32, no. 1, pp. 1–13, 2011.

[12] T. R. Bott, *Fouling of Heat Exchangers*. Amsterdam: Elsevier, 1995.

[13] W. Ebert and C. B. Panchal, "Analysis of Exxon crude-oil-slip stream coking data," Argonne National Laboratory, Argonne, IL, Tech. Rep. ANL/ES/CP-92175, 1995, presented at Fouling Mitigation of Industrial Heat Exchangers, San Luis Obispo, CA, June 18–23, 1995. OSTI identifier 453433.

[14] G. T. Polley, D. I. Wilson, B. L. Yeap, and S. J. Pugh, "Evaluation of laboratory crude oil threshold fouling data for application to refinery pre-heat trains," *Applied Thermal Engineering*, vol. 22, no. 7, pp. 777–788, 2002.

[15] B. L. Yeap, D. I. Wilson, G. T. Polley, and S. J. Pugh, "Mitigation of crude oil refinery heat exchanger fouling through retrofits based on thermo-hydraulic fouling models," *Chemical Engineering Research and Design*, vol. 82, no. 1, pp. 53–71, 2004.

[16] F. Coletti and S. Macchietto, "A dynamic, distributed model of shell-and-tube heat exchangers undergoing crude oil fouling," *Industrial & Engineering Chemistry Research*, vol. 50, no. 8, pp. 4515–4533, 2011.

[17] H. Müller-Steinhagen, M. R. Malayeri, and A. P. Watkinson, "Heat exchanger fouling: Environmental impacts," *Heat Transfer Engineering*, vol. 30, no. 10-11, pp. 773–776, 2009.

[18] E. M. Ishiyama, W. R. Paterson, and D. I. Wilson, "The effect of fouling on heat transfer, pressure drop, and throughput in refinery preheat trains: Optimization of cleaning schedules," *Heat Transfer Engineering*, vol. 30, no. 10-11, pp. 805–814, 2009.

[19] M. C. Georgiadis, L. G. Papageorgiou, and S. Macchietto, "Optimal cleaning policies in heat exchanger networks under rapid fouling," *Industrial & Engineering Chemistry Research*, vol. 39, no. 2, pp. 441–454, 2000.

[20] F. Smaïli, V. S. Vassiliadis, and D. I. Wilson, "Mitigation of fouling in refinery heat exchanger networks by optimal management of cleaning," *Energy & Fuels*, vol. 15, no. 5, pp. 1038–1056, 2001.

[21] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[22] S. Skogestad, "Advanced control using decomposition and simple elements," *Annual Reviews in Control*, vol. 56, p. 100903, 2023.

[23] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[24] K. R. Muske and T. A. Badgwell, "Disturbance modeling for offset-free linear model predictive control," *Journal of Process Control*, vol. 12, no. 5, pp. 617–632, 2002.

[25] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[26] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[27] E. Diaz-Bejarano, F. Coletti, and S. Macchietto, "A model-based method for visualization, monitoring, and diagnosis of fouling in heat exchangers," *Industrial & Engineering Chemistry Research*, vol. 59, no. 10, pp. 4602–4619, 2020.

[28] F. Lozano Santamaria and S. Macchietto, "Stability of optimal closed-loop cleaning scheduling and control with application to heat exchanger networks under fouling," *Processes*, vol. 8, no. 12, p. 1623, 2020.

[29] J. D. Hedengren, R. Asgharzadeh Shishavan, K. M. Powell, and T. F. Edgar, "Nonlinear modeling, estimation and predictive control in APMonitor," *Computers & Chemical Engineering*, vol. 70, pp. 133–148, 2014.

[30] L. D. R. Beal, D. C. Hill, R. A. Martin, and J. D. Hedengren, "GEKKO optimization suite," *Processes*, vol. 6, no. 8, p. 106, 2018.

[31] B. J. Allison and S. Ogawa, "Design and tuning of valve position controllers with industrial applications," *Transactions of the Institute of Measurement and Control*, vol. 25, no. 1, pp. 3–16, 2003.

3.10 Exercises

3.1 Concept check · 2 min

In scenario 4 the column inlet sits on 360.0 °C for three weeks and the integrated absolute error is 0.0 °C h, while E-101 and E-102 foul. Which reading on the panel shows that the unit is losing its margin?

(a) The fuel valve position, which opens without a step from about 81 to about 97 %

(b) The column inlet error tile, which grows a little each day as the two exchangers foul

(c) The integrated absolute error, which counts the duty the fouled exchangers no longer recover

(d) The average charge rate, which the feedforward trims down as the duty falls off

3.2 Concept check · 2 min

Equation (3.7) is the energy balance on the furnace coil. Which quantity in it does the feedforward of Equation (3.15) have no measurement of, and which scenario shows what that costs?

(a) The furnace inlet T_2, so the rate cut of scenario 3 defeats the feedforward

(b) The heating value ratio HV/HV_0, so the lean fuel gas of scenario 10 defeats it

(c) The charge rate F_c, so the three weeks of fouling in scenario 4 defeat it

(d) The coil capacitance C_{fur}, so the E-101 outage of scenario 5 defeats it

3.3 Concept check · 2 min

In scenario 6 the override holds the column inlet through the E-101 outage with its valve position target at 95 %. The target is set to 85 % before Run and the 48 h are run again. What changes?

(a) The worst column inlet improves by several degrees, and the crude processed is unchanged

(b) The override never acts, and the run repeats scenario 5 with the valve on its stop

(c) The worst column inlet is about the same, and roughly 370 t more crude is given up

(d) The override starts cutting at once, because the valve at 81.3 % is near the target

3.4 Theory and calculation · 7 min

Use Equations (3.1) and (3.2) with the E-101 values of Table 3.2: $U_{clean,1} = 1.26$, $R_{f\infty,1} = 0.55$, $\tau_{f,1} = 420\,\mathrm{h}$ and a cleaning threshold of 0.42, in the working units. At constant velocity the resistance relaxes as $R_f(t) = R_{f\infty}v^{-1.4} - \left(R_{f\infty}v^{-1.4} - R_f(0)\right)\mathrm{e}^{-tv/\tau_f}$.

(a) Find the coefficient of E-101 at its threshold and the share of the clean coefficient it has lost. *Checkpoint:* 0.824 and 34.6 %.

(b) The run opens with $R_{f,1} = 0.320$ at the design rate. Find the hour E-101 reaches its threshold. *Checkpoint:* 239.6 h.

(c) Repeat (b) at 300 t/h, finding the asymptote and the time constant first. *Checkpoint:* 0.823, 560 h and 124.2 h.

3.5 Theory and calculation · 6 min

The design condition has $F_c = 400\,\mathrm{t/h}$, $c_{p,c} = 2.2\,\mathrm{MJ/t/K}$, $T_{SP} = 360\,^\circ\mathrm{C}$, $UA_{tl} = 8\,\mathrm{MJ/h/K}$, $T_{air} = 25\,^\circ\mathrm{C}$, $Q_{fur,max} = 146\,000\,\mathrm{MJ/h}$ and a furnace inlet of 228.14 °C.

(a) Use Equation (3.10) to find the coil outlet the column needs, the transfer line loss and the starting valve position. *Checkpoint:* 363.05 °C, about 0.74 MW and 81.3 %.

(b) Find the steady gain from valve position to column inlet. *Checkpoint:* 1.644 °C per percent.

(c) In scenario 11, averaged over hours 12 to 24, the valve sits at 78.91 % with the thermocouple bias and at 82.56 % without it. Find the bias in degrees, and say where the true column inlet is. *Checkpoint:* 6.0 °C, true inlet near 354 °C.

3.6 Theory and calculation · 6 min

Before the E-101 outage of scenario 5, E-101 recovers 22.87 MW, E-102 26.79 MW and the furnace 32.98 MW of its 40.56 MW. At hour 6, with E-101 bypassed and the valve on its stop, E-102 recovers 37.64 MW.

(a) Find the duty the train loses, the share the furnace covers and the shortfall left over. *Checkpoint:* 12.02 MW, 7.58 MW and 4.44 MW.

(b) Find the column inlet the shortfall leaves at hour 6, with the flowing heat capacity and the transfer line loss coefficient together 888 MJ/h/K. *Checkpoint:* 18.0 °C short, about 342 °C.

(c) In the E-102 outage the crude reaches the coil at 124.6 °C. Find the charge rate at which the full furnace holds 360 °C if the furnace inlet stayed at that value, and compare it with the 295 t/h the override settles at. *Checkpoint:* about 277 t/h.

3.7 App challenge · 8 min

Load scenario 4, “Three weeks of fouling, nothing else”, at the default seed, set the speed to 100 and the trend window to 504 h.

(a) Run the full 504 h. Read the fuel valve position at 504 h (CSV column `fuel_valve_pct`), the Fuel fired tile, and the hour at which E-101 first reaches 1.0 on the fouling trend (column `foul_e101_frac`). *Checkpoint:* 97.5 %, 76.7 TJ and 239.6 h (±0.1 h).

(b) Press Reset, set Crude charge demand to 420 t/h before Run, and run the 504 h again. Find the hour the fuel valve first reaches its stop (FURNACE AT FIRING LIMIT on the trend, or the first 1 in `fuel_saturated`) and the column inlet at 504 h. *Checkpoint:* 215.2 h (±0.1 h) and 352.1 °C (±0.1 °C).

(c) In the 420 t/h run E-101 reaches its threshold at 290.5 h, later than in (a). Use Equation (3.2) to explain why, and why the valve still reaches its stop sooner. *Checkpoint:* $v = 1.05$: asymptote 0.514, time constant 400 h.

3.8 App challenge · 9 min

Compare the E-101 outage without and with the override, at the default seed, running each case 48 h at speed 10. Read the lowest column inlet from the trend or from the minimum of the CSV column `col_inlet_C`, and the crude processed from `crude_total_t` at 48 h.

(a) Run scenario 5, "Clean E-101 at full rate, with feedforward only". Read the lowest column inlet, the crude processed and the At firing limit tile. *Checkpoint:* 340.3 °C, 19 200 t and 33.3 % of the run.

(b) Run scenario 6, "The same clean under override control". Read the same values, the Rate cut tile and the lowest charge rate (`crude_rate_tph`). What did holding the temperature cost? *Checkpoint:* 358.6 °C, 18 645 t (±2 t), 363 t/h, so 555 t or 2.9 %.

(c) Press Reset, set Valve position target (Advanced) to 85 % before Run, and run again. Then do the same at 100 %. Which of the three targets would you commission? *Checkpoint:* 358.8 °C and 18 279 t at 85 %. At 100 % the run repeats (a).

3.9 App challenge · 8 min

Load scenario 9, "The same two cleanings, colliding", at the default seed, set the speed to 100 and run the full 504 h.

(a) Read the scorecard IAE, the lowest charge rate (`crude_rate_tph`), the lowest column inlet (`col_inlet_C`) and the crude processed (`crude_total_t`) at 504 h. *Checkpoint:* 559.1 °C h, 240 t/h, 262.6 °C and 199 056 t.

(b) Press Reset, set Cleaning duration in Show Model, Parameters, to 8 h before Run, and run the 504 h again. Read the same four values. *Checkpoint:* 13.9 °C h, 290.5 t/h, 344.9 °C and 200 315 t.

(c) The cleans are scripted at hours 240 and 248. Find how long the two outages overlap in each run, and explain why the integrated error falls by a factor of about forty. *Checkpoint:* 8 h of overlap, then none.

3.10 Application area · 4 min

Scenario 8 cleans both exchangers at their thresholds and processes 199.1 kt with 72.8 TJ of fuel. Scenario 4, the same three weeks with no cleaning, processes 201.6 kt with 76.7 TJ. Fuel gas costs $6/GJ.

(a) Find the fuel the two cleans save and its value. *Checkpoint:* 3.9 TJ, about $23k.

(b) Find the crude margin per tonne at which the fuel saved just pays for the crude given up. *Checkpoint:* about $9.4 per tonne.

(c) Name two things this break-even leaves out that would change the decision on a real unit.

3.11 Application area · 5 min

The model cleans perfectly. Mechanical cleaning of a real exchanger typically recovers 85 to 95 % of the clean coefficient.

(a) E-101 ($U_{clean,1} = 1.26$, threshold 0.42) is cleaned to 90 % of its clean coefficient. Find the residual resistance and its share of the threshold. *Checkpoint:* 0.088 and 21 %.

(b) Find the hours from that clean to the next threshold at the design rate ($R_{f\infty,1} = 0.55$, $\tau_{f,1} = 420\,\text{h}$), against 606 h from a perfect clean. *Checkpoint:* about 532 h.

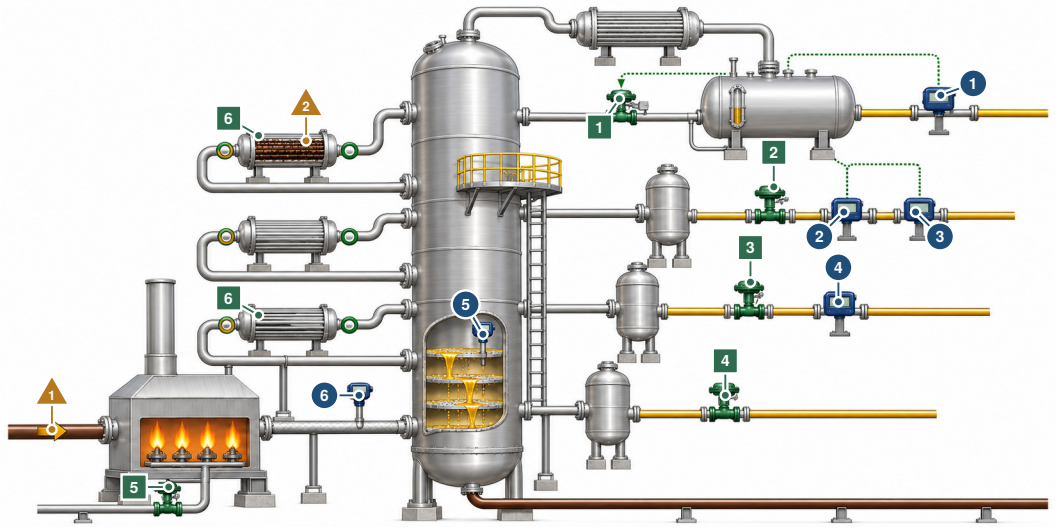
(c) What does this do to a cleaning schedule over several runs, and to the chance that two cycles collide?

3.12 Review the agent · 5 min

After running scenario 13 for 48 h, an AI assistant writes: “The MPC processed 17.703 kt against the override’s 17.616 kt and held the column inlet near 357 °C where the override dipped to 317.5 °C. The optimization finds 87 t the override cannot. Replace VPC-101 with the MPC, and for safety keep the override’s 5 % of valve in hand by setting the plan’s valve ceiling to 95 %.” Find the flaw, and say how you would check it in the app.

CHAPTER 4

A Crude Atmospheric Tower Simulator for Degrees of Freedom and Triangular Interaction



The controller moves 1 reflux valve 2 kerosene draw valve 3 diesel draw valve 4 gas oil draw valve 5 heater firing 6 heat taken out by the three pumparounds

The controller reads 1 naphtha end point 2 kerosene flash point 3 kerosene freeze point 4 diesel cloud point 5 overflash, the liquid that wets the wash zone 6 heater outlet temperature

What disturbs it ▲1 a heavier crude ▲2 a fouled top pumparound ▲ more charge

The dotted signal lines from the naphtha analyzer at the upper right and from the kerosene analyzers below it lead back toward the overhead drum and its one reflux valve, so only one of the readings can be held. Illustration, not to scale.

IN PLAIN TERMS Four Fuels, Four Flows, Eight Promises

The main idea. A crude tower boils crude oil and splits it into four fuels, from a light one near the top to a heavy one lower down. Operators decide where one fuel ends and the next begins by setting four flows.

The challenge. Each fuel must pass two tests, so there are eight promises and only four flows to keep them. It is like slicing a loaf from one end. A thicker slice moves every cut after it and none before it.

Why it matters. Only half of the promises can be held by control, and a plant that tries to hold all eight gets controllers that fight each other.

THREE WAYS TO READ THIS CHAPTER

Explore	10 min	This page, the Did you know? boxes in Sections 4.1 and 4.5.7, the History boxes in Sections 4.2 and 4.4, the Try it box in Section 4.5, and Key points in Section 4.9. No equations.
Study	45 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 4.5 with Figure 4.3, Section 4.6, Section 4.7 and the Open challenge box in Section 4.8.

The control question A crude tower cuts one boiling curve into four products, each with two end specifications, using four handles. Which four of the eight numbers can be held, which way do the draws interact, and in what order should the loops be brought into service?

Learning objectives

- Count the degrees of freedom of a tower that makes products in series, and show that the naphtha end point and the kerosene flash point are one cut point.
- Compute cut points, tails and the overflash from the assay and the material and heat balances.
- Derive the gains and the inverse response of the top cut from the reflux power law and the tail.
- Explain why the gas oil draw is paired with the overflash and how the override protects the wash zone.
- Choose a commissioning order from the direction of interaction, and compare four loops with one predictive controller on measured runs.

Prerequisites Sections 1.4.3 (cut points and degrees of freedom), 1.4.2 (heater balance), 1.6.2 (PI loops), 1.6.3 (predictive control) and 1.6.4 (cases).

The app https://apmonitor.com/apps/crude-tower/

See also Volume 6, Chapter 2 for pairing on a triangular gain matrix and Volume 6, Chapter 7 for choosing which flow a structure holds. Volume 4, Chapter 2 for a protective override and Volume 4, Chapter 4 for throughput pushed against machine limits.

4.1 Introduction

A crude atmospheric tower is the first unit in a refinery and the last one a control course reaches, and the reason is that it does not fit the shape a course teaches with. Four products leave the tower as consecutive slices of one boiling curve. Each carries two specifications, an end point and a property such as a flash point or a freeze point, so there are eight numbers to hold. There are four places the curve can be cut. A student who counts the loops before counting the degrees of freedom will pair a controller with the naphtha end point and another with the kerosene flash point and then discover that the two are the same variable. Measured on the tower modeled here, taking the top reflux from 135 t/h to 170 t/h moves the naphtha end point from 191.4 °C to 176.0 °C and the kerosene flash point from 58.2 °C to 48.8 °C. Not identical, and not independent either. Two loops on one valve is the outcome, and on a panel it looks like a loop that has stopped working.

The misconception is that a specification is a controlled variable. On a unit that makes products in series from a single feed, half the specifications are consequences of the other half, and which half is which is an operating decision that changes with the season and the market. Hold the naphtha end point when the reformer is the constraint. Hold the kerosene flash point when jet fuel is worth more than gasoline. What cannot be done is to hold both, and no amount of tuning changes that. Counting independent handles before pairing loops is the cheapest engineering available on any unit, and it is the step that a two-input two-output example does not motivate.

DID YOU KNOW? Thirty to Fifty Trays in a Steel Cylinder

A crude tower is a steel cylinder about 120 feet high [1] that holds 30 to 50 fractionation trays [2]. All the crude oil a refinery runs passes through it first. On 1 January 2026 the 130 operable refineries of the United States could charge 19.2 million barrels of crude oil per stream day [3].

The second structural fact is that the interaction here runs one way. A draw moves its own boundary and every boundary below it and nothing above it, which makes the plant lower triangular. Measured, twenty more tonnes an hour of diesel leaves the kerosene cut point exactly where it was, while twenty more tonnes an hour of kerosene moves the diesel cut point by 17.6 °C. Towers with side strippers and side rectifiers have this structure generally [4], and so do many other units that separate a feed into consecutive fractions. Section 1.4.3 derives both facts from the material balance of the running example.

This chapter works with an interactive crude tower whose assay, tail sharpening and heat balance are hidden, so that the degrees of freedom must be counted from experiment, the triangular structure discovered rather than asserted, and a commissioning order chosen with a measurement to check it against. The model is an assay-sliced tower with its pumparound heat recovery, its furnace and its overflash override, written out in the detail needed to rebuild it. The scenario sequence contains a commissioning result that contradicts a standard heuristic and is measured rather than argued. The application runs in a browser without installation and exports its complete time history. The assay model, pumparound heat recovery and override control are established practice. Section 4.2 gives the domain history. Section 4.3 states the model: the cut points on the assay, the tails and the internal reflux, the heat balance and the overflash. Section 4.4 describes the four loops and their directions, the top loop selector, the overflash override and the optional predictive controller. Section 4.5 reports the measured results, from the degrees of freedom and the triangular interaction through the commissioning order, the pumparounds, the constraint and the graded challenge to the comparison with predictive control. Section 4.6 sets the loops beside the instruments, flow loops and control system of a real unit, Section 4.7 states the limitations, Section 4.8 the three results that generalize and Section 4.9 the conclusion. The exercises at the end of the chapter work through the same scenarios.

4.2 Background

Crude oil is a continuum of hydrocarbons and a refinery begins by cutting it into fractions by boiling range. The atmospheric tower does that first cut, and everything downstream, the reformer, the hydrotreaters, the cracking units, is fed by one of its draws. The measurement that describes the feed is the true boiling point assay, a laboratory distillation curve that reports the temperature at which each cumulative volume fraction has boiled. Characterizing a crude is still an active measurement problem, with spectroscopic methods used to predict the curve without running the distillation [5], and the product specifications are written on a different and cheaper laboratory distillation, ASTM D86, whose relationship to the true curve has its own literature [6].

HISTORY 1908 to 1922: The Pipe Still Takes Over From the Bench of Stills

Early continuous refineries passed crude oil through a bench of stills, as many as twelve or more. A book of 1922 records that tubular stills had largely replaced the bench in recent years [7]. One of the first plants to top crude by other means than the conventional stills was erected in 1908. In the Trumble system that the book describes, the crude was heated against hot distillates and residue, passed through a coil of pipes in a furnace and entered a single evaporator, a vessel 6 feet in diameter and 25 feet high [7]. The vapors left together and were condensed in stages into seven fractions, whose quality could be varied to some extent through the steam and cooling water of each stage [7]. The furnace, the single tower and the preheat against hot products are those of this chapter, but the side cuts were still condensed outside the tower.

Table 4.1 Classification of Variables at the Design Operating Point

Symbol	Role	Nominal	Range
EP naphtha	CV or consequence	184.1 °C	150–215 °C
FP kerosene	CV or consequence	53.9 °C	38–75 °C
FR kerosene freeze	CV	−47.0 °C	−58 to −34 °C
CP diesel cloud	CV	−1.8 °C	−14 to 8 °C
Φ overflash	CV, constraint	3.28 vol%	> 2 vol%
L top reflux	MV	150 t/h	60–320 t/h
D_K	MV	74 t/h	20–160 t/h
D_D	MV	102 t/h	30–200 t/h
D_G	MV	51 t/h	0–130 t/h
COT	MV, economic	377 °C	340–400 °C
$Q_{TPA,KPA,DPA}$	MV, economic	18/14/10 MW	≤35/30/25 MW
charge	DV, measured	662 m^3/h	400–850 m^3/h
crude heaviness	DV, unmeasured	0	−0.4 to 0.8
Q_F	constraint	76.5 MW	≤ 85 MW

Ranges are the spans the application accepts, as setpoints for the controlled variables and as valve or slider positions for the others.

Control of the tower developed from tray temperatures. Before continuous analysis, each draw was set from a tray temperature and adjusted when the laboratory reported a product outside its specification, with the cut point inferred rather than measured. That arrangement is workable because the temperature at a draw tray is close to a cut point, and it fails whenever the crude changes, because a different crude has a different relationship between temperature and composition at the same tray. The modern answer is either an inferential built from several temperatures and flows or a direct analyzer, and in both cases the interesting work is the layer above (Section 1.4.6).

Multivariable control arrived on crude units early and for a specific reason. The tower has more measured qualities than handles, several hard constraints including the furnace and the wash zone, and an economic objective that changes with the crude slate, which is exactly the problem a model predictive controller with an inequality-constrained objective is built for [8]. Dynamic simulation and nonlinear model-based quality control of a crude tower were demonstrated in the mid 1990s [9], and subsequent work has concentrated on tuning those controllers [10], on integrating them with real time optimization [11], and on the reduced models a planning layer needs above them [12, 13]. Surrogate models for design optimization of the same unit are a current subject [14]. The regulatory layer underneath all of it is still four loops and an override, and it is that layer this application is about. One such multivariable controller is also offered as an option above that layer, and Section 4.5.7 measures it against the four loops.

Pumparound heat recovery arrived with energy prices. A pumparound withdraws hot liquid from a tray, exchanges it against incoming crude and returns it cooler, which both recovers heat and condenses vapor. Measured on this tower with all four loops in automatic, taking the top pumparound from 12 MW to 26 MW raises the crude preheat from 238.7 °C to 264.2 °C and cuts the furnace firing from 81.0 MW to 70.5 MW, so 10.5 MW of fuel follow from one setpoint. It is the largest single energy lever on the unit and it is not free, for reasons Section 4.7 states. The preheat train that receives this heat, and the fouling that erodes it, are the subject of Chapter 3.

4.3 Process Model

Tower T-101 charges 662 m^3/h of a light sweet crude through fired heater F-101 to a flash zone, and makes overhead naphtha, kerosene, diesel and atmospheric gas oil with a residue from the bottom. Three pumparounds recover heat into the crude preheat train. Figure 4.1 shows the process arrangement and principal manipulated variables.

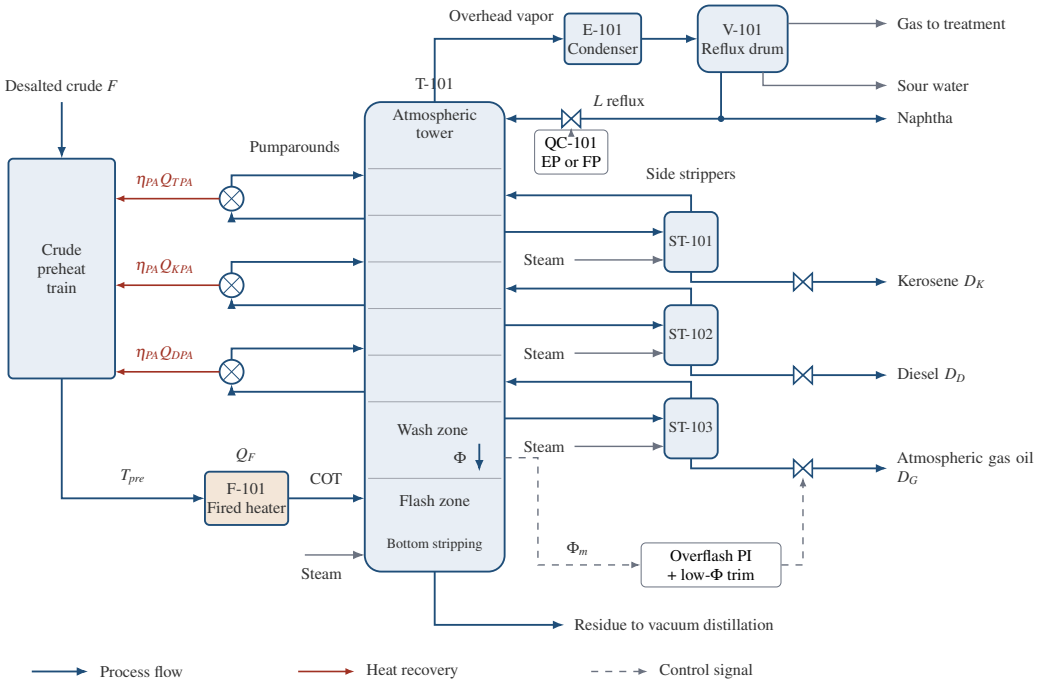


Figure 4.1 Simplified process and control schematic of atmospheric tower T-101. Desalted charge F receives a recovered fraction η_{PA} of the heat removed by three separate pumparound circuits (Q_{TPA}, Q_{KPA} and Q_{DPA}), each returning cooled liquid above its withdrawal tray. Preheated crude at T_{pre} passes through F-101, with firing duty Q_F, and enters the flash zone after reaching coil outlet temperature COT. Condenser E-101 and drum V-101 separate overhead gas, sour water and hydrocarbon liquid, which splits into naphtha and reflux L. QC-101 adjusts L using either naphtha end point EP or kerosene flash point FP. Side strippers receive steam, return vapor above their respective draw trays, and discharge kerosene D_K, diesel D_D and atmospheric gas oil D_G. Overflash Φ is the internal liquid descending through the wash zone. Its measured estimate Φ_m feeds the gas oil PI controller and a downstream protective trim that reduces D_G at low overflash in both Manual and Auto. The kerosene and diesel quality loops, pumps, level and pressure controls, and downstream product cooling are omitted for clarity. Steam and overhead separation are shown as process context, not as independently simulated units.

4.3.1 Cut Points and the Assay

The crude is described by a true boiling point curve $T_{bp}(v)$ relating cumulative volume fraction to boiling temperature, interpolated linearly between anchor points from 20 °C at $v = 0$ to 750 °C at $v = 1$. A heaviness h raises the curve by $h\,(40 - 25v)$ °C, more at the light end than at the heavy end. Four cut fractions $v_1 \dots v_4$ divide it, and each product is the slice between two of them. The naphtha cut is placed by the top reflux, Equation (4.1), which is a power law fitted so that more reflux pulls the naphtha cut down.

$$v_1 = v_{1,ref}\left(\frac{L_{ref}}{L}\right)^{p}(1 + 0.35\,h) \tag{4.1}$$

The remaining three are placed by the material balance. The cut below draw i lies at the naphtha fraction plus the volume of every draw above it, and each cut point T_i follows its place on the assay with the first-order lag τ_i, Equation (4.2).

$$v_{i+1} = v_1 + \sum_{j \le i}\frac{D_j}{\rho\,F}, \qquad \tau_i\,\frac{dT_i}{dt} = T_{bp}(v_i) - T_i \tag{4.2}$$

where D_1, D_2 and D_3 are the kerosene, diesel and gas oil draws (D_K, D_D, D_G), F is the volumetric charge and $\rho = 860\,\mathrm{kg/m^3}$ is one density for the crude and all of its products. That equation is the triangular structure. A draw enters the cut points below it and none above it, as in Equation (1.4), which writes the same balance with the naphtha counted as the first draw.

Pause. Try Exercise 4.2 before reading on.

4.3.2 Tails and Internal Reflux

A cut is not sharp. Each product carries a tail of material from the fraction on either side, and the tail width narrows as the internal reflux in that section rises, Equation (4.3).

$$\sigma_i = \frac{\beta}{\max(0.2, R_i)^n}, \qquad R_i = \frac{L_i}{D_i} \tag{4.3}$$

where σ_i is the tail spread in degrees Celsius, D_i is the product drawn at the top of section i (the overhead naphtha for the top section), $n = 1$, and L_i is the liquid running down through section i. The liquid starts as the top reflux and, section by section, gains the vapor condensed by the pumparound above it, Q_{PA}/λ with $\lambda = 250\,\mathrm{kJ/kg}$ (14.4 t/h per megawatt), and loses the draw taken off at its top. Product properties are linear correlations on the cut points and tails. The naphtha end point is $T_1 + \sigma_1$, the kerosene flash point is $T_1 - \sigma_1 - 112\,^\circ\mathrm{C}$, and the kerosene freeze point and the diesel cloud point are linear in the 95 percent points $T_2 + \sigma_2$ and $T_3 + \sigma_3$.

This is where a draw does its second job. Liquid drawn off at a tray no longer runs down through the trays below it, so a larger draw thins the section underneath and widens its own tail. Measured on the diesel draw, thirty more tonnes an hour moves the diesel and gas oil boundary from 337.6 °C to 363.9 °C, a shift of 26.4, while the diesel 95 percent point moves from 340.4 °C to 367.8 °C, which is 27.4. The extra degree is the tail, which widened from 2.82 °C to 3.91 °C. The second effect is about four percent of the first, and it means the loop gain is not constant.

4.3.3 Heat Balance and the Overflash

The furnace heats the preheated crude to the flash zone and vaporizes part of it, Equation (4.4), and the pumparounds set the preheat, Equation (4.5).

$$Q_F = \dot{m}\left[c_p\,(T_{fz} - T_{pre}) + \lambda\,\phi_v\right], \qquad T_{fz} = \mathrm{COT} - 2\,^\circ\mathrm{C} \tag{4.4}$$

$$T_{pre} = T_{cold} + \frac{\eta_{PA}\,(Q_{TPA} + Q_{KPA} + Q_{DPA})}{\dot{m}\,c_p} \tag{4.5}$$

where $\dot{m} = \rho F$ is the crude mass flow, T_{fz} is the flash zone temperature after a fixed transfer line drop, $\phi_v = T_{bp}^{-1}(T_{fz} + \Delta T_{fl})$ is the fraction vaporized, read on the assay $\Delta T_{fl} = 25\,^\circ\mathrm{C}$ above the flash zone temperature to stand for the stripping steam, and the preheat follows Equation (4.5) with the lag τ_{pre}. Equation (4.4) is the heater balance of Equation (1.1) with the vaporization term kept. The overflash is the liquid that overflows the wash zone above the flash zone, and it is what wets the packing there. It is the difference between what the flash zone vaporizes and what the draws take, $\Phi = 100\,(\phi_v - v_4)$ in volume percent, so it falls when the coil outlet falls or when the draws rise. With every draw held, the overflash moves 0.185 vol% per degree of coil outlet temperature, and eight degrees off the coil outlet takes it from 3.28 vol% to 1.80 vol%, below its 2 vol% minimum, while the kerosene freeze point does not move at all.

4.3.4 Assumptions

The tower is represented by cut points on an assay rather than by trays. There is no stage-by-stage calculation, no vapor-liquid equilibrium and no tray hydraulics, so flooding, weeping and the capacity limit that ends a real push are absent.

Table 4.2 Model Parameters

Symbol	Meaning	Value	Class
c_p	crude specific heat	2.60 kJ/kg/K	adjustable
η_{PA}	pumparound heat recovered	0.75	adjustable
T_{cold}	crude before pumparounds	173 °C	adjustable
ΔT_{fl}	flash curve below TBP	25 °C	adjustable
λ	latent heat of traffic	250 kJ/kg	fixed
β	tail spread at unit R	12.0 °C	adjustable
n	tail sharpening exponent	1.0	hidden
$v_{1,ref}$, L_{ref}	reference naphtha cut	0.200, 150 t/h	hidden
p	reflux power on the top cut	0.45	adjustable
τ_1–τ_4	cut point lags	4, 8, 12, 15 min	adjustable
τ_v	draw and reflux flow lag	1.2 min	adjustable
τ_{COT}	furnace thermal lag	5 min	adjustable
τ_{PA}	pumparound duty lag	6 min	adjustable
τ_{pre}	preheat train lag	14 min	adjustable
θ	product analyzer cycle	12 min	adjustable
τ_{an}	analyzer averaging lag	2.0 min	adjustable
Φ_{min}	wash zone dry below	2.0 vol%	visible
$Q_{F,max}$	furnace firing limit	85 MW	visible
COT_{max}	coil outlet limit	395 °C	visible

The tail model is empirical. A real separation sharpness depends on tray count, efficiency and stripping steam as well as on internal reflux, and the relationship between a tail and a D86 specification is a correlation rather than a calculation.

Stripping steam is folded into a fixed offset between the true boiling point curve and the flash curve, so steam rates cannot be studied.

The pumparound sections have no vapor traffic model above the return, which is the largest single limitation and is discussed in Section 4.7.

The crude is a single assay with a heaviness parameter that shifts the curve and tilts it linearly. A real crude switch changes the shape of the curve and not just its position and slope, and it changes sulfur, acid number and metals, none of which appear here.

Product properties are correlations fitted to the sliced assay rather than laboratory methods. The D86 relationship to the true curve is itself an approximation [6].

4.4 Controller Design

4.4.1 Velocity Form PI

Every loop uses the velocity-form PID of Section 1.6.2, Equations (1.16) to (1.19), executed every $\Delta t_c = 0.2$ min. No loop in this application uses derivative action ($\tau_D = 0$), so each is a PI controller, and the tracking time of the back-calculation is $T_t = \tau_I$. The controller direction is the $a = \pm 1$ of Equation (1.16). In manual the integral follows the operator's valve position, which makes the transfer to automatic bumpless, and in automatic the operator's station follows the controller output. The quality loops act on analyzer readings that are sampled every $\theta = 12$ min and averaged with a 2 min lag, while the overflash loop acts on a continuous measurement with a 0.8 min lag. Integration is fixed-step fourth-order Runge-Kutta at 0.05 min (Section 1.6.1).

4.4.2 The Four Loops and Their Directions

The plant gains at the design point, taken as local slopes of the steady-state model, are –0.448 °C/t/h from top reflux to naphtha end point (–0.425 °C/t/h averaged over a 10 t/h increase, the figure the application quotes), 0.692 °C/t/h from kerosene draw to kerosene freeze point, 0.411 °C/t/h from diesel draw to diesel cloud point, and –0.176 vol%/t/h from gas oil draw to overflash. The section

time constants are 4, 8, 12 and 15 min. The tunings the automatic scenarios load set each controller gain near the reciprocal of its plant gain, the internal-model-control gain for a closed-loop time constant equal to the open-loop one [15], giving K_c of 2.5, 1.4, 2.4 and 5.5. The integral times of 6, 12, 18 and 22 min are one and a half times the section time constants rather than equal to them.

Two of the four loops are direct acting and two are reverse acting, and getting the signs right requires reading the assay rather than the flow sheet. More reflux makes a lighter naphtha, so a rising end point calls for more reflux and the end point loop is direct acting ($a = -1$). The flash point controller on the same valve is direct acting for the same reason. More kerosene draw takes heavier material into the kerosene, so its freeze point rises with the draw and that loop is reverse acting ($a = +1$), as is the diesel loop. More gas oil draw takes liquid that would otherwise have been overflash, so a falling overflash calls for less draw and the overflash loop is direct acting.

> HISTORY **1973: Which Variables, Which Valves, Which Links**
>
> In 1966 the relative gain array gave engineers a measure of interaction for deciding which valve to pair with which measurement [16]. The wider question was put in a 1973 critique of chemical process control theory: which variables should be measured, which inputs manipulated, and what links made between the two sets. Academic work, the critique held, had largely neglected these structural decisions [16]. A series of papers in 1980 answered in part with a rule, to always control the active constraints, and the pairing of the gas oil draw with the wash zone in this chapter is in that spirit. The review that gives these dates still makes the practical case for single loops. They can be tuned one at a time without a model of the interactions, and the gain from a multivariable controller has to be weighed against the cost of obtaining and maintaining its model [16].

4.4.3 The Top Loop Has Two Objectives and One Valve

The top loop selector chooses whether the reflux holds the naphtha end point or the kerosene flash point. It cannot hold both, and the application makes the choice explicit rather than offering two loops that would fight. Each objective has its own controller, and the one not selected tracks the reflux so that a switch between them is bumpless. Measured in scenario 2, the end point loop takes the end point from 184.0 °C to 176.0 °C by raising the reflux from 150 t/h to 170 t/h, and the kerosene flash point follows from 53.9 °C to 48.8 °C, a ratio of 0.63.

Pause. Try Exercise 4.1 before reading on.

4.4.4 The Overflash Override

The gas oil draw is paired with the overflash rather than with a gas oil specification. A slightly heavy gas oil is giveaway. A dry wash zone is coke on the packing and a shortened run. A one-directional override acts on the gas oil draw downstream of its loop. Whenever the measured overflash Φ_m is below its minimum, the requested draw is multiplied by $1 - \min(0.6, (\Phi_{min} - \Phi_m)/1.5)$, so the cut grows by 6.7 percent of the draw per 0.1 vol% of shortfall up to 60 percent, without waiting for the loop to integrate. The override is not tied to the loop mode and also acts with the gas oil draw in manual. The PI controller neither sees the cut nor back-calculates against it, so its integral keeps acting on the overflash error alone, and because the cut vanishes continuously as the measured overflash returns to the minimum, the hand-back to the loop is bumpless. The override is proportional, so it limits the time below the minimum without removing it. Pairing a manipulated variable with a constraint rather than with a product is the right choice whenever the constraint is more expensive to violate than the specification is. Chapter 5 protects its wash bed with an override of the same kind, and Volume 4, Chapter 2 treats a protective override on a machine.

4.4.5 Model Predictive Control

A second structure is offered beside the four loops, never in their place. Section 1.6.3 describes the predictive control option shared by the series, and this section gives what is specific to the tower. With Control set to MPC and the tower in automatic, one plan moves the top reflux and the three

Table 4.3 Identified Coefficients of the Controller's Model

Handle	Quality	a	b	Gain
L	EP (°C)	−0.0881	−0.3603	−0.4484
L	FP (°C)	+0.0882	−0.3602	−0.2720
L	FR (°C)	−0.0059	−0.2252	−0.2311
L	CP (°C)	−0.0029	−0.1351	−0.1380
L	Φ (vol%)	+0.0600	—	+0.0600
D_K	FR (°C)	+0.0328	+0.6587	+0.6915
D_K	CP (°C)	+0.0029	+0.3952	+0.3981
D_K	Φ (vol%)	−0.1756	—	−0.1756
D_D	CP (°C)	+0.0153	+0.3952	+0.4105
D_D	Φ (vol%)	−0.1756	—	−0.1756
D_G	Φ (vol%)	−0.1756	—	−0.1756

Coefficients are per t/h of the handle, a on the flow lag and b on the section lag of Equation (4.6). The gain is $a+b$.

draws together. Every $T_s = 1$ min, five controller scans, a quadratic program chooses all four handles over $N = 120$ decisions, two hours ahead, and only its first move is applied, for one period [8, 17]. The loop switches then say only which loops take over if a plan fails, and the wash zone override still trims the gas oil draw downstream.

The controller's model maps $u = (L, D_K, D_D, D_G)$ to the five qualities that carry both a specification and an instrument the loops read, the end point, the flash, freeze and cloud points and the overflash. A first design with eight outputs, adding a gas oil end point and two tray temperatures, was not built, because the application has no gas oil end point analyzer or specification and its one tower temperature is read by no loop. Step tests at the design point, in manual with the analyzer noise off (5 t/h of reflux and 3 t/h of draw, up and down, 120 min of the true qualities), gave second-order-plus-dead-time fits that all placed a lag near 1 min beneath one of 4, 8 or 12 min. That is the structure of the plant equations, so the model is written on it, Equation (4.6), with coefficients fitted by linear least squares.

$$\begin{aligned} \tau_v \dot{x}_{v,j} &= u_j - u_{j,0} - x_{v,j}, \qquad \tau_s \dot{x}_{s,j} = x_{v,j} - x_{s,j} \\ \hat{y}_q &= y_{q,0} + \sum_j \left(a_{qj} x_{v,j} + b_{qj} x_{s(q),j} \right) + \beta_q \end{aligned} \tag{4.6}$$

Here $u_{j,0}$ is handle j at $t = 0$, $x_{v,j}$ its flow lag state with $\tau_v = 1.2$ min, and $x_{s(q),j}$ the lag of the section whose cut point quality q reads, τ_1 for the end point and flash point, τ_2 for the freeze point and τ_3 for the cloud point, while the overflash reads the flow lags alone. $y_{q,0}$ is the reading at the start and β_q the output bias. The lags are the declared values of Table 4.2, not the run's, so an edited plant parameter is a model mismatch. The ten-state model's coefficients are in Table 4.3, with residuals of at most 0.003 percent of each response, and a repeat with 1 t/h steps agrees within 0.08 percent of each gain from 1 to 120 min. The flash point's apparent 1.5 min dead time is an inverse response that the opposite signs of its a and b reproduce.

The plan predicts the true qualities, and the analyzers enter through the estimator. Each prediction passes through a copy of its instrument, sampled in the scan in which the real analyzer publishes and averaged with the declared 2 min lag, or for the overflash the declared 0.8 min lag, and every scan $\beta_q \leftarrow \beta_q + \lambda_q (z_q - y_{q,0} - \tilde{y}_q - \beta_q)$. Here z_q is the reading the loops see, $\tilde{y}_q$ the copy's output, $\lambda_q = \Delta t_c / \tau_b$ with $\tau_b = 6$ min for the analyzers and $\lambda_q = 1$ for the overflash, which makes the prediction offset free [18]. Charge, crude, coil outlet and pumparounds reach the plan only through the biases, and there is no preview. The plan minimizes Equation (4.7) over twelve move blocks of

1, 1, 1, 1, 2, 2, 3, 4, 6, 9, 15, 75 decisions.

$$J = \sum_{k=1}^{N} \left[q\,(\hat{y}_{t,k} - r_t)^2 - w\Big(D_{K,k} + \tfrac{p_D}{p_K} D_{D,k} + \tfrac{p_G}{p_K} D_{G,k}\Big) \right] + S \sum_{m=1}^{12} \|\Delta u_m\|^2 + \sum_i \rho_i \left(\sigma_i^2 + \sigma_i\right) \tag{4.7}$$

$\hat{y}_{t,k}$ is the top quality the selector holds, k decisions ahead, with target r_t, $q = 1$ per °C^2, $S = 0.1$ per (t/h)^2 with Δu_1 measured from the commands applied now, and $w = 0.05$ per t/h scales indicative values of $p_K = 260$, $p_D = 230$ and $p_G = 120$ dollars per tonne above residue. The constraints are Equation (4.8).

$$\begin{aligned} & u_{min} \le u_m \le u_{max}, && \sigma_i \ge 0 \\ & \widehat{FR}_k \le e_k^{FR} + \sigma_{FR}, && \widehat{CP}_k \le e_k^{CP} + \sigma_{CP} \\ & \hat{\Phi}_k \ge e_k^{\Phi} - \sigma_{\Phi}, && \hat{\Phi}_k \ge e_k^{W} - \sigma_W \end{aligned} \tag{4.8}$$

The handle ranges of Table 4.1 are hard. The soft one-sided bands sit on the loops' own targets (−47 and −1.8 °C, 3.3 vol%) rather than on the specifications, so that both structures work to the same numbers, with a second overflash band at $\Phi_{min} + 0.6$ vol% and a band on the unheld top quality at its specification (38 or 185 °C). Each slack is priced at $\rho_i = 10^4$, or 10^5 for the wash zone band. Each edge is a funnel, $e_k = \bar{y} + \max(0,\, h_k - \bar{y})\, f_k$ for an upper limit $\bar{y}$ and mirrored for a lower one, where h_k is the prediction with the present commands held, $f_k = 1$ for the first minute and $\exp(-(t_k - 1)/\tau_q)$ after it, and τ_q is the declared lag of the quality's section, or 1.2 min for the wash zone band. The top quality is a setpoint because only one of the end point and flash point can be held, and lined out on scenario 3 with the noise off the value holds the true end point 0.0190 °C below target (the plan's own horizon-end prediction is 0.0202 °C). A dual active-set solver [19] handles each plan, and S and τ_b were chosen once on scenario 10 for the least integrated error among settings that rejected no plan and moved the handles no more than the loops.

Under the plan each loop tracks the applied command with no preload, its output equal to the command at every scan of scenarios 6, 8 and 10 with the noise off. Any plan with a solver status other than optimal, or with a first move that is infinite, undefined or beyond a handle range, is discarded with a historian mark, and the loops take over. Switching to Loops at five instants moved each handle by exactly one loop execution, at most 0.0165 t/h of reflux and 0.1377 t/h of gas oil, and a switch to manual leaves every station where the plan put it. The plant runs on the run's parameters and the instruments sample, average and add noise, so neither coincides with this model. In Chrome a plan took 4046 and 4126 µs on average in two measurements on scenario 10, 0.41 percent of a core at real time and 48.6 to 49.5 percent at the 120 times top speed, where the application kept 58.5 to 59.7 fps against 59.5 to 59.9 fps under the loops.

4.5 Results

Eleven scenarios run from slicing the assay by hand to a graded challenge. Each scenario is credited separately (Section 1.6.4). A case is complete when the learner performs the actions its rule lists and the simulation reaches the listed state. The challenge is judged on the case as given, and changing a model parameter or a disturbance setting voids the attempt. Credit accumulates toward certificates of progress and of completion for the module.

4.5.1 Degrees of Freedom

Table 4.4 reports the top reflux swept across its working range with the top loop in manual and every draw held at its design value, read once the run has settled.

The measured ratio of the two movements is about 0.63 over 150 t/h to 170 t/h. It depends on the assay and on how the tail narrows as the traffic rises. The tower has four cut points and eight end

Table 4.4 One Valve Moving Two Specifications

Quantity	$L = 135\,\mathrm{t/h}$	$L = 170\,\mathrm{t/h}$
Naphtha end point (°C)	191.4	176.0
Kerosene flash point (°C)	58.2	48.8

specifications, so four of the eight are choices and four are consequences, and the only question is which four to choose.

4.5.2 Triangular Interaction

With every draw held, twenty more tonnes an hour of diesel leaves the kerosene cut point exactly where it was, and twenty more tonnes an hour of kerosene moves the diesel cut point by 17.6 °C. The closed loop shows the same direction in scenario 3. Lowering the freeze point target from −47 °C to −50 °C with all four loops in automatic takes the kerosene draw from 74.0 t/h to 69.8 t/h and the diesel draw from 102.1 t/h to 106.2 t/h, while the reflux and the gas oil draw (50.9 t/h) stay where they were. Lowering the cloud point target by 3 °C instead moves the diesel draw to 94.7 t/h and the gas oil draw to 58.1 t/h, and the kerosene draw stays at 74.1 t/h. The interaction runs one way, down the tower, which is what makes a tower with four interacting loops tunable one loop at a time. Volume 6, Chapter 2 finds the same triangular structure in the gain matrix of a reverse osmosis train.

4.5.3 Commissioning Order

Scenario 7 brings the four loops into automatic twenty minutes apart from the bottom of the tower upward, the gas oil loop at 10 min, the diesel loop at 30, the kerosene loop at 50 and the top loop at 70, and a heavier crude (heaviness 0.35) arrives at 120 min. Summed over the reflux and the three draws, the valve travel over the first 120 min is 3.8 t/h. Bringing the top loop in at 10 min and the kerosene loop at 30, alongside the scripted gas oil and diesel switches, raises it to 6.4 t/h, and most of the difference (2.1 t/h of 2.6 t/h) is in the reflux. A strictly top-down order (top at 10 min down to gas oil at 70, with the scripted switches removed) takes 5.7 t/h. After all four are in service the schemes behave identically, with peak overflash deviations of 4.99 vol% and 5.01 vol% and peak freeze point deviations of 12.02 °C and 12.05 °C on the crude change, within 0.3 percent of each other, so the entire difference is in the commissioning transient.

The reason is the triangular structure. Closing the top loop moves the naphtha boundary and therefore every boundary below it. Downstream loops closed first are already in service to absorb that, while downstream loops closed last let the drift accumulate in manual until an operator picks it up. The familiar heuristic, commission the inner or fastest loop first, is about cascades, where the interaction points from the inner loop outward. The statement that covers both is to bring up the loop that receives the disturbance before the loop that generates it.

4.5.4 Pumparound Heat Recovery

Table 4.5 reports scenario 5, in which the top pumparound steps from 18 MW to 26 MW at 30 min and to 12 MW at 180 min with all four loops in automatic, read at 30, 180 and 360 min. Each megawatt of duty raises the preheat by 1.82 °C and saves 0.75 MW of firing, which is η_{PA}. The kerosene loop takes 1.7 t/h more draw at 26 than at 12 MW, because the sharper split puts more kerosene into the cut at the same freeze point.

Ten and a half megawatts of fuel and a sharper split follow from one setpoint, which raises the obvious question of why not take all of it. In this model nothing stops it, and that is a limitation stated in Section 4.7.

Table 4.5 Top Pumparound Duty Against Preheat, Firing and Sharpness

Quantity	12 MW	18 MW	26 MW
Crude preheat (°C)	238.7	249.6	264.2
Furnace firing (MW)	81.0	76.5	70.5
Kerosene and diesel tail (°C)	3.51	2.65	2.00
Kerosene draw (t/h)	73.1	74.1	74.8

TRY IT IN THE APP **5 · Pumparounds: heat against sharpness**

Load the scenario and press Run. Nothing has to be set. At 30 min the top pumparound steps from 18 to 26 MW, and at 180 min it drops to 12 MW. Watch the furnace firing beside the tower. It settles at 70.5 MW with the duty at 26 MW and at 81.0 MW with it at 12 MW, because each megawatt recovered saves 0.75 MW of firing.

4.5.5 The Constraint Binds First

With every draw held, eight degrees off the coil outlet takes the overflash from 3.28 vol% to 1.80 vol%, below its minimum, while the kerosene freeze point does not move at all. The sensitivity is 0.185 vol% per degree and nothing else. With the four loops in automatic the gas oil loop absorbs an eight-degree loss without the override, and the overflash bottoms at 2.70 vol%, so scenario 6 takes 22 °C off the coil outlet at 30 min. The overflash crosses its minimum at 34.2 min, the override engages at 35.1 min and removes at most 2.95 t/h beyond the loop's request, and the overflash bottoms at 1.89 vol% at 42.2 min before the override releases at 58.4 min, as shown in Figure 4.2. The wash zone is dry for 23 min in all. The gas oil draw settles at 27.5 t/h before the coil outlet is restored at 200 min and is back at 50.4 t/h by 360 min, and the kerosene freeze point stays within 0.09 °C of its target throughout. The constraint binds long before any product notices, which is the argument for pairing the gas oil draw with it.

Pause. Try Exercise 4.3 before reading on.

4.5.6 The Graded Challenge

The challenge runs 700 min with analyzer noise of 0.8 °C and three disturbances, a heavier crude (heaviness 0.35) at 60 min, a charge increase from 662 m^3/h to 740 m^3/h at 260 min, and a fouled top pumparound (18 MW to 12 MW) at 460 min. It starts with the top loop alone in automatic and cautious tuning. It is met when, with the targets left as given, the overflash stays within 0.65 vol% of target and the kerosene freeze point within 1.5 °C over the last third of the run, and the wash zone is dry for no more than 15 min in all. With the shipped configuration (top loop alone in automatic, cautious tuning, coil outlet 377 °C) the worst deviations over the last third are 4.01 vol% and 1.92 °C and the wash zone is dry for 174.2 min, so the run fails on all three counts. With the suggested configuration (all four loops in automatic with K_c and τ_I of 1.5 and 10 min on the top loop, 0.8 and 24 min on the kerosene loop, 2.4 and 18 min on the diesel loop and 14 and 22 min on the gas oil loop, coil outlet 381 °C, kerosene and diesel pumparounds at 24 and 18 MW) they are 0.28 vol% and 0.65 °C with 1.0 min dry, inside the limits by 0.37 vol%, 0.85 °C and 14.0 min. Both runs use the analyzer noise seed the application reuses on every run. The suggested configuration also passes with each of seeds 1 to 400, at worst 0.46 vol%, 1.24 °C and 2.0 min dry. The suggested tuning differs from that of the other automatic scenarios. The top and kerosene loops act on the noisy analyzers and both move the kerosene cut, so they run softer (K_c 1.5 and 0.8 against 2.5 and 1.4), while the gas oil loop acts on the noise-free overflash and runs at $K_c = 14$ rather than 5.5. The wash zone is then dry only for the minute after the instantaneous crude switch. The heat moves are not scored. They keep the firing at 81.7 MW after the pumparound fouls, where the same tuning without them fires 93.0 MW.

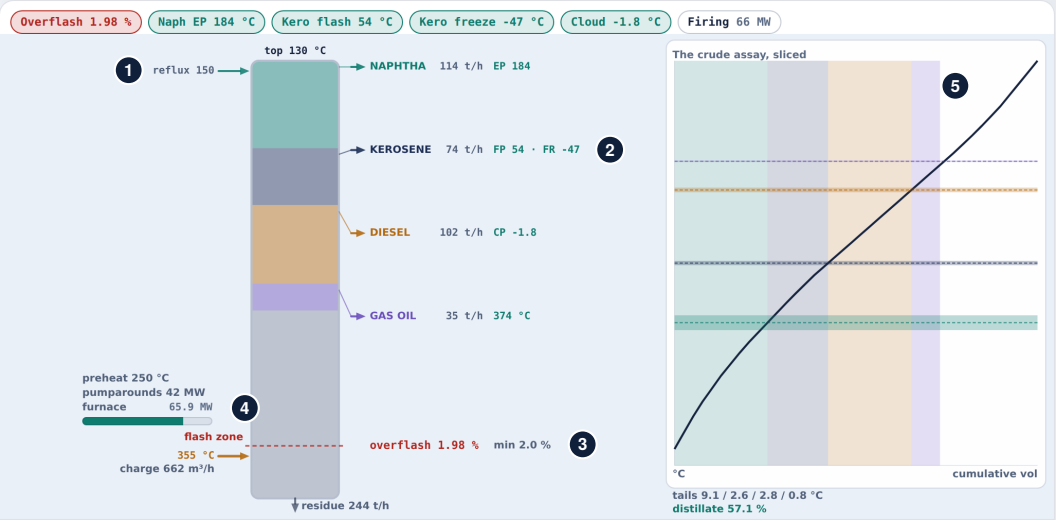


Figure 4.2 The crude tower application in scenario 6 at 55 min, 25 min after the coil outlet temperature stepped from 377 °C to 355 °C with all four loops in automatic, cropped to the plant view with numbered callouts. (1) The top reflux, the handle of the top loop. (2) The kerosene draw with its flash and freeze points. (3) The overflash at 1.98 vol%, just below its 2 vol% minimum, where the override holds the gas oil draw at 35.0 t/h against a loop request of 35.7 t/h. (4) The furnace firing of 65.9 MW. (5) The inset, which slices the assay at the four cut points. The naphtha end point, kerosene freeze point and diesel cloud point stay on their targets (184.0, −47.0 and −1.8 °C), and in the full window, outside this crop, the trends cover the last 30 min.

4.5.7 Comparison With Model Predictive Control

Table 4.6 sets the four loops against the plan of Section 4.4.5 on scenarios 10, 5, 6, 9 and 8 and on the challenge. Each row runs from $t = 0$ with the case's seed, instruments and noise, the loops at shipped tuning (suggested on the challenge), the plan at scenario 10's settings, with no preview and the same targets. IAE and travel are the scorecard's tiles on the true qualities. The off-specification minutes are not tiles but the CSV accumulators `metric_off_naph_min`, `metric_off_kero_min` (flash below 38 °C or freeze above −40 °C) and `metric_off_diesel_min`, integrated every 0.05 min step. Figure 4.3 draws three of the table's columns for the six cases that both structures run.

In scenario 10 only the reflux is asked to move, but a lighter naphtha moves every cut beneath it. Under the loops the freeze point falls to −50.69 °C at 39.4 min, because the reflux has moved the naphtha and kerosene cut and the kerosene draw has not, and swings back to −45.95 °C at 62.4 min. The plan's freeze point never falls below −48.10 °C, as shown in Figure 4.4, because its model carries the triangle and raises the kerosene draw with the reflux, to 74.98 t/h by 31 min while the loops' draw is still at 74.06. On the step back the loops carry the end point to 187.59 °C at 158.2 min, off specification for 8.35 min, and after 150 min the plan's end point never exceeds 183.98 °C. The plan wins all four integrated errors, the freeze point by 18.1 against 123.9 °C · min, with 98.8 against 270.2 t/h of travel and the same 223.0 k$ of draw value.

On scenario 9 the plan cuts the end point error from 304.7 to 35.1 °C · min but loses the cloud point, 318.7 against 179.6, because the higher charge lightens every cut and the plan's cloud point, never above its −1.8 °C target, gives away inside a one-sided band that only the value's pull closes, more slowly than the diesel loop's integral. On scenario 6 the wash zone is dry for 8.3 against 23.45 min, and on scenario 5 the plan loses only the freeze point, 30.8 against 26.8, with 10.3 t/h of travel against 33.8. The draw value stays within 0.3 percent of the loops' in every shared case.

Table 4.6 Four Loops Against Model Predictive Control

Case (window)	Structure	IAE, true quality from $t = 0$				Travel	Off specification (min)			Wash zone	Draw value	Plans,
		EP	FR	CP	Φ	(t/h)	naphtha	kerosene	diesel	dry (min)	(k$)	rejected
10, end point steps (270 min)	loops	148.2	123.9	77.4	39.4	270.2	8.35	0	0	0	223.0	—
	predictive	66.0	18.1	10.0	4.1	98.8	0	0	0	0	223.0	270, 0
5, pumparound traffic (360 min)	loops	12.6	26.8	20.8	10.0	33.8	0	0	0	0	292.9	—
	predictive	7.4	30.8	15.4	7.2	10.3	0	0	0	0	293.0	360, 0
6, coil outlet (360 min)	loops	12.6	6.5	3.6	182.7	63.0	0	0	0	23.45	285.0	—
	predictive	7.3	6.6	7.6	86.7	54.8	0	0	0	8.3	285.4	360, 0
9, charge (600 min)	loops	304.7	223.0	179.6	99.0	519.4	45.8	0	0	0	572.7	—
	predictive	35.1	200.7	318.7	52.4	89.4	13.4	0	0	0	571.4	600, 0
8, heavier crude (480 min)	loops	371.7	294.5	193.1	98.2	542.5	16.5	8.9	21.25	19.2	385.9	—
	predictive	769.3	443.3	250.0	81.6	212.2	79.8	26.0	47.45	3.5	385.0	480, 0
	bias every scan[a]	511.1	287.0	167.5	75.7	308.2	57.45	20.2	34.75	3.5	—	480, 0
	reflux gains ×0.72[a]	567.5	332.5	182.3	71.7	201.9	50.7	20.55	34.2	3.5	—	480, 0
	both[a]	361.2	206.8	119.6	63.0	265.9	34.0	11.55	26.85	3.5	—	480, 0
11, challenge, suggested (700 min)	loops	766.1	569.8	399.4	95.9	924.9	91.6	16.05	17.65	1.0	611.0	—
	predictive	820.9	638.7	543.9	125.3	639.4	71.55	21.35	37.2	2.95	610.7	700, 0
11, challenge, shipped[b] (700 min)	predictive	821.2	638.5	539.5	119.5	636.8	71.65	21.35	38.25	3.0	603.2	700, 0

IAE is in °C · min, and in vol% · min for the overflash Φ. Travel sums the four handles' travel tiles. Off-specification minutes are the CSV accumulators named in the text, at their 0.05 min resolution. The draw value prices the kerosene, diesel and gas oil made at 260, 230 and 120 dollars per tonne above residue. [a] Diagnosis copies of the predictive controller, not in the application: the analyzer biases updated every scan instead of through the 6 min filter, the model's reflux coefficients on the four analyzer qualities scaled by 0.72, or both. [b] Control set to MPC and every other setting as shipped. The scored target is met (worst deviations 0.55 vol% and 0.94 °C over the last third, 3.0 min dry), while the furnace fires 93.0 MW at 700 min, above its 85 MW limit for the last 439.4 min, exactly as the loops fire with the same settings. Solve times are given in Section 4.4.5.

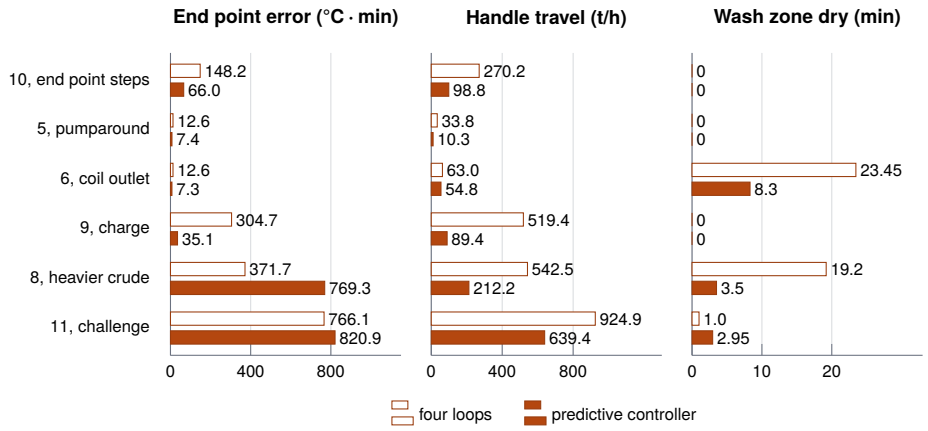


Figure 4.3 Three columns of Table 4.6 for the six cases that both structures run, the challenge with the suggested settings. White bars are the four loops and solid bars the predictive controller. Notice that the plan travels less in every case, wins the end point error everywhere but on the heavier crude and the challenge, and shortens the dry time of the wash zone in scenarios 6 and 8.

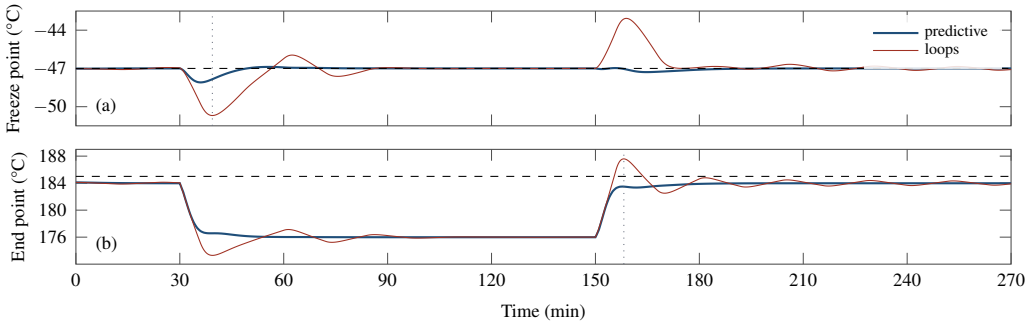


Figure 4.4 Scenario 10 under the predictive controller and under the four loops, with the end point target stepped from 184 to 176 °C at 30 min and back at 150 min. (a) True kerosene freeze point against its −47 °C target. (b) True naphtha end point against the 185 °C specification. The dotted verticals mark 39.4 min, the loops' lowest freeze point, and 158.2 min, the loops' highest end point.

The plan loses scenario 8 on every quality, an end point error of 769.3 against 371.7 °C · min and 79.8 against 16.5 min of naphtha off specification, while keeping the wash zone dry for 3.5 against 19.2 min. The first cause is the model's gain, though not through the crude itself. At the design reflux of 150 t/h the heavier crude makes the end point 1.14 times as sensitive to the reflux as the model says. The climb of the reflux to about 260 t/h that the crude forces flattens the power law of Equation (4.1), to 0.47 of the model's gain there and 0.72 over the climb, so each analyzer result is answered with too little reflux. The second cause is the 6 min bias filter, which adds its lag to the analyzer's. In diagnosis copies that are not in the application the end point error falls to 511.1 with the bias updated every scan, to 567.5 with the reflux gains scaled by 0.72, and to 361.2 with both, against the loops' 371.7, yet the naphtha is still off specification for 34.0 against 16.5 min.

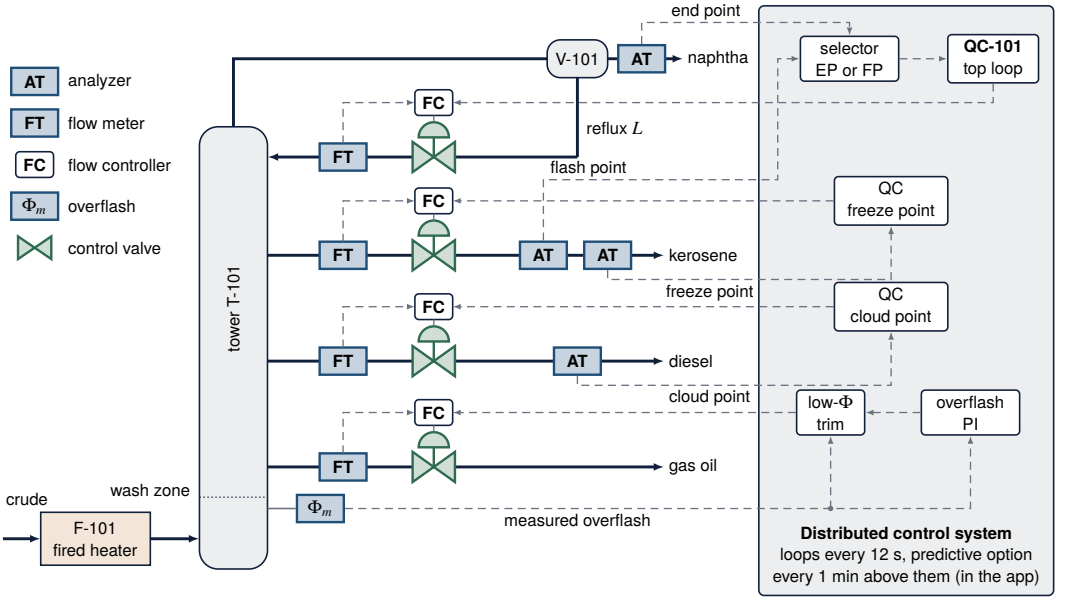


Figure 4.5 The four loops with their hardware. Lettered boxes are sensing devices, valve symbols are final elements and dashed arrows are signals. Notice that every quality controller writes the set point of a flow controller (FC, drawn at its valve) and no valve directly, that the top loop reads two analyzers through a selector, and that the wash zone trim sits between the overflash controller and its flow loop. Not to scale, and not a piping or safety diagram.

> ✦ **DID YOU KNOW? The Controller That Won the Step Lost the Crude**
>
> From the simulator (Table 4.6). On the end point steps of scenario 10 the predictive controller integrates 66.0 °C · min of end point error against the loops' 148.2. On the heavier crude of scenario 8 the same controller integrates 769.3 against 371.7, yet it keeps the wash zone dry for 3.5 against 19.2 min.

For information, the challenge with Control on MPC and everything else as shipped meets its scored target, a worst overflash deviation of 0.55 vol% at 564.0 min, a worst freeze point deviation of 0.94 °C at 533.0 min and 3.0 min dry. In the same run the furnace fires 93.0 MW, over its 85 MW limit for the last 439.4 min, and the naphtha is off specification for 71.65 min, neither of which the target scores. The firing is identical under the loops, because the duty of Equation (4.4) depends only on the coil outlet, the preheat and the crude, never on a handle of either structure. With the suggested settings both fire 81.7 MW and both meet the target, the plan with 0.54, 0.94 and 2.95 against the loops' 0.28, 0.65 and 1.0. The challenge is unchanged. Halving the integration step leaves the winner of every tile unchanged in all six cases, and playback at 1, 20 and 120 times real time is byte identical to the headless run.

4.6 Technology in Practice

Figure 4.5 draws the four loops of Section 4.4 with the hardware a refinery installs in place of the model's ideal instruments.

What Measures It

Four of the five controlled variables are product qualities, read by analyzers that take a sample and report once a cycle. The cycle θ and the lag τ_{an} of Table 4.2 stand for the sample system and the

analysis time of such an instrument. The fifth, the overflash, is read in the model without a cycle, through a 0.8 min lag. A variable that can be determined only at a low sampling rate or by off-line analysis is the original and still dominant use of the soft sensor, a model that predicts it from faster measurements [20].

> **TECHNOLOGY Sensor: An Analyzer That Reports Once a Cycle**
>
> Many industrial analyzers sit away from the process, and tubes, heaters, filters and pumps bring them a steady stream of sample. Problems with process analyzers quite often stem from that sample system [21]. If the line plugs, the analyzer measures stale liquid and has no way of knowing it, so its reading no longer follows the process [21]. The process gas chromatograph shows why such readings come in steps. It injects a small fixed volume of sample, vaporized first if it is a liquid, and separates the species by the time each needs to pass a column. Cycle times of minutes are not uncommon, and some dead time is inevitable [21]. The model's analyzers publish every 12 min through a 2 min averaging lag (Table 4.2).

What Moves It

The top reflux and the three draws are flows, and the model gives each the flow lag τ_v. The coil outlet temperature and the pumparound duties are operator settings in the application, with lags of 5 and 6 min. A guide for refinery energy managers puts the crude that enters the tower at 340 to 390 °C [2], a range that holds the 377 °C of Table 4.1.

> **TECHNOLOGY Final Element: A Flow Loop Under Every Handle**
>
> The handles of this chapter are flows in tonnes per hour, so each is a loop of its own: a meter, a flow controller and a control valve. The quality loop above sends the flow set point, and the flow controller counters a change in supply pressure before the outer loop has to react [21]. Such a cascade requires an inner variable faster than the outer one, and liquid flow answers a valve at once [21]. The most common pressure-based flow element is the orifice plate, whose differential pressure rises with the square of the flow. Such meters are commonly understood not to hold good accuracy much below a third of their full range [21]. The model gives each flow loop one lag, 1.2 min, and an exact meter.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level sends set points to distributed controllers that act through the actuators from the feedback of the sensors [22]. A survey of industrial practice puts basic control at every second and constraint control at every minute, and builds the conventional constraint layer from PID algorithms, lead-lag blocks and high and low select logic [8]. The four loops of this chapter execute every 0.2 min in the application, and the selector on the top loop and the trim on the gas oil draw belong to that conventional layer. The predictive option plans every 1 min, the period the survey names.

> **IN PRACTICE On a Working Unit, a Temperature Loop Stands Under the Quality**
>
> An account of one industrial crude unit shows what stands between a quality and a valve [11]. A temperature controller at the tower top sits under the predictive controller, which holds the flash point of the diesel product, and keeps the top reflux inside its bounds, by moving that controller's set point. The 95 % point of the diesel, an ASTM D86 specification, is held mainly through the heavy diesel reflux and the furnace outlet temperature. Ten outputs are kept by eight inputs once a minute, most of them inside zones. The main disturbance is the composition of the crude, and the optimizer above acts once a crude analyzer has detected a change and the unit has reached steady state [11].

What Is Modernizing It

The first of three developments is predictive control with an optimizer above it. A survey taken in mid-1999 counted 4542 linear predictive applications, 1985 of them in refining, the largest single

block [8]. On the unit of the In practice box a rigorous steady-state model sets targets for the crude feed rate, the kerosene draw, the furnace outlet temperature and five more variables [11]. The second is the soft sensor. Its weak point is upkeep: after launch its performance usually deteriorates as the process changes, and most soft sensors have no automated maintenance, which leaves the work to people and makes it a significant cost [20]. The third is the faster analyzer. Programming the column temperature of a process chromatograph during the cycle gives the separation of a long column with the dead time of a much shorter one [21].

4.7 Limitations

The largest limitation is the pumparound. The application models the heat recovered and the liquid condensed below the return, and it does not model the vapor traffic lost above it. On a real tower, taking more pumparound duty makes the section above it softer, so the trend in Table 4.5 would reach a point of diminishing return that this model does not show. A real tower also has a hydraulic capacity, and a pumparound circulating rate is a large flow through trays that can flood.

The tower is cut points on an assay rather than trays. Nothing hydraulic can be studied, tray efficiency is absent, and stripping steam is a fixed offset rather than a manipulated variable. The furnace firing is the duty the coil outlet requires rather than a limited actuator, so crossing the 85 MW limit raises an alarm without capping the firing or pulling the coil outlet down.

Product properties are correlations on the sliced assay. A student would not learn here that a kerosene can fail on sulfur or on smoke point, or that a diesel cetane number depends on composition rather than on boiling range alone.

The crude heaviness disturbance shifts and tilts the assay rather than reshaping it, so a real crude switch is more complicated than the one modeled, and none of the corrosion, fouling or catalyst poisoning consequences of a crude change appear.

The regulatory layer implemented here is what a real unit carries underneath a multivariable controller. The application includes one such controller only as an option, whose bands sit on the loops' own targets and whose product values are indicative, so the economic trade between throughput and quality that a real unit makes continuously [11] is still set by the student through those targets rather than by the software.

The predictive controller's model is linear, fitted at one design point and blind to every disturbance, which reaches it only through the analyzer biases, and on the heavier crude of scenario 8 the plant's end point gain at 260 t/h is 0.47 of the model's. The furnace firing is in neither structure, so neither can act on the 85 MW limit.

The analyzer and overflash readings start at the design targets, not at the tower's state, so a setting chosen before Run reaches the readings only over their averaging lag, and the first minutes of such a run show the change late.

4.8 Discussion

Three results generalize past crude towers.

The first is that specifications and controlled variables are different things. On any unit making consecutive fractions from one feed there are more specifications than handles, the surplus ones are consequences, and which are which is an operating decision rather than a control design. The arithmetic is simple and the consequence of skipping it is two loops on one valve, which presents on a panel as a loop that has stopped working.

The second is that the direction of interaction determines the commissioning order, and the familiar rule is stated for the wrong direction. Measured here, bringing the top two loops in first costs 1.7 times as much valve travel as commissioning from the bottom up (6.4 t/h against 3.8 t/h). The general form covers cascades and triangular plants both: bring up the loop that receives the disturbance before the loop that generates it.

The third is that a manipulated variable is often better paired with a constraint than with a product. The overflash moves 0.185 vol% per degree of coil outlet while the kerosene freeze point moves by nothing, so the constraint is the fast and sensitive measurement on that handle as well as the expensive one to violate. Where those two coincide the pairing is obvious, and where they do not the question of what the valve is for has to be answered explicitly.

OPEN CHALLENGE **Optimizing a Tower That Is Seldom at Rest**

The conventional optimizer above a process unit updates a steady-state model from data taken at steady state and then computes new set points [23]. A 2022 review reports the wait for steady state as a fundamental limiting factor of the method, which it finds less used in practice than one would expect [23]. On a crude tower the main disturbance is the crude itself. On one industrial unit, model uncertainty between the layers left some variables oscillating short of the optimum until the middle layer was retuned [11]. Scenario 8 shows the same weakness in small. Progress would be an optimizer that learns from transient data, a family of methods the review surveys [23], proven through a crude switch on a running unit.

4.9 Conclusion

A crude tower is where counting loops stops being enough. The simulated tower of this chapter makes four products from one assay with four handles and eight specifications, and moving the top reflux from 135 t/h to 170 t/h takes the naphtha end point from 191.4 °C to 176.0 °C and the kerosene flash point from 58.2 °C to 48.8 °C together, so the two cannot be held independently. The interaction between the four draws runs one way down the tower, and commissioning along that direction costs 3.8 t/h of valve travel against 6.4 t/h with the top two loops brought in first, and once every loop is in service the peak deviations on a crude change agree within 0.3 percent. The gas oil draw is paired with the overflash rather than with a product because eight degrees off the coil outlet takes the overflash from 3.28 vol% to 1.80 vol% with the draws held while the kerosene freeze point does not move, and because a dry wash zone costs a run rather than a giveaway. Ten and a half megawatts of furnace fuel follow from fourteen megawatts of top pumparound duty, and the same duty narrows the kerosene and diesel tail from 3.51 °C to 2.00 °C. An optional predictive controller that carries the triangle in its model integrates 18.1 against 123.9 °C · min of freeze point error on an end point step with 37 percent of the loops' travel, and loses the heavier crude. A fuller model would add a vapor traffic model above each pumparound return so that the heat recovery sweep reaches its own limit, a hydraulic capacity so that flooding can end a push, and a second crude whose assay changes shape rather than position. As it stands, the simulated tower is a place to count degrees of freedom before pairing loops.

KEY POINTS

- A tower with four handles and eight specifications can hold four of them. The other four are consequences.
- The naphtha end point and the kerosene flash point are one cut point and cannot be held independently.
- The interaction runs one way, down the tower. Bring up the loop that receives the disturbance before the loop that generates it.
- The gas oil draw is paired with the overflash: the constraint binds long before any product notices.
- A predictive controller with the triangle in its model wins the end point step and loses the heavier crude.

Where else this applies. A triangular gain matrix returns in Volume 6, Chapter 2, a wash bed protected by an override in Chapter 5, and the choice of which flow a structure holds in Volume 6, Chapter 7.

References

[1] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[2] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[3] U.S. Energy Information Administration, "Refinery capacity report 2026, table 1: Number and capacity of operable petroleum refineries by PAD district and state as of January 1, 2026," U.S. Department of Energy, Washington, DC, Tech. Rep., 2026.

[4] W. L. Luyben, "Control of a distillation column with side stripper and side rectifier," *Chemical Engineering Research and Design*, vol. 161, pp. 38–44, 2020.

[5] C. Pasquini and A. F. Bueno, "Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity," *Fuel*, vol. 86, no. 12-13, pp. 1927–1934, 2007.

[6] R. N. Santos, E. R. Lima, and M. L. Paredes, "Astm d86 distillation curve: Experimental analysis and premises for literature modeling," *Fuel*, vol. 284, p. 118958, 2021.

[7] J. Kewley, "Fractional distillation as applied in the petroleum industry," in *Distillation Principles and Processes*, S. Young, Ed. London: Macmillan, 1922, pp. 319–360, chapters XXXI and XXXII, pp. 334–351.

[8] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[9] C.-B. Chung and J. B. Riggs, "Dynamic simulation and nonlinear-model-based product quality control of a crude tower," *AIChE Journal*, vol. 41, no. 1, pp. 122–134, 1995.

[10] A. S. Yamashita, A. C. Zanin, and D. Odloak, "Tuning the model predictive control of a crude distillation unit," *ISA Transactions*, vol. 60, pp. 178–190, 2016.

[11] P. A. Martin, A. C. Zanin, and D. Odloak, "Integrating real time optimization and model predictive control of a crude distillation unit," *Brazilian Journal of Chemical Engineering*, vol. 36, no. 3, pp. 1205–1222, 2019.

[12] F. Li, F. Qian, M. Yang, W. Du, and V. Mahalec, "Product tri-section based crude distillation unit model for refinery production planning and refinery optimization," *AIChE Journal*, vol. 67, no. 2, 2020.

[13] Y. Yang and P. I. Barton, "Refinery optimization integrated with a nonlinear crude distillation unit model," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 205–210, 2015.

[14] Y. Xiong, X. Shi, Y. Ma, and Y. Chen, "Optimization design of crude oil distillation unit using bi-level surrogate model," *Frontiers in Control Engineering*, vol. 4, 2023.

[15] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *Journal of Process Control*, vol. 13, no. 4, pp. 291–309, 2003.

[16] ——, "Advanced control using decomposition and simple elements," *Annual Reviews in Control*, vol. 56, p. 100903, 2023.

[17] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[18] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free model-predictive control," *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[19] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[20] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[21] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[22] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[23] D. Krishnamoorthy and S. Skogestad, "Real-time optimization as a feedback control problem – a review," *Computers & Chemical Engineering*, vol. 161, p. 107723, 2022.

4.10 Exercises

4.1 Concept check · 2 min

An engineer adds a second PI loop on the top reflux of this tower, so that the reflux is asked to hold the naphtha end point at 184 °C and the kerosene flash point at 50 °C at the same time. What should the engineer expect?

(a) Both targets are met, because the two loops read different analyzers on separate cycles

(b) Both targets are met after retuning, because a decoupler separates two outputs of one handle

(c) The flash point loop wins, because the flash point moves more than the end point per tonne of reflux

(d) At most one target is met, because both are read from the one cut point that the reflux places

4.2 Concept check · 1 min

With every other draw and the reflux held at the design line-up, the diesel draw is raised by 20 t/h. Which cut points have moved once the tower settles?

(a) The diesel and gas oil cut and the gas oil and residue cut, and none above them

(b) All four cuts, the naphtha and kerosene cut least because it is farthest away

(c) Only the diesel and gas oil cut, because the gas oil draw itself has not changed

(d) The kerosene and diesel cut and the diesel and gas oil cut, on either side of the diesel

4.3 Concept check · 2 min

In scenario 6 the coil outlet falls 22 °C with all four loops in automatic, and the measured overflash drops below its 2 vol% minimum. Which statement describes the override on the gas oil draw?

(a) It replaces the gas oil loop and holds the overflash at its minimum until the coil outlet returns

(b) It scales down the loop's request more as the shortfall grows, which shortens the time below the minimum without removing it

(c) It writes the gas oil loop's integral, so that the loop does not wind up while the draw is cut

(d) It acts only with the gas oil loop in automatic, so a draw held in manual has no protection

4.4 Theory and calculation · 9 min

The top cut follows Equation (4.1) with $v_{1,ref} = 0.200$, $L_{ref} = 150\,\mathrm{t/h}$, $p = 0.45$ and $h = 0$, and the tail follows Equation (4.3) with $\beta = 12.0\,^{\circ}\mathrm{C}$ and $n = 1$, where $R_1 = L/D_N$ and the naphtha made is $D_N = v_1 \rho F$ with $\rho F = 569.3\,\mathrm{t/h}$. Between $v = 0.15$ and $v = 0.25$ the assay is the straight line through 145 °C at 0.15, 175 °C at 0.20 and 205 °C at 0.25. Take steady state, so $T_1 = T_{bp}(v_1)$.

(a) Find v_1, T_1, σ_1, the naphtha end point and the kerosene flash point at $L = 150\,\mathrm{t/h}$. *Checkpoint:* 0.200, 175.0 °C, 9.11 °C, 184.1 °C and 53.9 °C.

(b) Repeat at $L = 170\,\mathrm{t/h}$, and find the ratio of the flash point's change to the end point's change. *Checkpoint:* 0.1890, 168.4 °C, 7.60 °C, 176.0 °C, 48.8 °C and 0.63.

(c) Differentiate at $L = 150\,\mathrm{t/h}$ to find dT_1/dL and $d\sigma_1/dL$. Use them to find the gains of the end point and the flash point, and compare with a and b of Table 4.3. *Checkpoint:* -0.360 and $-0.0881\,^{\circ}\mathrm{C/t/h}$; gains -0.448 and $-0.272\,^{\circ}\mathrm{C/t/h}$.

4.5 Theory and calculation · 8 min

Use the design line-up: charge $F = 662\,\mathrm{m^3/h}$ with $\rho F = 569.3\,\mathrm{t/h}$, $v_1 = 0.200$, draws of 74, 102 and 51 t/h, coil outlet 377 °C, pumparound duties summing to 42 MW and the parameters of Table 4.2. The assay passes through 333 °C at 0.50, 358 °C at 0.55, 383 °C at 0.60 and 410 °C at 0.65, linearly between them.

(a) Find the diesel and gas oil cut point, and how far it moves when the kerosene draw rises by 20 t/h. *Checkpoint:* 337.6 °C, moving 17.6 °C.

(b) Find the fraction vaporized, the overflash, the overflash's sensitivity to the coil outlet and the coil outlet drop, with every draw held, that brings the overflash to 2.0 vol%. *Checkpoint:* 0.6315, 3.28 vol%, 0.185 vol% per °C and 6.9 °C.

(c) Find the rise in preheat and the firing saved per megawatt of pumparound duty, and apply them to the scenario 5 step from 12 to 26 MW. *Checkpoint:* 1.82 °C and 0.75 MW per megawatt; 25.5 °C and 10.5 MW.

4.6 Theory and calculation · 5 min

The loops' plant gains at the design point are −0.448 °C/t/h (reflux to end point), 0.692 °C/t/h (kerosene draw to freeze point), 0.411 °C/t/h (diesel draw to cloud point) and −0.176 vol%/t/h (gas oil draw to overflash), with section time constants of 4, 8, 12 and 15 min.

(a) Find the controller gain $1/|K|$ for each loop and compare with the gains the automatic scenarios load. *Checkpoint:* 2.23, 1.45, 2.43 and 5.68, against 2.5, 1.4, 2.4 and 5.5.

(b) The measured overflash reads 1.9 vol% while the gas oil loop requests 40 t/h. Find the draw the override sends, and the measured overflash below which the cut stops growing. *Checkpoint:* 37.3 t/h; 1.1 vol%.

4.7 App challenge · 8 min

Load scenario 2, One valve, two specifications, and change no setting. The naphtha end point target steps from 184 to 176 °C at 30 min, and at 150 min the selector switches the top loop to the kerosene flash point with a 54 °C target. Run to 360 min at 20 min per wall second (about 18 s) and export the CSV file. Read the columns `naph_ep_C`, `kero_flash_C` and `reflux_tph`.

(a) Read the three at 30 and at 150 min, and find the ratio of the flash point's change to the end point's change. *Checkpoint:* 184.0, 53.9, 150.3 t/h; 176.0, 48.8, 170.1 t/h; ratio 0.63.

(b) After the switch, find the highest end point and its time, and say how long the end point stays above its 185 °C specification. *Checkpoint:* 188.0 °C at 170.4 min, above the specification from 165.6 to 184.8 min.

(c) Read the three at 360 min. *Checkpoint:* 184.3 °C, 54.0 °C and 149.6 t/h.

Accept temperatures within 0.1 °C, the reflux within 0.1 t/h and times within 0.2 min.

4.8 App challenge · 8 min

Load scenario 6, The loop that is not on a product, and change no setting. The coil outlet steps from 377 to 355 °C at 30 min and back at 200 min. Run to 360 min at 20 min per wall second and export the CSV file.

(a) From the columns `overflash_pct` and `ago_override`, find the first row below 2 vol%, the first row with the override on and the last row with it on. *Checkpoint:* 34.2, 35.1 and 58.4 min.

(b) Find the lowest overflash and its time, and read the Wash zone dry tile. *Checkpoint:* 1.887 vol% at 42.2 min; 23 min.

(c) Read `ago_tph` at 200 and at 360 min, and find the largest distance of `kero_freeze_C` from its −47 °C target. *Checkpoint:* 27.5 t/h, 50.4 t/h and 0.09 °C.

Accept times within 0.1 min and other values within one unit in the last digit given.

4.9 App challenge · 9 min

Load scenario 7, Commissioning order, and why the usual rule is backwards here, and keep Control on Loops. The script switches the gas oil, diesel, kerosene and top loops to automatic at 10, 30, 50 and 70 min, and a heavier crude arrives at 120 min.

(a) Run as shipped, at 1 min per wall second from 115 min, and pause between 119.5 and 120.2 min. Add the four MV travel tiles. *Checkpoint:* 3.79 to 3.82 t/h, 3.81 t/h at 120 min.

(b) Reload, run at 1 min per wall second, and switch the Top loop to Auto as the clock reads 10 min and the Kerosene loop as it reads 30 min. Pause as in (a) and add the tiles again. *Checkpoint:* 6.39 t/h with the top loop switched at 10.0 min, about 5.97 to 6.57 t/h for a switch between 9.5 and 10.5 min.

(c) Run both cases on to 180 min, export each file, and find after 120 min the largest distance of `overflash_pct` from 3.3 vol% and of `kero_freeze_C` from –47 °C. *Checkpoint:* 4.99 and 12.02 °C bottom-up, 5.01 and 12.05 °C with the top two loops early.

4.10 Application area · 3 min

In scenario 5 more top pumparound duty only helps: the preheat rises, the firing falls and the kerosene and diesel split sharpens.

(a) Name the effects that bound pumparound duty on a real tower and that this model leaves out. *Checkpoint:* The softer section above the return, flooding and the exchanger network.

(b) The step from 12 to 26 MW saves 10.5 MW of firing. Find the fired duty saved in a year of continuous operation. *Checkpoint:* About 92 000 MWh.

4.11 Application area · 3 min

A planner proposes to use the application to decide whether a new crude cargo will make jet fuel and diesel on specification.

(a) Give two reasons from the model why the application cannot answer this. *Checkpoint:* The heaviness only shifts and tilts the assay, and the properties are boiling-range correlations.

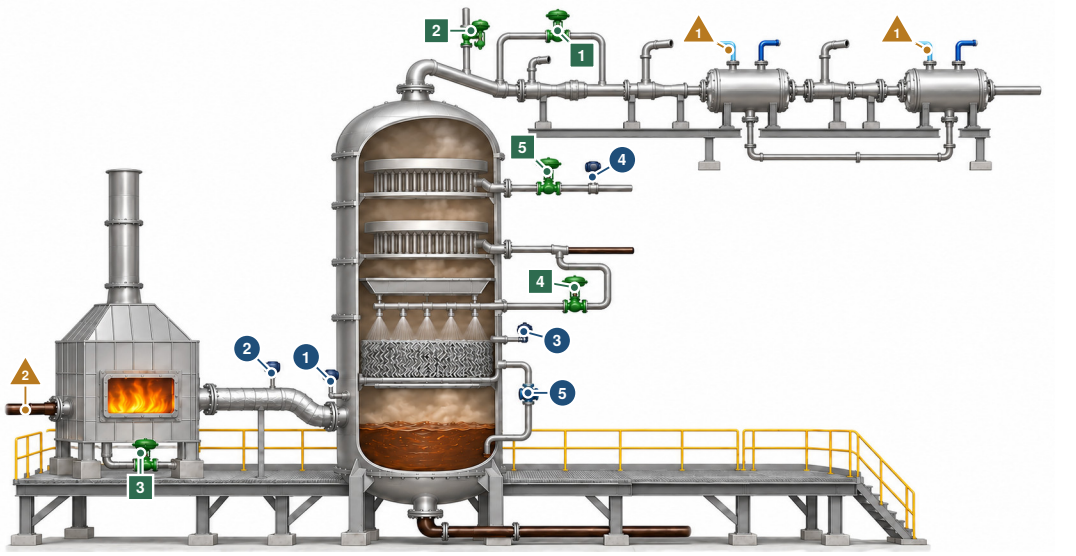
(b) Name two specifications a real kerosene or diesel can fail that the application does not compute. *Checkpoint:* Sulfur or smoke point for kerosene, cetane number for diesel.

4.12 Review the agent · 5 min

An AI assistant is asked how to pass the graded challenge with the least effort and replies as follows. “Set Control to MPC and change nothing else. The scored target is met, with a worst overflash deviation of 0.55 vol%, a worst freeze point deviation of 0.94 °C and 3.0 min dry. The predictive controller moves all four handles together, so it also keeps the furnace inside its 85 MW limit, and the heat moves of the suggested configuration are no longer needed.” Find the flaw and say how to check it.

CHAPTER 5

Vacuum Tower Cut Points Held by an Ejector Curve That Breaks



The controller moves 1 recycle valve around the first steam jet 2 air bleed valve 3 heater firing 4 wash oil valve 5 light gas oil draw valve

The controller reads 1 pressure where the feed flashes 2 heater outlet temperature 3 temperature above the wash bed 4 light gas oil temperature 5 overflash, the liquid that drains from the wash bed

What disturbs it ▲1 warm cooling water ▲2 a heavier feed ▲ air leaking in

Both valves at the first steam jet, at the top of the picture, can only raise the pressure in the tower, so the deepest vacuum is whatever the jets and their cooling water allow. Illustration, not to scale.

IN PLAIN TERMS **A Vacuum That Can Break**

The main idea. The heaviest part of crude oil would break down from the heat before it boiled, so a refinery boils it under a deep vacuum. A deeper vacuum lifts more oil for very little extra heat.

The challenge. Jets of steam make the vacuum, and they can carry away only so much leaked air and gas. Just past that limit the vacuum collapses within a minute, and it returns only when the load is cut well below the breaking point.

Why it matters. Warm cooling water lowers the limit, so a controller that pushes for more oil must watch how close the jets are to it.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the Did you know? boxes in Sections 5.1 and 5.5.2, the History boxes in Sections 5.2 and 5.3.2, the Try it boxes in Sections 5.5.3 and 5.5.4, and Key points in Section 5.8. No equations.
Study	45 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Section 5.5 with Figure 5.4, Section 5.6 and Section 5.7 with its Open challenge box.

The control question The flash-zone pressure is the cheapest handle on a vacuum tower's cut, but the ejectors that hold it break at a load the cooling water sets and recover only well below it. How far can the cut be pushed, and what must a lift controller watch?

Learning objectives

- Compute the shift of the cut per halving of the pressure from a vapor pressure line.
- Calculate an ejector set's critical and recovery loads from its curve and the cooling water.
- Explain why the wash oil holds the heavy gas oil's 95 % point and the overflash protects the wash bed.
- Measure a break and its hysteresis, and explain why an air bleed spends the critical load.
- Explain why a lift controller blind to the ejector load runs the heater to its limit.

Prerequisites Sections 1.4.3, 1.4.2, 1.4.6, 1.4.7 and 1.6.2, and Chapter 4 for the crude curve.

The app https://apmonitor.com/apps/vacuum-tower/

See also Volume 4, Chapter 2 for a limit cycle under a protective override, Volume 4, Chapter 4 for overrides that push throughput, Volume 5, Chapter 10 for an inferred quality corrected by a late laboratory and Volume 5, Chapter 3 for control through a slow analyzer.

5.1 Introduction

The residue of an atmospheric crude tower boils above 383 °C, and its heavier half would crack long before it boiled at atmospheric pressure. The vacuum tower simulated here takes 350 t/h of that residue through a heater to 405 °C and flashes it at 40.0 mbar a (30.0 mmHg), where 63.8 % of it vaporizes. Held at 30 mbar a instead of 50 with the same heater outlet and wash oil, the heavy vacuum gas oil (HVGO) grows from 130.3 to 155.1 t/h and its 95 % point climbs from 545.8 to 570.5 °C, past its 565 °C limit, for 0.7 MW more firing. No other handle on the unit moves the cut that far for so little heat. The pressure is made by three steam ejectors in series. Raising the air leaking into the tower from 120 to 600 kg/h lifts the flash zone only from 40.0 to 45.3 mbar a, but at about 695 kg/h the first stage breaks and the flash zone passes 70 mbar a within a minute, and the stage does not recover until the leak is cut to between 560 and 564 kg/h.

DID YOU KNOW? A Second Tower for the Heavy End of the Barrel

On 1 January 2026 the refineries of the United States could charge 8.4 million barrels per stream day to vacuum distillation, beside 19.2 million barrels of crude oil [1]. At the reduced pressure the columns are built with larger diameters than atmospheric ones, to keep comparable vapor velocities [2].

Two beliefs are what the simulator sets out to correct. The first is that a pressure loop on a vacuum tower is a free handle on the cut, whose cost grows smoothly with how hard it is pushed. In this plant the constraint underneath the loop has a cliff, a jump of 35 mbar on the ejector suction once the intercondenser after the first stage reaches the stage's maximum discharge pressure, and a hysteresis that makes the fix go further than the fault did. The second is that more heat always means more lift. Once the stage has broken the lift falls, and a lift controller that does not watch the ejector load answers by firing harder, which only adds cracked gas, so it drives the heater outlet to its limit. Plants with the same shape are common, among them compressors near surge (Volume 4, Chapter 2), condensers near their cooling limit and any unit whose capacity limit is a fold rather than a ceiling. Selector logic for active constraint switching is the general answer [3], and the industrial practice of pushing a unit against its constraints is usually written as model predictive control [4]. The vacuum tower adds that the constraint's position moves with the afternoon's cooling water.

However, the material a student meets keeps these pieces apart. Ejector analysis treats one stage and its mixing [5], unstable-mode testing treats the ejector set without the tower above it [6], troubleshooting texts describe the tower and its vacuum system from the field [7], and optimization studies of the column and its ejectors work at steady state [8]. This chapter works with a simulated

vacuum unit in which the cut, the ejector curve, the wash bed and the heater interact on one clock. It measures what a pressure step, an air leak, a starved wash bed, a hotter heater, warm cooling water, a heavier residue and a lift controller do to it. The equilibrium flash [9], the boiling point characterization [10, 11], the water vapor pressure [12] and the proportional-integral (PI) loops are all established practice. The results are measurements, taken on one generic unit at one seed. Among them, the app's teaching vapor pressure line shifts the cut 26 to 41 K per halving of the pressure where the design brief for the app expected about 10 K, and this is about a fifth more than the same flash gives with the Maxwell–Bonnell correlation [13]. The model has 17 states and a residue balance that closes to 2.1×10^{-14}. It is built on the same crude curve as the atmospheric tower of Chapter 4 and on a water vapor pressure copied unchanged from the flash steam vessel of Volume 1, Chapter 6, and the chapter gives every equation and default needed to rebuild it. Each of nine cases carries a rule that grants credit only for what the learner does and what the unit then reaches (Section 1.6.4).

Section 5.2 follows the domain from the first vacuum stills to the field practice of ejector troubleshooting. Section 5.3 sets out the unit: the residue and its vapor pressure line, the flash zone, the ejectors with their break, the wash bed and draws, the heater and the instruments, and the numerics and assumptions. Section 5.4 gives the four loops, the high select on the wash oil, the maximizer and the three models in play. Section 5.5 measures the cases: pressure and the cut (Section 5.5.1), the break and its hysteresis (Section 5.5.2), the wash oil, the heater and the cracked gas (Section 5.5.3), summer water and a heavier residue (Section 5.5.4) and the graded challenge (Section 5.5.5). Section 5.6 sets the loops beside the hardware of a working unit. Sections 5.7 and 5.8 state the limitations and conclude.

5.2 Vacuum, Ejectors and the Heavy Gas Oil

Vacuum distillation of petroleum began with lubricating oils, as the History box below relates. Fractions were characterized by their boiling ranges, and a 1933 method approximated the critical and thermal properties of petroleum fractions [10]. A temperature read under vacuum still had to be turned into a normal boiling point before it could be compared with a crude assay, and in 1957 Maxwell and Bonnell published a vapor pressure correlation for petroleum hydrocarbons that does that conversion [13] and that an open-source process simulator still implements. Computers then took over the flash itself, as the History box in Section 5.3.2 relates. A 2005 ASTM manual characterizes petroleum fractions as pseudo-components cut from the true boiling point curve, each with a correlated molar mass, latent heat and vapor pressure [11].

> ⧖ HISTORY **1883 to 1932: Vacuum Stills Begin With Lubricating Oil**
>
> Vacuum distillation of petroleum began with lubricating oils. A book of 1922 lists vacuum stills designed as far back as Henderson's of 1883, all working at relatively low vacuum, and a plant for high vacuum patented by Steinschneider in 1911 [14]. The still was strengthened against the outside pressure, the oil was always distilled with steam, and an air pump held the vacuum on a condenser set at the top of a barometric water column. The same book judged that vacuum had not received the attention it deserved, and it saw no advantage in applying it to the volatile fractions of crude oil [14]. A 1932 account credits the vacuum process with less cracking, a better separation of fractions and a lower heat input and condensing load [15]. Steam still enters the tower of this chapter, but steam ejectors hold its vacuum.

The vacuum came from steam. A one-dimensional analysis of ejectors with constant-pressure and constant-area mixing, checked against experiments, was published in 1950 [5], and multistage condensing ejectors with intercondensers are the subject of a 1994 monograph for the process industries [16]. Such ejectors have an unstable mode in which a critical stage loses its pumping capacity, and a 1994 test procedure measured it and predicted the cyclic region between the stable and unstable modes [6]. Field practice names the same events, the critical discharge pressure exceeded,

Table 5.1 Variables by Role at the Design Point (Case 2 at $t = 0$)

Symbol	Role	Design	Range
p_{fz}	CV, flash-zone pressure (mbar a)	40.0	SP 25–80
COT	CV, coil outlet (°C)	405.0	SP 385–440
$T_{95,inf}$	CV, HVGO 95 % point, inferred (°C)	558.0	SP 530–580
$T_{LV,inf}$	CV, LVGO cut point, inferred (°C)	428	SP 400–460
OF	CV (override), overflash (% of F)	4.43	floor 1–6
u_r, u_b	MV, recycle, air bleed valve (%)	46, 0	0–100
Q	MV, absorbed duty (MW)	22.6	0–40
r	MV, wash oil ratio (% of gas oil)	10.8	2–25
D_L	MV, LVGO draw (t/h)	70.0	20–140
G_{leak}	DV, air leak (kg/h)	120	0–1200
T_{cw}	DV, cooling water (°C)	28	15–40
S_{strip}	DV, stripping steam (t/h)	1.0	0–8
F	DV, residue charge (t/h)	350	250–420
τ_{tail}, m_{AR}	DV, tail stretch (°C), metals (ppm)	0, 40	0–100, 10–100
D_H	output, HVGO on pan level (t/h)	142.3	—
VR	output, vacuum residue (t/h)	137.4	—
T_{fz}, V/F	output, flash zone (°C), lift (%)	389.3, 63.8	—
G_{nc}, G_{crit}	output, vent gas, critical load (kg/h)	384, 964	—
m_H	output, HVGO Ni + V (ppm)	0.47	limit 1.0
T_{skin}	output, tube skin (°C)	489	limit 560

vacuum breaking, the water vapor pressure as a limit, warm cooling water, air leaks, cracked gas from the heater outlet temperature, and a positive feedback loop in vacuum tower pressure control [7].

In a fuels vacuum unit the wash section above the flash zone decides how deep a cut can go at a given gas oil quality [17]. Entrainment from industrial flash zones is typically 3 to 5 %, at least 98 to 99 % of it can be knocked back by a well-designed wash bed, and what passes carries the residue's metals into the HVGO [17]. Designers set the wash oil by an overflash target at the bottom of the bed, typically 0.1 to 0.2 gpm/ft^2 at 60 °F, and on heavy crudes six-year runs between wash-bed replacements are typical [18]. The reasons for a low HVGO yield from residual fuel oil have been examined [19], process models of the whole vacuum unit are part of integrated refinery modeling [20], and the column and its ejector system have been optimized together [8]. The control problem on the regulatory layer, a pressure loop that can only add load, a quality loop that moves the wash oil and a lift pushed against the ejectors, is the subject of this chapter.

5.3 The Unit as Modeled

The simulated unit is generic, with no plant behind it, and its sizes are teaching values. Its layout is shown in Figure 5.1. Atmospheric residue at 350 t/h and 360 °C passes a fired heater to the coil outlet temperature COT, and a transfer line drops it 120 mbar into the flash zone of a damp tower with 1.0 t/h of stripping steam and 0.5 t/h of coil steam. The vapor rises through a wash bed, an HVGO pumparound section with its pan, and a light vacuum gas oil (LVGO) pumparound section. The LVGO is drawn on flow, the HVGO drains its pan on level, and part of it returns as wash oil. What reaches the bottom of the wash bed is the overflash, which joins the vacuum residue to the coker of Chapter 11. Steam, air and cracked gas leave the top to three ejector stages with intercondensers, and a recycle from the first stage's discharge or an air bleed sets the pressure.

Table 5.1 lists the variables by role. The model integrates 17 states: the first-stage suction p_s, the beds' pressure drop Δp, the recycle and bleed valve positions u_r and u_b, the absorbed duty Q, the COT, the two lagged cut points T_{cut} and T_{LH}, the LVGO draw D_L, the wash oil W, the wash bed holdup H_w and coke K_w, the HVGO pan inventory M_{pan}, the coil coke δ, the top temperature T_{top}, and two ledgers of residue fed and products made. Time is in minutes, flows in t/h (gas loads in kg/h), pressures in mbar absolute and temperatures in °C.

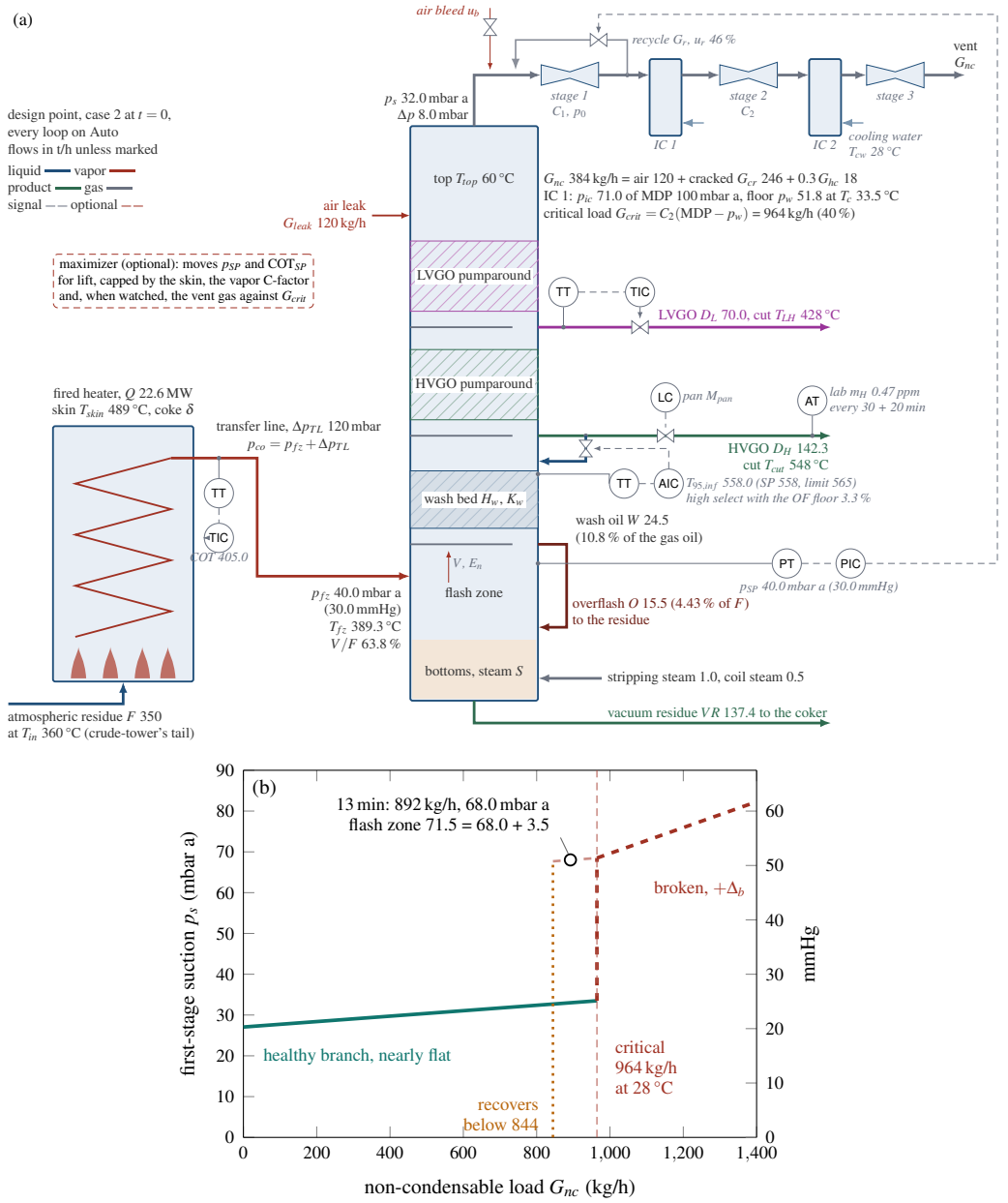


Figure 5.1 The vacuum unit as modeled. (a) Drawn after the Tower view of case 2 at $t = 0$, the design point with every loop on Auto, with the heater at 22.6 MW, the flash zone at 40.0 mbar a and 389.3 °C, the wash bed with its overflash collector, the HVGO pan on level with the wash oil returned above the bed, the LVGO draw, and the ejector set with the recycle around stage 1 and the air bleed. Flows in t/h unless marked. Process model only, not to scale. (b) The first-stage suction against the non-condensable load as the Ejectors view of case 2 draws it at 13 min, after the leak was raised to 700 kg/h and returned to 600 kg/h: the healthy branch to the critical load (thick solid), the 35 mbar jump and the broken branch (thick dashed), the hysteresis band (faint dashed) back to the recovery load (dotted), the critical load (thin dashed) and the operating point (open circle), still broken.

5.3.1 The Residue and Its Vapor Pressure

The feed is the tail of the crude curve of the atmospheric tower app, imported from that app unchanged so that the two towers slice the same crude. On it 40.0 % of the crude boils above 383 °C, 65.3 % of that residue boils below 565 °C and 16.7 % above 650 °C. The residue is cut into 13 slices of about 20 K up to 650 °C and one pitch lump that never vaporizes, and each slice has a vapor pressure on a line through its normal boiling point $T_{b,i}$, as shown in Equation (5.1).

$$\begin{aligned} z_i &= c(T_{i+1}) - c(T_i), \quad c(T) = \frac{v(T) - v_{383}}{1 - v_{383}} \\ \ln \frac{p_i^{sat}}{p_{atm}} &= A\left(1 - \frac{T_{b,i}}{T}\right), \qquad \lambda_i = \frac{ART_{b,i}}{M_i} \\ M_i &= 330\,\mathrm{g/mol}\,(T_{b,i}/673.15\,\mathrm{K})^{2.5} \end{aligned} \tag{5.1}$$

where $v(T)$ is the cumulative fraction of the crude boiling below T on the atmospheric tower's curve and v_{383} its value at 383 °C, z_i the mass fraction of slice i (taken equal to its volume fraction), T and $T_{b,i}$ absolute, $p_{atm} = 1013.25\,\mathrm{mbar}\,a$, R the gas constant, λ_i the latent heat and M_i a teaching molar mass law. A heavier residue stretches the curve above 383 °C by $\tau_{tail} x$ at a distilled fraction x of the residue. With $A = 16$ the design case cuts near 550 °C. The atmospheric equivalent temperature of a temperature T read at a hydrocarbon partial pressure p follows from the same line, $\mathrm{AET}(T,p) = T\,(1 + \ln(p_{atm}/p)/A)$ in kelvin, which is Equation (1.13) of Chapter 1, and its inverse $\mathrm{boil}(T_b, p)$ is the temperature at which a slice of normal boiling point T_b boils at p.

The line is a teaching value, not the Maxwell–Bonnell correlation [13]. Compared with that correlation in the form an open-source process simulator implements it, with a Watson factor of 12, which reproduces n-hexadecane's normal boiling point from its 10 mmHg point within 0.9 K, the line puts the normal boiling point of what boils at 391 °C under 40 mbar a at 525.2 °C against 528.1 °C, and over 380 to 400 °C and 20 to 80 mbar a the two agree within 7.1 K. Their slopes differ more. Halving the pressure at 391 °C lets a component 28.8 K heavier boil on the line, $T \ln 2/A$, and 24.0 K by Maxwell–Bonnell, so the line overstates the shift by about a fifth, the same as a line constant of 19.2.

5.3.2 The Flash Zone

The flash zone is one equilibrium stage with the steam as an inert vapor, solved by the Rachford–Rice equation [9] and reached from the coil outlet at constant enthalpy, as shown in Equation (5.2).

$$\begin{aligned} &\sum_i \frac{n_i(K_i - 1)}{1 + \beta(K_i - 1)} + \frac{n_s}{\beta} = 0, \qquad K_i = \frac{p_i^{sat}(T_{fz})}{p_{fz}} \\ &F c_p\,(\mathrm{COT} - T_{fz}) = H_V(T_{fz}, p_{fz}) \\ &\qquad\qquad - H_V(\mathrm{COT}, p_{fz} + \Delta p_{TL}) \end{aligned} \tag{5.2}$$

where $n_i = F z_i / M_i$ and n_s are the molar flows of the slices and the steam, β is the vapor share of all moles, $K = 0$ for the pitch, $H_V = \sum_i \lambda_i V_i$ is the latent heat carried by the vapor V_i at the named conditions, $c_p = 2.8\,\mathrm{kJ/(kg\,K)}$ and $\Delta p_{TL} = 120\,\mathrm{mbar}$. The temperature is found by the Illinois method on a fixed bracket and the vapor fraction by Newton steps held inside $(0,1)$, so the flash is a pure function of its inputs. At design $T_{fz} = 389.32\,^{\circ}\mathrm{C}$, 15.68 K under the outlet, 223.4 t/h of oil vaporizes, and the vapor is 86.7 % hydrocarbon by moles ($y_0 = 0.867$). A lower pressure therefore lifts more oil at a slightly cooler flash zone.

> **HISTORY 1952: A Flash Calculation Arranged for a Machine**
>
> In 1952 a technical note in a petroleum journal observed an increasing interest in electronic digital computers for research and engineering calculations [9]. A problem, it said, must often be formulated for a machine quite differently from the way it would be arranged for manual solution, to use the machine's speed and to limit the need for number storage. The note arranged the vapor-liquid equilibrium of a flash vaporization as one implicit equation in one variable, a form well suited to sequence-controlled computing equipment [9]. That equation is the form the flash zone of this chapter solves, with the steam added as an inert vapor, in Equation (5.2).

5.3.3 Ejectors, the Break and the Pressure

The first stage empties the tower's vapor space at a capacity proportional to its suction pressure, and the intercondenser after it sets the pressure the stage discharges against, as shown in Equation (5.3).

$$\begin{aligned}
\frac{\mathrm{d}p_s}{\mathrm{d}t} &= \kappa\left[G_1 - 1000\,C_1\,(p_s - p_0 - \Delta_b)\right] \\
G_1 &= 1000\,S + G_{nc} + 0.7\,G_{hc} + G_r \\
G_{nc} &= G_{leak} + G_{cr} + B_{max}u_b + 0.3\,G_{hc} \\
p_{ic} &= p_w(T_c) + G_{nc}/C_2, \quad T_c = T_{cw} + \Delta T_{app} + \Delta T_{st}S \\
G_{crit} &= C_2\,(\mathrm{MDP} - p_w) \\
G_r &= 1000\,R_{max}u_r\sqrt{(p_{ic} - p_s)/50}
\end{aligned} \tag{5.3}$$

where $\kappa = RT/(V_g\bar{M})$ with 333 K, the vapor space $V_g = 2000\,\mathrm{m}^3$ and 19 g/mol, or 0.729 mbar per kg, G_1 is the first stage's load in kg/h (steam S in t/h, light oil G_{hc} of which 30 % is taken as non-condensable, and recycle G_r), $C_1 = 0.15\,\mathrm{t/(h\,mbar)}$, $p_0 = 3\,\mathrm{mbar}\,a$, $C_2 = 20\,\mathrm{kg/(h\,mbar)}$, the maximum discharge pressure $\mathrm{MDP} = 100\,\mathrm{mbar}\,a$, $\Delta T_{app} = 4\,\mathrm{K}$, $\Delta T_{st} = 1\,\mathrm{K}$ per t/h and p_w is the water vapor pressure of IAPWS-IF97 Region 4 [12]. That routine is copied byte for byte from the flash steam app, and its error against the release's verification values is below 1.8×10^{-9} at 300, 500 and 600 K. The stage breaks when p_{ic} reaches the MDP, which sets $\Delta_b = \Delta_{b,0} + k_b \max(0, p_{ic} - \mathrm{MDP})$ with $\Delta_{b,0} = 35\,\mathrm{mbar}$ and $k_b = 0.5$, and it recovers only when p_{ic} falls $h = 6\,\mathrm{mbar}$ under the MDP. Against the non-condensable load the suction $p_s = p_0 + \Delta_b + G_1/(1000C_1)$ is therefore nearly flat to the critical load and then jumps, the declared teaching curve of Figure 5.1b, with G_{crit} falling as the water warms, 964 kg/h at 28 °C, 709 at 32 and 637 at 33 °C. The stage's own slope gives a pressure time constant of $60/(\kappa\,1000\,C_1) = 0.55\,\mathrm{min}$.

The flash zone sits above the suction by the pressure drop of the overhead and the beds, which grows with the vapor's kinetic energy and with coke in the wash bed, as shown in Equation (5.4).

$$\begin{aligned}
p_{fz} &= p_s + \Delta p, \qquad \phi_K = 1 + 0.35\left[(1 - K_w)^{-2} - 1\right] \\
\tau_{\Delta p}\frac{\mathrm{d}\Delta p}{\mathrm{d}t} &= \Delta p_0\,\frac{n^2\bar{M}T/p_{fz}}{(n^2\bar{M}T/p_{fz})_0}\,\phi_K - \Delta p
\end{aligned} \tag{5.4}$$

where n is the vapor's molar flow, $\bar{M}$ its molar mass, T the flash zone temperature, the subscript 0 the design, $\Delta p_0 = 8\,\mathrm{mbar}$ and $\tau_{\Delta p} = 0.1\,\mathrm{min}$. Cracked gas comes from the coil and from the hot bottoms, $G_{cr} = G_{h0}2^{(\mathrm{COT}-405)/12} + G_{b0}2^{(T_{fz}-390)/12}$ with $G_{h0} = 150$ and $G_{b0} = 100\,\mathrm{kg/h}$, and the light oil the top pumparound does not condense is $G_{hc} = 60\exp((T_{top} - 60)/8)$ in kg/h, capped at 4000 kg/h, with the top relaxing in 2 min toward 60 °C, or 175 °C with the pumparound tripped. At design $G_{nc} = 384\,\mathrm{kg/h}$ (air 120, cracked gas 246, light oil 18), 39.8 % of the critical load, $p_{ic} = 71.0\,\mathrm{mbar}\,a$ over a water floor of 51.8 mbar a at 33.5 °C, and the recycle valve at 45.7 % returns 2.42 t/h.

5.3.4 Wash Bed, Draws and the 95 % Point

Residue droplets ride up with the vapor, and the wash bed holds a liquid inventory fed by the wash oil and the droplets it catches, emptied by evaporation and by drainage, as shown in Equation (5.5).

$$\begin{aligned} \frac{\mathrm{d}H_w}{\mathrm{d}t} &= W + (1-P)E_n - E_v - O, \quad E_v = \frac{\varepsilon V H_w}{H_w + H_h} \\ O &= \frac{H_w}{\tau_w}, \quad \omega = \min\Big(1, \frac{O}{O_{wet}}\Big), \quad E_n = e_0\Big(\frac{C}{C_0}\Big)^{n_E} V \\ P &= \min\Big(1,\ P_0\Big(\frac{C}{C_0}\Big)^{1.5}\sqrt{\frac{W_{ref}}{W}} + (1-\omega)^2\Big) \\ \frac{\mathrm{d}K_w}{\mathrm{d}t} &= k_K\,(1-\omega)^2 \end{aligned} \tag{5.5}$$

where $C = n\sqrt{T\bar{M}/p_{fz}}$ is a C-factor index, $e_0 = 0.02$, $n_E = 2$, $P_0 = 0.03$, $\varepsilon = 0.06$, $H_h = 0.01\,\mathrm{t}$, $\tau_w = 8\,\mathrm{min}$, the wetting minimum O_{wet} is 1.5 % of the feed, W_{ref} is 12 % of the design vapor and $k_K = 0.0025$ per minute. The wash oil actually delivered, written W below, is $W\min(1, M_{pan}/M_{lo})$ with $M_{lo} = 2\,\mathrm{t}$, so a starved pan starves the bed. At design 2.0 % of the vapor is entrained, the level a packing vendor reports for a well-designed vapor horn in tests and below the 3 to 5 % of industrial flash zones, and a wet bed catches 96.9 % of it, between the 95 % of that vendor's poorest case and the 98 to 99 % it calls attainable [17]. A dry bed passes 99.4 % and loses 17 % of its open area to coke in the 80 min case 4 keeps it dry, a compressed teaching rate against the six-year runs reported on heavy crudes [18].

The HVGO drains its pan on level, so its yield is what the lift leaves after the LVGO, the gas and the wash oil, and the cuts follow the products along the crude curve, as shown in Equation (5.6).

$$\begin{aligned} \dot{M}_{pan} &= V + E_v + PE_n - D_L - G_{hc} - W - D_H \\ D_H &= (M_{pan} - M_{lo})/\tau_{pan}, \quad VR = L_f - E_n + O \\ \tau_{cut}\dot{T}_{cut} &= T_{AR}\big((D_L + D_H)/F\big) - T_{cut} \\ T_{95} &= T_{cut} + \beta_H/\sqrt{W/D_H} \end{aligned} \tag{5.6}$$

where E_v is the evaporation of Equation (5.5), $T_{AR}(x)$ the residue's curve at distilled fraction x, $L_f = F - V - G_{cr}$ the flash-zone liquid after cracking, $\tau_{pan} = 3\,\mathrm{min}$, $\tau_{cut} = 6\,\mathrm{min}$ and $\beta_H = 4\,\mathrm{K}$. The LVGO cut T_{LH} follows $T_{AR}(D_L/F)$ in 4 min. The HVGO metals are what distills with a deep cut plus the droplets that pass the bed, $m_H = [0.4\,2^{(T_{95}-560)/15}D_H + PE_n m_{AR}F/L_f]/D_H$ in ppm. Of the design 0.47 ppm, 0.36 distills and the rest is entrainment. The residue ledgers close to 2.1×10^{-14} of the throughput over a disturbed 45 min run, and to at most 4.5×10^{-14} over four other seeds and the halved step.

5.3.5 Heater and Instruments

The firing reaches the coil through a 1 min firebox lag, and the coil's heat capacity integrates the duty against the sensible heat and the latent heat of what the coil outlet has already vaporized, as shown in Equation (5.7).

$$\begin{aligned} C_{coil}\frac{\mathrm{d}\,\mathrm{COT}}{\mathrm{d}t} &= 60\Big[Q - \frac{F h_{need}}{3600}\Big] \\ h_{need} &= c_p(\mathrm{COT} - T_{in}) + H_V^{co}/F \\ T_{skin} &= \mathrm{COT} + 45\,Q/Q_{ref} + \beta_c\,\delta \\ \frac{\mathrm{d}\delta}{\mathrm{d}t} &= k_c\,2^{(T_{film}-450)/10} \end{aligned} \tag{5.7}$$

where H_V^{co} is the latent heat of the vapor already made at the coil outlet (the skin follows the form of Equation (1.3)), $C_{coil} = 70\,\mathrm{MJ/K}$, Q in MW, F in t/h, $Q_{ref} = 26\,\mathrm{MW}$, $T_{film} = \mathrm{COT} + 45Q/Q_{ref}$,

Table 5.2 Parameters: A, Adjustable in the Parameter Panel. H, Hidden. S, Set in the Source. All Are Teaching Values

Symbol	Meaning	Value	Unit	
A	vapor pressure constant	16	—	A
c_p, T_{in}	residue heat capacity, inlet	2.8, 360	kJ/kg/K, °C	A
Δp_{TL}	transfer line drop	120	mbar	A
C_{coil}, τ_{fb}	coil capacity, firebox lag	70, 1.0	MJ/K, min	A
Q_{max}	absorbed duty limit	40	MW	A
β_c, δ_0	skin rise per coke, start coke	15, 3	K/mm, mm	A
k_c	coking rate at 450 °C film	0.004	mm/day	A
Q_{ref}, film	film rise reference	26, 45	MW, K	H
G_{h0}, G_{b0}	cracked gas, coil, bottoms	150, 100	kg/h	A
—	rise that doubles cracking	12	K	A
p_0, C_1	stage 1 shutoff, slope	3, 0.15	mbar a, t/h/mbar	A
MDP, C_2	max discharge, stage 2 slope	100, 20	mbar a, kg/h/mbar	A
ΔT_{app}, ΔT_{st}	approach, rise per t/h steam	4, 1	K	A
$\Delta_{b,0}$, h	jump at break, hysteresis	35, 6	mbar	A
k_b	jump per mbar over the MDP	0.5	—	H
V_g	tower vapor space	2000	m^3	A
R_{max}, B_{max}	recycle, bleed capacity	6, 1500	t/h, kg/h	A
Δp_0	beds' drop at design	8	mbar	A
e_0, n_E	entrainment, exponent	0.02, 2	—	A
P_0, ε	wet-bed pass, evaporation	0.03, 0.06	—	A
τ_w, O_{wet}	bed drain time, wetting minimum	8, 1.5	min, % of F	A
k_K	dry-bed coking rate	0.0025	per min	A
τ_{pan}, τ_{cut}, τ_{LH}	pan, cut lags	3, 6, 4	min	A
β_H, m_{d0}	tail spread, metals at 560 °C	4, 0.4	K, ppm	A
—	95 % point, metals, OF limits	565, 1.0, 3	°C, ppm, %	A
$T_{skin,max}$	tube skin limit	560	°C	A
—	slice width, pitch cut	20, 650	K, °C	S
—	light oil taken as non-condensable	0.3	—	S

$\beta_c = 15\,\mathrm{K/mm}$, $\delta_0 = 3\,\mathrm{mm}$ and $k_c = 0.004\,\mathrm{mm/day}$. The run left is the skin's margin to 560 °C over its rate of rise, 1768 days at design. Only a measurement packet reaches the controllers. Pressures lag 0.03 min, the COT 0.1 min, the wash-bed and LVGO thermocouples 0.5 min and the flows 0.1 to 0.2 min, with colored noise of 0.2 min correlation scaled to 0.08 mbar on p_{fz}, 0.06 mbar on p_s, 0.15 K on the COT, 0.25 K on the two thermocouples and 3 kg/h on the vent gas. The metals laboratory samples every 30 min and reports 20 min later, a sampled delay of the kind Equation (1.12) describes. Both cut points are inferred, as shown in Equation (5.8).

$$T_{95,inf} = \mathrm{AET}(T_{wb}, y_0 p_{wb}), \quad p_{wb} = p_{fz} - 0.35\,(p_{fz} - p_s) \tag{5.8}$$

where $T_{wb} = \mathrm{boil}(T_{95}, y\,p_{wb})$ is the vapor leaving the wash bed with y the true hydrocarbon share. The LVGO cut is inferred the same way from its draw temperature at p_s. A change in steam changes y, which the inferential does not see (Section 1.4.6).

5.3.6 Initialization, Numerics and Assumptions

Each case starts lined out on its setpoints. A steady-state solve puts the flash zone on its setpoint with the selected valve where it must be, the heater on its COT, the wash bed and pan at their balances and, for the quality loops on Auto, the LVGO draw and the wash ratio where the inferred cut points sit on their setpoints, never below the ratio that holds the overflash floor. The integrator is the engine's classical fourth-order Runge–Kutta method (Section 1.6.1) at 0.02 min (1.2 s), the controllers run every 0.04 min and the historian logs ten rows a minute. Halving the step moves the 235 quantities of the case measurements by at most one printed digit, such as case 2's break time (10.04 to 10.02 min) and the challenge's break (352.4 to 352.3 min), except the count of case 3's limit cycle, 21 to 22 in 20 min. Table 5.2 gives the parameters.

Table 5.3 Loops, Their Defaults and Limits

Loop	Measurement → handle	K_c, τ_I	Output
PIC	p_{fz} → recycle or bleed	2.5 %/mbar, 0.3 min	0–100 %
TIC	COT → absorbed duty	0.5 MW/K, 3 min	0–40 MW
AIC	$T_{95,inf}$ → wash ratio	0.4 %/K, 10 min	2–25 %
OF floor	overflash → wash ratio	1 %/%, 10 min	2–25 %
TIC	$T_{LV,inf}$ → LVGO draw	1.0 t/h per K, 5 min	20–140 t/h

The model is lumped. The flash zone is one equilibrium stage, and the stripping, wash and pumparound sections are lagged cut points with one tail spread, so it says nothing about tray or packing hydraulics, flooding or weeping. Mass fractions equal volume fractions, and the pitch above 650 °C never boils. The vapor pressure line runs about a fifth steeper than Maxwell–Bonnell. The ejector set is a declared curve, real curves and maximum discharge pressures come from the vendor and differ stage by stage [16], and the unstable region of a real ejector is a band of cyclic operation [6] rather than a latched flag. Cracked gas, entrainment, wash bed coking, metals and coil coking are teaching laws with declared constants, and the dry wash bed cokes in minutes so a case can show it. Only the mass balance closes exactly, while the steam passes through and the energy balance is used only for the coil and the transfer line.

5.4 Loops and Maximizer as Implemented

5.4.1 The Loops

Four loops close on the measurement packet, listed in Table 5.3. The Manual and Auto switch in the header overrides all, and each loop also has its own switch, so a loop on Man leaves its station with the operator. The pressure, heater and LVGO loops run the velocity-form PID law of Section 1.6.2 with the derivative off, as shown in Equation (5.9).

$$\begin{aligned} e_k &= \sigma\,(r_k - y_k), \qquad s_k = s_{k-1} + K_c\,(e_k - e_{k-1}) \\ \hat{u}_k &= s_k + I_{k-1}, \qquad u_k = \min(u_{max}, \max(u_{min}, \hat{u}_k)) \\ I_k &= I_{k-1} + \left[\frac{K_c}{\tau_I} e_k + \frac{u_k - \hat{u}_k}{\tau_I}\right]\Delta t \end{aligned} \tag{5.9}$$

where r_k is the setpoint, y_k the measurement, $\sigma = +1$ for all three (more recycle raises the pressure, more firing the COT, more LVGO the LVGO cut) and $\Delta t = 0.04$ min. In the symbols of Chapter 1, s_k is the running sum S_k of Equation (1.17), $\hat{u}_k$ is the output v_k before limiting of Equation (1.18) and the tracking time T_t of Equation (1.19) equals τ_I. The recycle only adds load to a stage that pulls as hard as it can, so the loop can raise the pressure above what the ejectors pull but never below it. A loop that is not driving its valve tracks it, so every transfer is bumpless.

The HVGO quality loop has no draw to move, so it moves the wash oil. Its ratio is the high select of two positional PI branches, the 95 % point and the overflash floor of 3.3 %, $r = \max(K_W e_{95} + I_W,\ K_O e_{OF} + I_O)$ with $e_{95} = T_{95,inf} - T_{95,SP}$ and $e_{OF} = \mathrm{OF}_{min} - \mathrm{OF}$. Each branch's integral is pulled toward the applied ratio over its own integral time instead of being set onto it, because when the app was built the shared PID's exact tracking rectified the 95 % point's noise through the select into a one-way climb of the wash ratio. The wash oil is r times the measured lifted gas oil, $D_L + D_H + O$.

5.4.2 The Maximizer

In Maximizer mode a constraint-pushing selector moves the pressure and COT setpoints every minute. It is not an optimization and solves nothing. Its lift signal is $e = (r^* - r)/r^* + (T_{95,SP} - T_{95,inf})/20$, filtered over 10 min, with $r^* = 11$ % the target wash ratio. Short of lift by more than 0.05, it lowers the pressure setpoint first, while the valve has room and the C-factor index is under 125 % of design, and then raises the COT. With lift to spare it lowers the COT first, down to 395 °C. Each move is at

Table 5.4 The Nine Cases

#	Settings and events	Teaching point
1	50 mbar a, wash oil by hand, learner lowers p_{SP} to 30	pressure moves the cut
2	PIC on Man, learner raises the air leak	the break, hysteresis
3	p_{SP} 40, 36, 44, 40 at 3, 8, 13 min, then air bleed	recycle against bleed
4	AIC on Man, learner cuts the wash oil to 4 %	a dry bed, metals, coke
5	design, learner raises the COT to 420 °C	the heater limit
6	leak 520 kg/h, PIC on Man, maximizer with load ignored	controller runaway
7	leak 250 kg/h, water 28 to 35 °C over 10–70 min	summer vacuum
8	tail +60 °C, 60 ppm metals at 10 min	an infeasible yield
9	challenge, 8 h: leak +300 at 2 h, water +5 °C from 5 h	the whole problem

most 0.5 K or 0.5 mbar per minute. Overrides cut the COT when the skin comes within 5 K of its limit and, when the ejector load is watched, when the measured vent gas comes within 15 % of the critical load the vendor curve gives at the measured water, and by 1 K a minute while the stage's discharge alarm reports it broken, and the pressure setpoint rises when the C-factor index passes its limit. With the load ignored, the maximizer does not know the ejectors exist. This is a selector and not model predictive control [4], and its structure is the active constraint switching of selector logic [3] that Section 1.4.7 places between the regulatory loops and optimization.

5.4.3 Three Models

The plant is Section 5.3. The identified model a learner can fit is the gain from pressure to cut in case 1 or from leak to pressure in case 2. The controllers' models are fixed numbers, the inferentials' design hydrocarbon share $y_0 = 0.867$ and the vendor curve $C_2(\mathrm{MDP} - p_w(T_c))$ evaluated with the parameters the case started with. The 565 °C 95 % point, the 1.0 ppm metals limit, the 3 % overflash minimum, the 100 mbar a MDP and the 560 °C skin are teaching values. The real ones come from the HVGO specification of the cracker downstream, whose catalyst the metals poison [17], the overflash the wash section is designed for [18], the ejector vendor's curves [16] and the tube alloy's design metal temperature.

5.5 Nine Cases, Measured

The cases are listed in Table 5.4, and Table 5.5 reads the application's own scorecard tiles at the instants the text uses. The measurements were made headless at the default seed with the engine stepping as the page does, the learner's moves applied at the times given. All nine cases produce identical logs whether opened first, opened after a different case or rerun with Reset, and the application's 43 checks pass. The exported file logs every instrument reading next to the simulated value it reads, the latter under a truth prefix (Section 1.6.4). Numbers that move with the noise were repeated at four more seeds (0, 1, 2, 3) and with the step halved.

5.5.1 Pressure Moves the Cut

Case 1 holds the COT at 405 °C with the wash oil fixed and lowers the flash zone from 50 to 30 mbar a at 5 min. An hour later the cut has moved 24.5 K and the 95 % point 24.7 K, from 545.8 to 570.5 °C, passing 565 °C 12.8 min after the change. The line alone gives 21.2 K for 50 to 30 mbar a. The HVGO rises from 130.3 to 155.1 t/h, the residue falls from 149.2 to 125.2 t/h, the firing rises from 22.2 to 22.9 MW because more of the charge is vapor at the coil outlet, and the metals rise from 0.27 to 0.87 ppm with the faster vapor. The flash zone runs 14.5 K under the outlet at 50 mbar a and 17.2 K at 30.

Pause. Try Exercise 5.1 before reading on.

Over a wider range the shift is shown in Figure 5.2a. At a fixed 405 °C outlet the cut of the lifted oil moves 30.3 K from 40 to 20 mbar a, 41.2 K from 80 to 40 and 26.0 K from 50 to 30, and the same flash with Maxwell–Bonnell moves it 26.1, 35.4 and 22.1 K. The design brief for the app expected

Table 5.5 The Application's Scorecard Tiles at the Instants the Text Reads

Case	Settings and instant	HVGO (t/h)	Ejector load (% of critical)	HVGO 95 % (°C)	Metals (ppm)	Overflash (% of F)	Run left (days)	Also
1	50 mbar a, 5 min	130.3	41	545.8	0.27	4.00	1877	
1	30 mbar a from 5, 65 min	155.1	39	570.5	0.87	5.52	1689	Q 22.9 MW
2	leak 600 from 6, 10 min	138.0	90	556.5	0.42	4.11	1796	p_{fz} 45.3
2	leak 700 at 10, 600 at 18, 24 min	111.3	92	532.0	0.13	3.10	2062	broken, p_{fz} 71.3
2	leak 500 from 24, 40 min	138.9	80	551.4	0.36	3.54	1795	recovered
3	bleed from 15, 35 min	141.8	81	557.9	0.48	4.55	1777	22 breaks
4	wash 4 % from 5, 65 min	157.9	40	580.7	4.16	0.04	1779	bed dry
4	wash 11 % from 95, 130 min	141.8	40	557.6	0.47	4.50	1798	coke stays
5	COT 420 from 5, 35 min	146.2	78	560.3	0.56	10.80	192	Q 28.7 MW
6	target 18 % from 5, 100 min	142.2	149	558.0	0.40	3.42	105	COT 425, p_{fz} 85.0
6	watched from 100, 200 min	143.1	85	558.0	0.47	5.49	1271	COT 407.4
7	as shipped, 70 min	129.7	111	556.3	0.38	4.12	1782	broken, p_{fz} 57.8
7	COT 398 from 40, 76 min	130.4	89	546.6	0.30	3.35	4792	never broke
8	as shipped, 120 min	114.5	41	546.5	0.28	3.30	2524	
8	maximizer from 30, 180 min	127.5	34	560.0	0.59	3.91	4150	p_{fz} 25.1, COT 400.8
9	shipped, 8 h	122.2	109	537.9	0.18	3.30	1939	1097 t, broke
9	suggested, 8 h	138.6	86	555.9	0.52	3.30	9921	1154 t, no break

about 10 K per halving, which would need a line constant of 46, far outside an entropy of vaporization, so the app reports the shift as the physics gives it and states the fifth by which its line overstates it. A 10 K hotter outlet moves the cut 18.8 K.

5.5.2 The Break and Its Hysteresis

Case 2 puts the pressure loop on Man, so the recycle valve stays where it is. From 120 to 600 kg/h of leak the flash zone rises only from 40.0 to 45.3 mbar a, because the first stage's suction follows its total load and the air is a small part of it. The intercondenser rises 5 mbar per 100 kg/h of non-condensable gas over its 51.8 mbar a floor, and held 20 min a leak of 690 kg/h leaves the stage critical while 694 breaks it (696 at steady state). Stepped to 700 kg/h at 10 min, the stage breaks 0.04 min later, the flash zone passes 60 mbar a within 20 s, reads 70.2 mbar a after a minute and 72.4 mbar a after a minute and a half, and the HVGO falls from 138.0 to 112.3 t/h by 18 min. Returned to 600 kg/h the stage stays broken, the situation of Figure 5.1b at 13 min, with 892 kg/h of non-condensable gas, a suction of 68.0 mbar a and a flash zone of 71.5 mbar a, 3.5 mbar above it because the beds' drop falls with the denser vapor. On a staircase of 2 kg/h every 3 min the stage breaks at 696 kg/h and recovers at 560 kg/h, as shown in Figure 5.2b, and raised every 2 min instead it breaks at 700 to 702 kg/h and recovers at 562 to 564 kg/h across four further seeds. The hysteresis alone is 6 mbar, or 120 kg/h at the second stage's slope, and the loop is wider because a broken stage's higher pressure raises the bottoms temperature and with it the cracked gas.

🕮 **Pause.** Try Exercise 5.2 before reading on.

Case 3 compares the two ways of raising the pressure. On the recycle, which loads the first stage only, the 36 mbar a step settles within 0.3 mbar in 1.06 min with the vent gas at 380 kg/h of 964. With both valves shut the ejectors pull the flash zone to 28.4 mbar a. An air bleed loads every stage, and holding 30 mbar a on it takes 370 kg/h of air and 77 % of the critical load, against 39 % on the recycle, while 31 mbar a would need 593 kg/h, 101 %. At 40 mbar a on the bleed the loop opens it, the stage breaks, the pressure overshoots, the bleed closes and the stage recovers, 21 times in 20 min with the bleed selected from the start and 22 in the 20 min after a switch at 15 min. The cycle is the cyclic region a real ejector shows between its stable and unstable modes [6] only in appearance, because here it comes from a latched curve under a loop whose setpoint the bleed cannot reach, and its count depends on the step, 21 at the default and 22 at half of it. No tuning removes it.

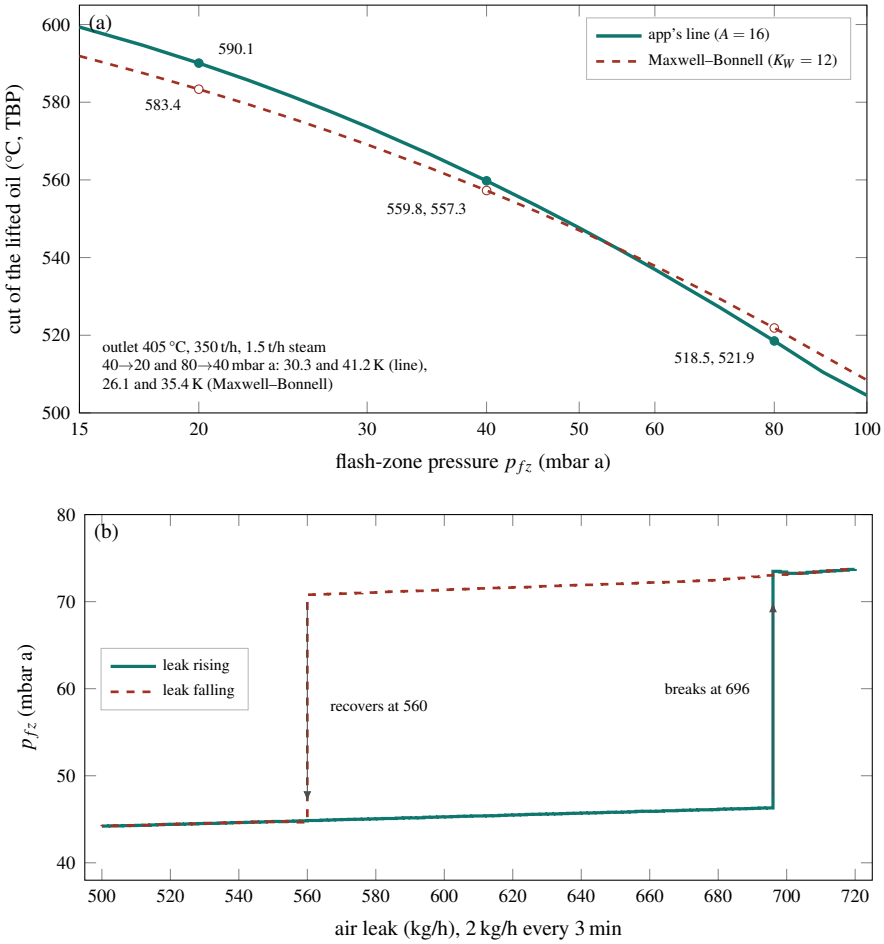


Figure 5.2 (a) The cut of the lifted oil on the crude curve against the flash-zone pressure at a 405 °C outlet, from the application's flash with its vapor pressure line (solid, filled markers) and from the same flash with Maxwell–Bonnell (dashed, open markers). (b) Case 2 with the pressure loop on Man and the air leak stepped 2 kg/h every 3 min from 500 to 720 and back, the flash-zone pressure at the end of each step, with the leak rising (solid) and falling (dashed).

✦ DID YOU KNOW? The Same Leak, Two Pressures

From the simulator (Table 5.5). With 600 kg/h of air leaking in, the flash zone holds 45.3 mbar a and the tower makes 138.0 t/h of heavy gas oil. After the leak has touched 700 kg/h and returned to the same 600 kg/h, the flash zone stands at 71.3 mbar a and the tower makes 111.3 t/h. Only a cut to 500 kg/h brings the stage back.

5.5.3 Wash Oil, Heater and Cracked Gas

Case 4 cuts the wash ratio from 11 to 4 % at 5 min. The 95 % point passes 565 °C 2.2 min later (2.1 to 2.2 at four other seeds), the inferred value 0.3 min after that, and the bed dries 9.5 min after the cut. The metals pass 1 ppm a minute later and settle at 4.2 ppm, nine times the design, while the 95 % point climbs to 581 °C and droplets make 2.8 % of the HVGO. The laboratory reports 4.18 ppm at 50 min, from the sample drawn at 30, 34.5 min after the truth crossed the limit. By 95 min coke covers 17 % of the bed's open area and the beds' drop has risen from 8.0 to 9.3 mbar. Restored to

11 %, the bed is wet in 3.7 min and the metals return to 0.46 ppm, but the coke stays at 18 % and the recycle holds the flash zone with 41.5 % of valve instead of 46.0. This bed dries and cokes as an extreme for teaching, against the multi-year runs a real one makes [18].

Case 5 raises the COT from 405 to 420 °C. With the 95 % point held, half an hour later the HVGO has risen only from 142.3 to 146.2 t/h, while the overflash has grown from 4.4 to 10.8 % of the feed. The extra lift goes to the wash. The coking rate rises 5.9 times, from 0.0026 to 0.0157 mm/day, the run left falls from 1786 to 192 days on 28.7 MW instead of 22.6, the skin rises from 489 to 515 °C, and the cracked gas rises 2.5 times, so the vent gas goes from 384 to 756 kg/h, 78 % of the critical load.

> **TRY IT IN THE APP** **5 · The heater limit**
>
> Load the scenario and press Run. Then raise the Heater outlet setpoint from 405 to 420 °C. Half an hour later the heavy gas oil has risen only from 142.3 to 146.2 t/h, while the overflash has grown from 4.4 to 10.8 % of the feed. The extra lift went to the wash.

Case 6 runs a tower leaking 520 kg/h of air, 81 % of the critical load, with the maximizer in charge and its ejector guard off, and the learner raises its wash ratio target to 18 % at 5 min. The COT climbs and the flash zone creeps up only to 42.1 mbar a by 30 min. At 33.3 min (33.2 to 33.7 over seeds), with the COT at 413.9 °C and 428 kg/h of cracked gas, the stage breaks, the lift falls, and the maximizer, still short of lift, raises the COT to its 425 °C ceiling by 57.0 min. At 100 min the flash zone is at 85.0 mbar a with 901 kg/h of cracked gas, 3.7 times the design. Watched from 100 min, the maximizer cuts the COT until the stage recovers 19.4 min later and then holds the vent gas 15 % under the critical load near 407 °C. The runaway goes only through the controller. Measured as the closed-loop response to 40 kg/h leak steps, the physical loop gain from vent gas through pressure and flash zone temperature back to cracked gas is 0.0057 to 0.0060 on the flat curve and 0.0131 to 0.0140 on the broken branch with the COT on its loop, and 0.022 and 0.063 with the firing fixed, far below the 1 a runaway needs, so the amplification $1/(1-L)$ of Equation (1.8) stays close to one. As shipped, with the target at 11 %, the COT setpoint never moves in 8 h, at the default seed or at any of four others, the COT stays within 0.11 K of 405 °C in the first hour, and the stage never breaks. Field practice describes a positive feedback loop in vacuum tower pressure control [7], and the case teaches that the loop that runs away here is the one a constraint pusher closes when it does not know the sign of its gain has changed.

Pause. Try Exercise 5.3 before reading on.

5.5.4 Summer Water and a Heavier Residue

Case 7 warms the cooling water from 28 to 35 °C over an hour from 10 min on a tower leaking 250 kg/h. Seven degrees add about 24 mbar to the water floor out of a 48 mbar margin. The critical load falls from 964 to 744 kg/h by 40 min and to 604 at 56.6 min, when CONDENSER FLOOR appears, and at 66.7 min, with the water at 34.6 °C, it meets the 515 kg/h the tower sends and the stage breaks. The flash zone goes to 57.6 mbar a and the HVGO falls from 142.3 to 123.9 t/h. A 398 °C COT at 40 min cuts the cracked gas to 162 kg/h and the stage never breaks, for 130.4 t/h. Cutting the stripping steam to zero also holds, with 142.2 t/h, but the inferred 95 % point then reads 558 °C against a true 554 °C, because the inferential assumes the design steam share. A soft sensor calibrated for one condition misleads in another in the same way in the spray dryer of Volume 5, Chapter 10.

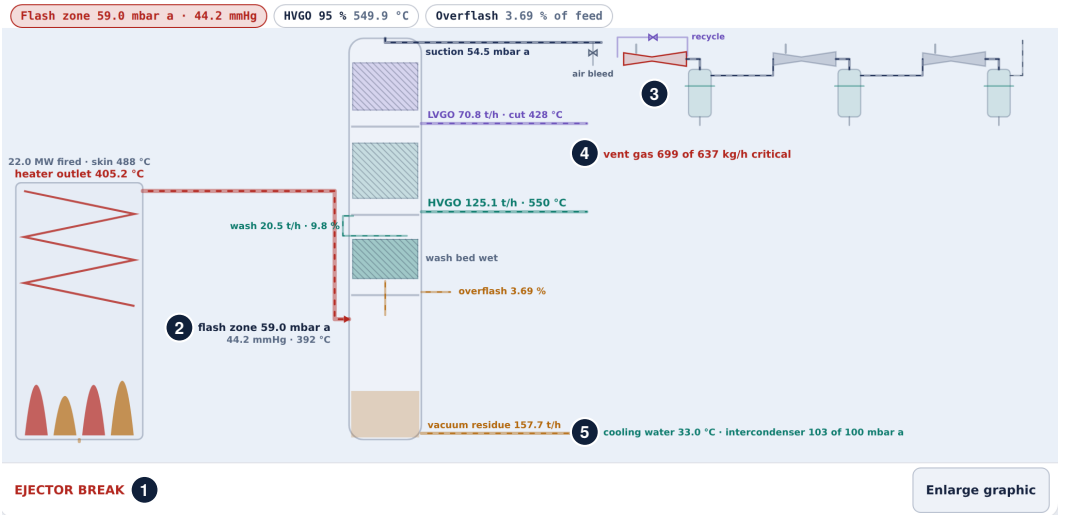


Figure 5.3 The challenge as shipped at $t = 6$ h on the Tower view, light theme, cropped to the plant view with numbered callouts. The stage broke at 352 min, so the status strip reads EJECTOR BREAK (1) and the flash zone reads 59.0 mbar a (44.2 mmHg) at 392 °C (2) against its 40 mbar a setpoint, with the suction at 54.5 mbar a. The first stage is drawn in red, with the recycle around it and the air bleed ahead of it (3). The vent gas of 699 kg/h stands over the critical load of 637 kg/h that the vendor curve gives at that water (4). The cooling water has warmed to 33.0 °C and the intercondenser reads 103 of 100 mbar a (5). The HVGO has fallen to 125.1 t/h with its inferred 95 % point at 549.9 °C and the residue risen to 157.7 t/h, with the COT held at 405.2 °C on 22.0 MW. In the full window, outside this crop, the recycle valve stands shut at 0 %, and the trends over a 4 h window show the pressure jump at the right edge and the vent gas crossing the falling critical line, with 262 kg/h of cracked gas and a true 95 % point of 549.2.

> **TRY IT IN THE APP** **7 · Summer cooling water**
>
> Load the scenario and press Run. Nothing has to be set. Watch the critical load line on the Ejectors view while the cooling water warms from 28 to 35 °C. CONDENSER FLOOR appears at 56.6 min. At 66.7 min the critical load meets the 515 kg/h the tower sends, the stage breaks, and the flash zone goes to 57.6 mbar a.

Case 8 switches to a residue whose curve is stretched 60 °C at its heavy end and carries 60 ppm of metals. Its share below 565 °C falls from 65 to 56 % and its pitch rises from 17 to 24 %. The HVGO falls from 142.3 to 114.5 t/h and the residue rises from 137.4 to 177.8 t/h, with the overflash on its 3.3 % floor and the 95 % point at 546.5 °C, under its 558 °C setpoint. The maximizer, on from 30 min, takes the pressure to its 25 mbar a floor and the COT near 401 °C for 127.5 t/h, 13 of the 28 lost. A sweep of steady states over the COT from 395 to 440 °C and every reachable pressure, each with the least wash meeting the 95 % point and the overflash, finds at most 134.2 t/h inside every limit, at 425 °C and 30.5 mbar a with the ejectors at 99 % of critical, against the design's 142.3 t/h. Above 425 °C nothing is feasible because the ejectors break. The case is honestly infeasible, and the design residue on the same sweep reaches 151.9 t/h.

5.5.5 The Challenge

The challenge runs 8 h. An air leak of 300 kg/h opens at 2 h, leaving the ejectors at 71 % of their critical load, and the cooling water warms 5 °C over an hour from 5 h. It passes with at least 1140 t of HVGO, no break, the 95 % point at most 565 °C, metals at most 1.0 ppm, overflash at least 3 % and

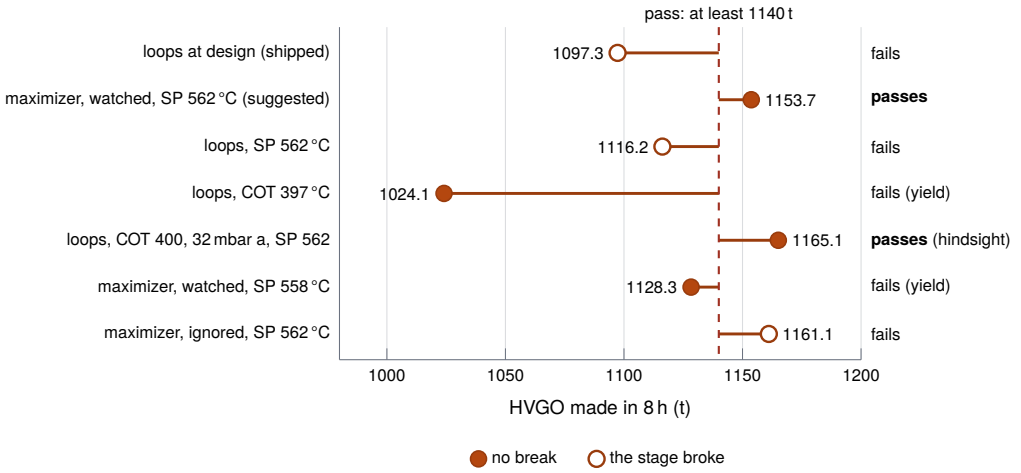


Figure 5.4 The runs of Table 5.6 as a chart: the HVGO made in 8 h, each stem drawn from the 1140 t the challenge asks for, with a filled marker where the stage never broke and an open marker where it broke. Notice that a run passes only to the right of the dashed line with a filled marker, and that the maximizer with the ejector load ignored makes more than the suggested run and still fails.

Table 5.6 The Challenge Over 8 h, Default Seed

Settings	HVGO (t)	Break	Verdict
loops at design (shipped)	1097.3	352.4 min	fails
maximizer, watched, SP 562 °C (suggested)	1153.7	none	passes
loops, SP 562 °C	1116.2	yes	fails
loops, COT 397 °C	1024.1	none	fails (yield)
loops, COT 400, 32 mbar a, SP 562	1165.1	none	passes (hindsight)
maximizer, watched, SP 558 °C	1128.3	none	fails (yield)
maximizer, ignored, SP 562 °C	1161.1	yes	fails

the skin under its limit, judged on the truth at every step after 10 min. As shipped, on fixed setpoints at the design point, the stage breaks at 352.4 min, when the water has reached 32.4 °C and the critical load has fallen to 683 kg/h against 684 kg/h of vent gas, the flash zone rises to 59 mbar a, and the run makes 1097.3 t, as shown in Figure 5.3. The settings tried are listed in Table 5.6, and Figure 5.4 draws them against the yield the challenge asks for. The suggested maximizer, watched, with the 95 % setpoint at 562 °C, cuts the COT as the critical load falls and lowers the pressure to win the lift back, and it passes with 1153.7 t and the 95 % point never above 563.1 °C. Over four further seeds and the halved step the shipped run makes 1096.9 to 1097.4 t and breaks between 352.0 and 352.4 min, and the suggested one makes 1152.7 to 1153.7 t with no break. Its benefit is adaptation without hindsight. A fixed 400 °C at 32 mbar a with the same setpoint passes with more, chosen knowing the afternoon.

5.6 Technology in Practice

Figure 5.5 draws the loops of Section 5.4 with the hardware a refinery installs in place of the model's ideal instruments.

What Measures It

The flash-zone pressure is read by an absolute pressure transmitter, and the 0.03 min lag of Section 5.3.5 stands for that instrument and its connection. The two cut points are thermocouple readings corrected for pressure by Equation (5.8). The vent gas is the hardest of the model's readings to obtain on a working unit. One refiner's account states that comparing an ejector with its curve is not easy,

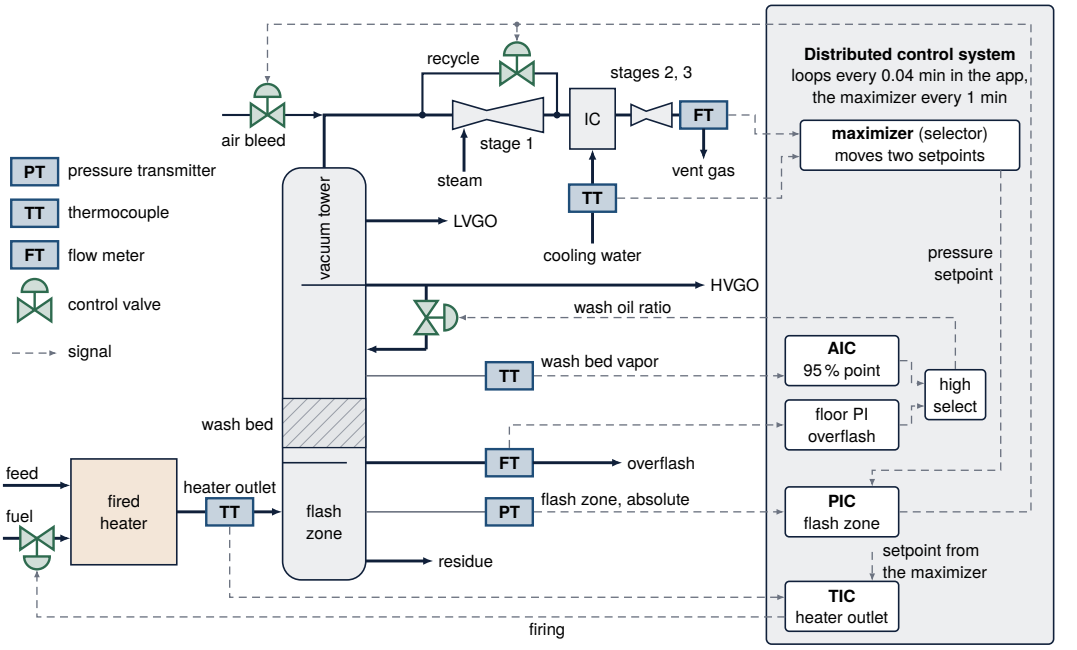


Figure 5.5 The pressure loop, the wash oil loop, the heater outlet loop and the maximizer with their hardware. Lettered boxes are sensing devices, valve symbols are final elements and dashed arrows are signals. IC is the first intercondenser. Notice that the pressure controller can only add load to the first stage, through the recycle or the air bleed, that two controllers reach the wash oil valve through one high select, and that the maximizer reads the vent gas and the cooling water but moves only setpoints. Not to scale, and not a piping or safety diagram.

because estimating the gas flow often poses difficulties, and that the vacuum level and the steam consumption are monitored instead [21].

> **TECHNOLOGY Sensor: Absolute Pressure Transmitter**
>
> The boiling temperature of a liquid is a function of the absolute pressure on it. Where boiling has to be controlled precisely the pressure is therefore measured absolute, and an open instrumentation textbook gives the vacuum distillation of crude oil as its example [22]. Such a transmitter is a differential pressure instrument with its low side sealed onto a vacuum chamber, so any pressure above a perfect vacuum reads as a positive difference. In a common design a taut metal diaphragm sits between two fixed plates, a fill fluid carries the pressure to it from an isolating diaphragm, and the two capacitances change as it flexes [22]. The weak point is the connection. A line that plugs makes the instrument respond too slowly or not at all, and a purge that flows too fast adds its own pressure drop to the reading [22].

What Moves It

The pressure controller moves a recycle valve around the first ejector stage, or an air bleed. Both only add load, so the ejectors set the lowest pressure, 28.4 mbar a in the model with both valves shut (Section 5.5.2). The design case holds 40.0 mbar a with the recycle 46 % open. The refiner's account above reports the opposite habit. It tries to keep its spill-back valve closed at all times and saves steam by taking parallel ejectors off line [21]. The firing reaches the coil through a firebox lag of 1 min in the model.

TECHNOLOGY Final Element: Steam Jet Ejector

An ejector has no moving part. A primary flow, here steam, accelerates and expands through a nozzle and entrains a secondary flow from a suction chamber. The flows mix, and a diffuser compresses the stream [23]. In its critical mode the entrained flow is choked and does not depend on the pressure downstream. As that back pressure rises a shock wave moves into the mixing chamber, and with a further rise the primary flow reverses into the suction chamber [23]. Refiners say that a stage breaks. One is expected to break when its motive steam is below the design pressure or its discharge is above the maximum on the maker's curve [21]. Vacuum units typically have three stages in series, each paired with a condenser that reduces the load on the next [21].

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level sends set points to distributed controllers that act through the actuators from the feedback of the sensors [24]. A survey of industrial practice puts basic control at every second and constraint control at every minute, and builds the conventional constraint layer from PID algorithms, lead-lag blocks and high and low select logic [4]. The loops of this chapter execute every 0.04 min and the maximizer moves its setpoints every minute, the period the survey names. The high select on the wash ratio is a block of that layer. A controller that a selector has de-selected tends to wind up, so override schemes need integral limits that engage then [22]. The pull of each branch toward the applied ratio in Section 5.4.1 is such a limit.

What Is Modernizing It

Three developments bear on this unit. The first is the online analyzer in place of the sample. The air that leaks into a vacuum system shows in its off gas, and one refiner reported in 2012 that it required an online oxygen analyzer on the tail gas of units that send it to a fuel gas header, and had installed several in five years [25]. The second is predictive control, which typically carries the valve positions of the loops below it as extra controlled variables held inside their limits [4]. A 2023 review that defends selectors expects significant benefits from it where many constraints switch [26]. The third has not arrived. Dry vacuum pumps reached the chemical industry of the United States in the late 1980s, and a 2005 energy guide did not expect them to make inroads in refining except in special applications [27].

IN PRACTICE The Overflash Is Sampled, Not Trusted to a Loop

In the model a controller holds a floor under the measured overflash. A 2016 answer in a refiners' question record describes the practice [28]. The overflash is metered in a gravity loop with an orifice plate. The answer advises against cascading the wash oil from the measured rate. A sample is analyzed instead, its share of gas oil is calculated, and the gas oil leaving the bottom of the bed per unit of bed area is compared with a target. The wash oil stays on flow control and its setpoint is changed by hand. Sampling is hard, because the stream is hot and below atmospheric pressure, and the analysis must tell spent wash oil from liquid entrained from the flash zone [28].

5.7 Limitations

The unit is generic, its sizes, rates and limits are chosen for teaching, so its yields and break loads are not predictions for any tower. The vapor pressure line overstates the shift per halving by about a fifth against Maxwell–Bonnell, the flash is one equilibrium stage, the sections are lagged cut points with one tail spread, and the pitch never boils. The ejector curve is declared rather than fitted to a vendor's, its break is a latched flag with a fixed jump, and the cyclic region of a real ejector appears only as case 3's limit cycle, whose count changes with the integration step. The wash bed's dry-out and coking are compressed for teaching, and its entrainment and capture are typical values rather

than a design method. Cracked gas follows the COT and flash zone temperature through a doubling law with no residence time.

The controllers carry no process model beyond two fixed numbers, the inferentials' design steam share and the vendor curve, so they meet model error only in the cases built to show it, as in case 7's stripping steam. The maximizer is a rule-based selector with no prediction. Disturbances arrive as steps and ramps.

OPEN CHALLENGE A Design Method for Selector Logic

The high select on the wash oil and the maximizer of this chapter are built from selectors. A 2023 review finds that such structures have proven themselves over a century with little academic attention, so that proper design methods have not been developed and the engineer works mainly from the pattern of earlier designs [26]. Its list of neglected research tasks includes the structure of selector logic and the tuning of anti-windup for selectors, and it calls the analysis of switched systems mathematically very difficult [26]. The wash oil select of Section 5.4.1, which turned measurement noise into a one-way climb until its tracking was changed, is a small case. Progress would be a method that predicts such behavior before commissioning.

5.8 Conclusion

This chapter studied a vacuum tower on 350 t/h of residue from the crude curve of the atmospheric tower, with an equilibrium flash, a wash bed, two gas oil draws and a three-stage ejector set, and gave every equation and default needed to rebuild it. Lowering the flash zone from 50 to 30 mbar a at a fixed heater outlet moves the HVGO 95 % point 24.7 K, and the cut shifts 26 to 41 K per halving on the app's line against 22 to 35 K with Maxwell–Bonnell. The ejectors hold the pressure nearly flat to a critical load of 964 kg/h, break with a 35 mbar jump near 695 kg/h of leak and recover only near 560 kg/h. Starving the wash bed raises the HVGO metals nine times while the laboratory reports half an hour late, and a hotter heater buys mostly overflash. A lift controller blind to the ejector load runs the heater to its ceiling, although the physical loop gain is only 0.014, and watching the load prevents it. The challenge fails as shipped with 1097 t and passes with the watched maximizer at 1154 t. The model is lumped and its constants are teaching values. A fuller model would add a staged wash and stripping section, a vendor ejector curve with its cyclic region, a heavy-oil vapor pressure correlation in place of the line, and a predictive controller for comparison. The simulator shows that the deepest cut on a vacuum unit is bought with a margin the cooling water can take away in an afternoon.

KEY POINTS

- Lowering the pressure in the tower lifts more gas oil for very little extra heat. No other handle moves the cut that far.
- The ejectors hold the pressure nearly flat up to a critical load, break with a jump, and recover only well below the load that broke them.
- The critical load falls as the cooling water warms, so the limit moves during the day.
- A hotter heater buys mostly overflash, and it adds cracked gas to the ejector load.
- A lift controller blind to the ejector load runs the heater to its ceiling. Watching the load prevents it.

Where else this applies. A limit that ends in a limit cycle under a protective override returns in Volume 4, Chapter 2, overrides that push throughput in Volume 4, Chapter 4, and an inferred quality that is right for one condition only in Volume 5, Chapter 10.

References

[1] U.S. Energy Information Administration, "Refinery capacity report 2026, table 1: Number and capacity of operable petroleum refineries by PAD district and state as of January 1, 2026," U.S. Department of Energy, Washington, DC, Tech. Rep., 2026.

[2] Occupational Safety and Health Administration, “OSHA technical manual, section IV, chapter 2: Petroleum refining process,” U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[3] D. Krishnamoorthy and S. Skogestad, “Systematic design of active constraint switching using selectors,” *Computers & Chemical Engineering*, vol. 143, p. 107106, 2020.

[4] S. J. Qin and T. A. Badgwell, “A survey of industrial model predictive control technology,” *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[5] J. H. Keenan, E. P. Neumann, and F. Lustwerk, “An investigation of ejector design by analysis and experiment,” *Journal of Applied Mechanics*, vol. 17, no. 3, pp. 299–309, 1950.

[6] R. B. Power, “Predicting unstable-mode performance of a steam jet ejector,” in *Fluid Power Systems and Technology: Collected Papers*. American Society of Mechanical Engineers, 1994, pp. 11–15.

[7] N. P. Lieberman, *Troubleshooting Vacuum Systems: Steam Turbine Surface Condensers and Refinery Vacuum Towers*. Beverly, MA and Hoboken, NJ: Scrivener Publishing and Wiley, 2012.

[8] H. K. Khafaji, A. Shahsavand, and S. H. Rajaee Shooshtari, “Simultaneous optimization of crude oil refinery vacuum distillation column and corresponding ejector system,” *Energy*, vol. 294, p. 130702, 2024.

[9] H. H. Rachford and J. D. Rice, “Procedure for use of electronic digital computers in calculating flash vaporization hydrocarbon equilibrium,” *Journal of Petroleum Technology*, vol. 4, no. 10, pp. 19–3, 1952.

[10] K. M. Watson and E. F. Nelson, “Improved methods for approximating critical and thermal properties of petroleum fractions,” *Industrial & Engineering Chemistry*, vol. 25, no. 8, pp. 880–887, 1933.

[11] M. R. Riazi, *Characterization and Properties of Petroleum Fractions*, 1st ed. West Conshohocken, PA: ASTM International, 2005.

[12] W. Wagner, J. R. Cooper, A. Dittmann, J. Kijima, H.-J. Kretzschmar, A. Kruse, R. Mareš, K. Oguchi, H. Sato, I. Stöcker, O. Šifner, Y. Takaishi, I. Tanishita, J. Trübenbach, and T. Willkommen, “The IAPWS industrial formulation 1997 for the thermodynamic properties of water and steam,” *Journal of Engineering for Gas Turbines and Power*, vol. 122, no. 1, pp. 150–184, 2000.

[13] J. B. Maxwell and L. S. Bonnell, “Derivation and precision of a new vapor pressure correlation for petroleum hydrocarbons,” *Industrial & Engineering Chemistry*, vol. 49, no. 7, pp. 1187–1196, 1957.

[14] J. Kewley, “Fractional distillation as applied in the petroleum industry,” in *Distillation Principles and Processes*, S. Young, Ed. London: Macmillan, 1922, pp. 319–360, chapter XXXIII, Distillation under reduced pressure, pp. 352–354.

[15] R. S. Danforth, “Effect of vacuum on lubricating-oil distillation,” *Journal of Fluids Engineering*, vol. 54, no. 9, pp. 39–43, 1932.

[16] R. B. Power, *Steam Jet Ejectors for the Process Industries*. New York: McGraw-Hill, 1994.

[17] M. Roza and M. Pilling, “Vacuum tower wash section design: Myth and reality,” in *Distillation Absorption 2010*, A. B. de Haan, H. Kooijman, and A. Górak, Eds., 2010, pp. 693–698.

[18] American Fuel & Petrochemical Manufacturers, “Question 65: Our vacuum column wash bed has lasted seven years in service and now needs to be replaced due to excessive coking and pressure drop,” AFPM Q&A Database, panel answer by KBC Advanced Technologies, 2012. [Online]. Available: https://afpm.org/data-reports/technical-papers/qa-search/question-65-our-vacuum-column-wash-bed-has-lasted-seven

[19] V. Georgiev, D. Stratiev, K. Kirilov, K. Petkov, and D. Minkov, “Reasons for low heavy vacuum gas oil yield in vacuum distillation of residual fuel oil,” *Chemistry and Technology of Fuels and Oils*, vol. 45, no. 3, pp. 164–169, 2009.

[20] A.-F. Chang, K. Pashikanti, and Y. A. Liu, “Vacuum distillation unit,” in *Refinery Engineering: Integrated Process Modeling and Optimization*. Wiley, 2012, pp. 117–144.

[21] American Fuel & Petrochemical Manufacturers, “Question 63: What is your refinery’s first stage vacuum ejector configuration? how do you evaluate your ejector system to determine if it is properly sized and performing optimally?” AFPM Q&A Database, panel answers by two refiners, 2011, afpm.org, question and answer search.

[22] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[23] G. Besagni, R. Mereu, and F. Inzoli, “Ejector refrigeration: A comprehensive review,” *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 373–407, 2016.

[24] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[25] American Fuel & Petrochemical Manufacturers, "Question 55: Leaks in a vacuum system can result in oxygen ingress into the process. what are the panelists' practices with respect to safe sampling and monitoring of vacuum system off gas quality?" AFPM Q&A Database, panel answers by three refiners, 2012, afpm.org, question and answer search.

[26] S. Skogestad, "Advanced control using decomposition and simple elements," *Annual Reviews in Control*, vol. 56, p. 100903, 2023.

[27] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[28] American Fuel & Petrochemical Manufacturers, "Question 55: What are your effective ways of measuring vacuum overflash flow in a gravity seal loop (not pumped)? please comment on overflash measurement for controlling wash oil flow," AFPM Q&A Database, panel answer by an engineering contractor, 2016, afpm.org, node 42219.

5.9 Exercises

5.1 Concept check · 2 min

Case 1 holds the heater outlet at 405 °C with the wash oil fixed by hand and lowers the flash zone from 50 to 30 mbar a. Which statement describes what the unit does in the next hour?

(a) The HVGO gains little, because the heater must fire several megawatts more to hold the outlet

(b) The 95 % point stays at its setpoint, because the HVGO loop raises the wash oil to hold it

(c) The cut moves about 10 K, the shift per halving the design brief for the app expected

(d) The HVGO gains about 25 t/h and its 95 % point passes its limit, for under 1 MW more firing

5.2 Concept check · 2 min

With the pressure loop on Man, raising the air leak from 120 to 600 kg/h lifts the flash zone only from 40.0 to 45.3 mbar a, yet at about 695 kg/h the first ejector stage breaks. Which explanation is right?

(a) Air is a small part of the first stage's load, but the second stage must pump it, so the intercondenser rises to the first stage's maximum discharge pressure

(b) The pressure loop opens the recycle valve to offset the leak until the valve reaches its stop and the pressure escapes

(c) The cracked gas falls as the leak rises and offsets the extra load, until the heater outlet loop can no longer hold the outlet

(d) The air raises the water vapor floor of the intercondenser until the floor itself reaches 100 mbar a

5.3 Concept check · 2 min

In case 6 the maximizer, with the ejector load ignored, drives the heater outlet to its 425 °C ceiling after the stage breaks. The measured physical loop gain from vent gas back to cracked gas is 0.0131 to 0.0140 on the broken branch. What runs away?

(a) The physical cracked-gas loop, whose gain rises above one once the stage has broken

(b) The loop closed through the maximizer, which keeps firing for lift after more heat has begun to give less lift

(c) The pressure loop, whose integral winds up while the recycle valve sits shut at its stop

(d) The HVGO loop, whose high select with the overflash floor rectifies noise into a climbing wash ratio

5.4 Theory and calculation · 6 min

The app's vapor pressure line is $\ln(p^{sat}/p_{atm}) = A(1 - T_b/T)$ with $A = 16$ and $p_{atm} = 1013.25\,\text{mbar}\,a$, so its atmospheric equivalent temperature is $\text{AET}(T, p) = T\,(1 + \ln(p_{atm}/p)/A)$ in kelvin.

(a) What normal boiling point does the line give to a component that boils at 391 °C under 40 mbar a? *Checkpoint:* 525.2 °C.

(b) At 391 °C, how much heavier a component boils when the pressure is halved, and when it falls from 50 to 30 mbar a? *Checkpoint:* 28.8 K and 21.2 K.

(c) What line constant would give 10 K per halving at 391 °C, and what constant matches the 24.0 K of Maxwell–Bonnell? *Checkpoint:* 46 and 19.2.

5.5 Theory and calculation · 8 min

Use Equation (5.3) and its constants: vapor space 2000 m^3 at 333 K with a molar mass of 19 g/mol, $C_1 = 0.15\,\mathrm{t/(h\,mbar)}$, $C_2 = 20\,\mathrm{kg/(h\,mbar)}$, MDP 100 mbar a, hysteresis 6 mbar. At design the water floor is $p_w = 51.8\,\mathrm{mbar}\,a$ and the non-condensable load is $G_{nc} = 384\,\mathrm{kg/h}$, of which 120 kg/h is air.

(a) Compute $\kappa = RT/(V_g\bar{M})$ in mbar per kg and the suction's time constant $60/(\kappa\,1000\,C_1)$ in minutes. *Checkpoint:* 0.729 mbar per kg and 0.55 min.

(b) Compute the intercondenser pressure and the critical load at design, and the load at which a broken stage recovers. *Checkpoint:* 71.0 mbar a, 964 kg/h and 844 kg/h.

(c) If only the air changed, at what leak would the stage break and at what leak would it recover? The app breaks near 694 to 696 kg/h and recovers near 560 kg/h. Explain both differences. *Checkpoint:* about 700 kg/h and 580 kg/h.

5.6 Theory and calculation · 6 min

The heater of Equation (5.7) has a film temperature $T_{film} = \mathrm{COT} + 45\,Q/26$ (°C, Q in MW), a coking rate $0.004 \times 2^{(T_{film}-450)/10}$ in mm/day, a skin $T_{skin} = T_{film} + 15\,\delta$ with $\delta = 3$ mm of coke, and a skin limit of 560 °C. The run left is the margin over the rate of rise.

(a) At design, 405 °C on 22.6 MW, compute the film temperature, the coking rate, the skin and the run left. *Checkpoint:* 444.1 °C, 0.002 66 mm/day, 489.1 °C, about 1776 days.

(b) Repeat for case 5 half an hour after the outlet is raised to 420 °C, when the heater fires 28.7 MW. *Checkpoint:* 469.7 °C, 0.0156 mm/day, 514.7 °C, about 193 days.

(c) By what factor did the coking rate rise, and why does the run left fall by more than that factor? *Checkpoint:* 5.9.

5.7 App challenge · 8 min

Load scenario 1, "Pressure moves the cut (wash oil by hand)", at the default seed and the default speed of 5. Press Run, and at about 5 min lower the Flash-zone pressure setpoint to 30 mbar a.

(a) Just before the change, read the HVGO and HVGO 95 % (truth) tiles. *Checkpoint:* 130.3 t/h and 545.8 °C.

(b) One hour after the change, read the HVGO, HVGO 95 % (truth) and Metals (truth) tiles. How far did the 95 % point move, and how does that compare with the line's 21.2 K for 50 to 30 mbar a? *Checkpoint:* about 155 t/h, 570.5 °C, 0.87 ppm; a move of 24.7 K.

(c) Export the CSV and find, in column `truth_t95_C`, how long after the change the 95 % point first exceeds 565 °C. *Checkpoint:* 12.8 min (±0.3 min).

5.8 App challenge · 9 min

Load scenario 2, "The ejector curve breaks" (the pressure loop is on Man), at the default seed and speed 5. Press Run and at about 2 min raise the Air leak into the tower to 600 kg/h.

(a) Four minutes later, read the flash-zone pressure on the PIC card and the Ejector load tile. *Checkpoint:* 45.3 mbar a (±0.2 mbar) and 90 % of critical.

(b) Raise the leak to 700 kg/h. How soon does EJECTOR BREAK appear, and what does the PIC card read a minute after it? *Checkpoint:* within a few seconds; about 70 mbar a (±1 mbar).

(c) About eight minutes later return the leak to 600 kg/h. Three minutes after that, is the stage still broken? Read the flash-zone pressure and the Ejector load tile. *Checkpoint:* still broken, 71.3 mbar a (±0.3 mbar), 92 %.

(d) Lower the leak by 10 kg/h every 3 min until the ejectors recover. At which leak does that happen, and why is it below the 580 kg/h that the hysteresis alone would give? *Checkpoint:* 550 to 560 kg/h.

5.9 App challenge · 8 min

Load scenario 9, "Challenge: a leak and a hot afternoon", at the default seed and speed 120. The run stops by itself at 8 h.

(a) Run it as shipped. Read the So far tile (hover over it for the full verdict). *Checkpoint:* 1097 t, broke, first miss an ejector break at 352 min.

(b) Reload the scenario, set the Architecture to Maximizer (Ejector load stays Watched) and the HVGO 95 % point setpoint to 562 °C before Run, and run 8 h. *Checkpoint:* 1154 t, no break, 95 % point at most 563.1 °C, passes.

(c) Reload, keep the Architecture on Loops, set only the Heater outlet setpoint to 397 °C before Run, and run 8 h. Why does this safe setting fail while the run of (b) passes? *Checkpoint:* 1024 t, no break, fails on the yield.

5.10 Application area · 4 min

In case 4 the wash oil is cut to 4 % at 5 min. The bed dries 9.5 min later and the true HVGO metals pass 1 ppm a minute after that, but the laboratory, which samples every 30 min and reports 20 min later, first shows the excursion at 50 min.

(a) How long after the true crossing does the laboratory report it, and what effective delay does Equation (1.12) give for this laboratory? What are the shortest and longest possible delays? *Checkpoint:* 34.5 min; 35 min; 20 and 50 min.

(b) The metals poison the catalyst of the cracker downstream. Which measured quantity warns the operator first, and which part of the control structure already acts on it? *Checkpoint:* the overflash; the floor override on the wash oil.

5.11 Application area · 4 min

On a hot afternoon the cooling water to the intercondensers warms. The condensing temperature is the water plus 4 K plus 1 K per t/h of steam (1.5 t/h here). Water's vapor pressure is 64.6 mbar a at 37.5 °C and 68.1 mbar a at 38.5 °C.

(a) Compute the critical load with the water at 32 and at 33 °C. In the challenge the vent gas is about 684 kg/h. At what water temperature does the stage break? *Checkpoint:* about 709 and 637 kg/h; between 32 and 33 °C.

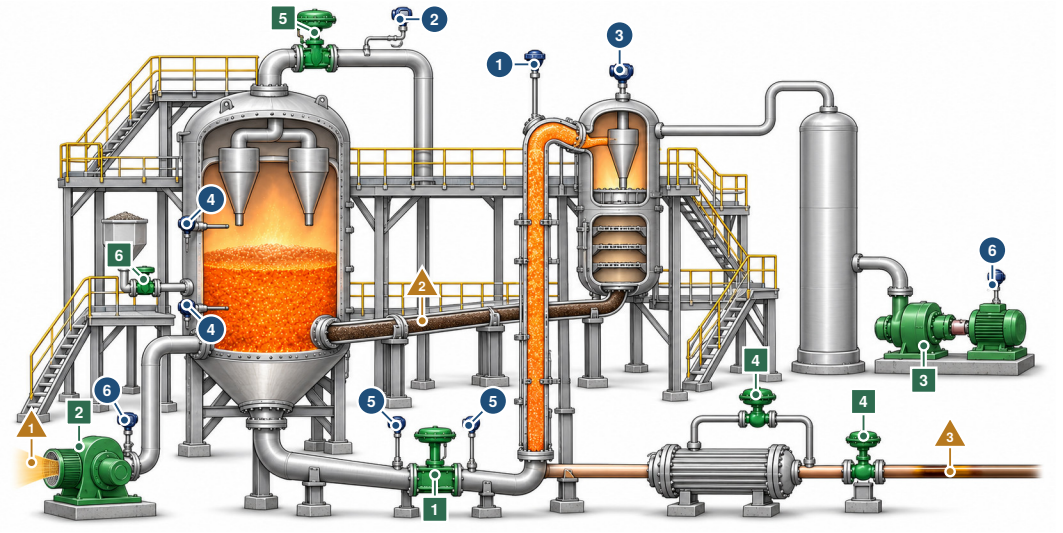
(b) Name two loads an operator can take off the ejectors before the break, and what each one costs. *Checkpoint:* lower COT (cracked gas), less stripping steam.

5.12 Review the agent · 5 min

After running scenario 3, "Pressure loop: recycle against air bleed", an AI assistant writes: "On the air bleed the pressure loop cycles, 22 breaks in the 20 min after the switch. The loop is tuned too tight for the ejectors. Halve the Pressure gain to 1.2 % per mbar and double the Pressure integral time to 0.6 min, and the bleed will hold 40 mbar a smoothly." Find the flaw, and say how you would check it in the app.

CHAPTER 6

Fluid Catalytic Cracking Simulator for Internal Heat Balance Feedback and Selective Control



The controller moves 1 catalyst slide valve 2 air blower 3 wet gas compressor 4 feed rate and feed preheat 5 flue gas valve 6 fresh catalyst addition

The controller reads 1 riser outlet temperature 2 oxygen in the flue gas 3 reactor pressure 4 bed and plenum temperatures 5 pressure drop across the slide valve 6 air flow and compressor speed

What disturbs it 1 hot afternoon air 2 catalyst activity 3 a heavier feed

The orange bed of the regenerator, where the coke burns off, is the only source of heat for the riser at the center, which the hot catalyst reaches through the slide valve at the bottom. Illustration, not to scale.

IN PLAIN TERMS **A Cracker That Is Its Own Furnace**

The main idea. A cat cracker breaks heavy oil into gasoline on a hot powder called catalyst. The powder picks up a carbon crust called coke, and burning the coke off is the only heat the plant has.

The challenge. Hotter cracking makes more coke and more coke makes more heat, so one small valve move about doubles in effect over ninety minutes. The plant also runs against the limits of its machines, and which one gives out first changes with the weather and the oil.

Why it matters. Running by one fixed number leaves fuel unmade on good days and overheats the plant on bad ones.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the History boxes in Sections 6.2 and 6.5, the Did you know? boxes in Sections 6.1 and 6.6, the Try it boxes in Sections 6.3 and 6.4, and Key points in Section 6.10. No equations.
Study	65 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 6.6 with Figure 6.5, Section 6.7, Section 6.8 and the Open challenge box in Section 6.9.

The control question A cat cracker has no furnace: its only heat is the coke it burns. How much does that internal heat balance amplify and slow every move, and how should one feed rate be shared among equipment limits that move with the weather, the feed and the severity?

Learning objectives

- Measure the heat balance loop gain from one slide valve step, and explain why the first two minutes understate the process gain.
- Explain from the coke balance why more air lowers the carbon on the catalyst but not the settled burn rate, and why afterburn peaks between partial and full burn.
- Compute the blower capacity, the catalyst circulation and the slide valve differential from the model, and the feed that a moving limit costs.
- Explain how a low selector with external reset shares the feed among four constraint controllers, and why the regenerator bed needs an override on the feed preheat instead.
- Compare the selector structure with a predictive controller on time above limits, riser outlet error, feed and margin.

Prerequisites Sections 1.4.4 (FCC heat balance and loop gain), 1.4.7 (selectors and external reset), 1.6.2 (PI loops), 1.6.3 (predictive control) and 1.6.4 (cases), and the throughput override of Chapter 3.

The app https://apmonitor.com/apps/fcc-unit/

See also Volume 4, Chapter 3 for a selector whose integrators track the applied command, Volume 4, Chapter 4 and Volume 4, Chapter 7 for a feed rate pushed against limits that change. Volume 6, Chapter 7 for amplification by a recycle.

6.1 Introduction

Opening one valve by three percent on the unit described here moves the riser outlet temperature 2.11 °C in the first two minutes and 7.17 °C in the end, and the end arrives ninety minutes later. Nothing else was touched. The regenerator dense bed, which no controller and no operator asked to move, settles 5.25 °C higher than it started. An engineer who reads the first two minutes of that trend and fits a model to it has measured a plant 3.4 times weaker and very much faster than the one underneath, and a controller tuned on that reading meets the real gain an hour later.

The cause is that the unit is its own furnace. Vaporized gas oil meets hot catalyst at the bottom of a riser and cracks for three seconds on the way up, the reaction is endothermic, and the catalyst is the only source of heat. The reaction leaves coke on the catalyst as it goes. The coked catalyst is stripped and dropped into a regenerator where the coke is burned with air, and that combustion is the only thing that reheats the catalyst. Extra catalyst therefore makes the riser hotter, a hotter riser cracks harder, harder cracking makes more coke, and the coke is the fuel. The fraction of any disturbance that the plant returns to itself is 0.52 on this model, so by Equation (1.8) the amplification is $1/(1-0.52) = 2.09$ and the settling time belongs to the regenerator rather than to the riser.

DID YOU KNOW? Most of the World's Gasoline

Fluid catalytic cracking produces the majority of the world's gasoline, and at the start of 2014 it ran in over 300 of the world's 646 refineries [1]. About 2300 tonnes of its catalyst are made every day [1]. The unit simulated in this chapter charges 250 t/h of gas oil.

The class of plant is larger than refining. Any unit that supplies its own reaction heat from a stream it produced closes the same path, and the family includes autothermal reactors, regenerative burners, feed and effluent exchangers around exothermic beds such as the one of Chapter 8, and every hot recycle, whose material counterpart is the snowball of Volume 6, Chapter 7. What such a plant shares is a loop gain below unity that inflates the steady-state gain, stretches the settling time and climbs with severity, and the same arithmetic produces multiple steady states once the gain approaches one, which has been documented for this unit specifically [2, 3]. A second general class sits on top

of the first. Severity and throughput both earn money on a cat cracker, the economic optimum lies outside the feasible region, and which of several machines runs out first changes with the weather, the feed and the catalyst. A low selector over one shared throughput handle is the classical answer [4] (Section 1.4.7), deciding which loop should own that handle inside each operating region has a documented design procedure [5], and raising a plant's rate until its slowest machine says stop is what a coordinating layer above the units is built to do [6].

However, a transfer function fitted to a step test carries neither behavior. The internal feedback is invisible because it is already inside the identified gain, and the identified gain is the one the test was long enough to see. The constraint set is invisible because a linear model has no limits in it, and the limit that matters today is not the one that mattered yesterday. This chapter works with a browser-based riser and regenerator described completely enough for a reader to rebuild it, and takes five measurements on it. They are the heat balance loop gain and the two very different process gains one loop shows at two minutes and at steady state, the gap between what more air does to the burn rate at a frozen carbon inventory and what it does once the unit settles, the afterburn peak that sits between partial burn and full burn so that the road passes through it in both directions, and what a low selector with external reset over four constraint controllers, with an override on the fifth constraint the shared handle cannot reach, is worth against a fixed throughput, and what a predictive controller on the same four handles costs against it. The measured results are that the heat balance amplification is 2.09 with the steady main loop gain 2.6 °C per percent of valve against 0.74 °C at two minutes, that the settled coke burn rate moves 0.3 % for a 9.3 % increase in air while the carbon on the regenerated catalyst falls by a third, and that over the 420 min graded shift the selective layer earns $95,391 with no time above the cyclone limit against $84,532 and 296 min above it for a fixed line-up. The plant has fifteen states, a closed-form coke burn, a closed-form slide valve and a damped Newton steady-state solver that has to be kept off a second, cold solution branch. The four-lump kinetics, the well-mixed regenerator and the selector arithmetic are all standard.

Section 6.2 traces how a cracking unit that burns its own coke came to be controlled the way it is. Section 6.3 builds the riser, stripper, regenerator, pressure and slide valve balances, puts the five limits on one scale and lists every constant. Section 6.4 measures what the heat balance does: the loop gain, the burn rate against the air, the afterburn peak and the one limit the feed rate cannot reach. Section 6.5 describes the three regulatory loops, the low selector with external reset, the override on the feed preheat and the predictive controller. Section 6.6 runs the fourteen scenarios through severity, constraint handover, the two failures that hold every setpoint, the graded challenge and the comparison with the predictive controller. Section 6.7 sets the instruments and final elements of a real unit beside the model. Sections 6.8 to 6.10 give the limitations, the economics of the constraint layer and a summary, and the exercises follow.

6.2 How the Cat Cracker Came To Be Controlled

Catalytic cracking began as a fixed-bed process in which the catalyst was regenerated in place, and it moved to a circulating solid in the 1940s, a step the History box below dates, precisely so that the coke could be burned off continuously in a second vessel while the first vessel kept cracking [1, 7]. The circulating solid also carried the combustion heat back, and that is where the control problem was created rather than solved. A design that recovers reaction heat from the product of the reaction has no independent heat input left, so every operating variable acts through the same energy balance and the reactor temperature became a quantity that is arrived at rather than set. Zeolite catalyst in the 1960s made the reaction fast enough that essentially all of the cracking happens in the transfer line, which shortened the reaction side to seconds and left the regenerator holding the whole of the plant's inertia [1, 7].

> ⧖ HISTORY **1942: The Catalyst Starts to Circulate**
>
> The first catalytic crackers held their catalyst still. The Houdry process, started up in 1936 in a New Jersey plant that processed 15 000 barrels a day, used a fixed bed reactor [1]. In 1938 a consortium of oil and engineering companies set out to develop a new cracking process. It built on a suggestion of Lewis and Gilliland, that a slow flow of gas through a powder might lift it enough to make it flow like a liquid. A pilot plant of 100 barrels a day at Baton Rouge, Louisiana, moved powdered catalyst through a reactor and a regenerator, and the first commercial unit started up in the middle of 1942. By the end of the war 34 units were running in the United States. The second Baton Rouge unit, started in June 1943, was still in operation when the review cited here appeared in 2015 [1].

The consequences of the internal energy balance were mapped much later than the technology. The observation that a cat cracker has more than one steady state at the same set of external conditions, a hot well-regenerated branch and a cold coked-up branch, was published as an operating fact with the underlying heat balance arithmetic behind it [2], and the full picture of where the branches lie, which are stable and how the boundary moves with feed quality was worked out in a four-part study of the current generation of units [3, 8]. That study also introduced the argument that a unit with far more measurements than manipulated variables should be controlled partially, holding a small number of dominant variables and letting the rest go where the balances put them [9]. The kinetic side developed in parallel and converged on lumping. Cracking selectivity was reduced to a small number of pseudo-components with second-order gas oil cracking and first-order overcracking of the gasoline lump [10], and the four-lump form that adds coke as a separate product and gives the gasoline yield a maximum against severity is the version a heat balance needs [11]. Combustion in the regenerator acquired its own literature once units began running lean, because carbon monoxide that leaves the dense bed finds oxygen in the freeboard where there is no catalyst to absorb the heat, and the freeboard was analyzed as an ignition problem in its own right [12].

Control practice moved in three steps. The first was regulatory, and it is still underneath everything else. Riser outlet temperature is held on the regenerated catalyst slide valve, flue gas oxygen on the air blower, and reactor pressure on the wet gas compressor, with a fourth loop holding the regenerator pressure on the flue gas slide valve. Which variables to pair, and which of the many available measurements are worth closing at all, was posed as a control structure selection problem and answered for the FCC specifically [9, 13]. The second step was a public dynamic model. A full simulator of a Model IV unit, with the catalyst circulation set by pressure balance rather than by a valve, was published so that control methods could be compared on the same plant [14], and it became the reference case for a generation of FCC control studies. The third step was multivariable constraint control. An industrial reactor and regenerator controller reported in 1993 pushed severity and throughput against the machine limits rather than tracking setpoints, and reported the payout in exactly those terms [15]. The regulatory structures that preceded it did not disappear, because a low selector and an override need one function block and no model, and they remain the layer that acts when the multivariable controller is off [4, 5]. The question of which fence a unit is allowed to stand against, and how much throughput the answer is worth, is the same question a coordinating layer asks [6].

6.3 Process Model

Regenerator R-101 and reactor R-102 exchange catalyst around one loop. 250 t/h of gas oil preheated to 220 °C meets 1700 t/h of catalyst at the bottom of the riser, a catalyst-to-oil ratio of 6.80, and leaves the riser at 525 °C with 75.00 % of the feed converted by weight into 46.00 % gasoline, 23.58 % light gas and LPG and 5.42 % coke. The catalyst delivers 95.0 MW to the riser, of which 72.7 MW heats and vaporizes the feed, 18.4 MW cracks it, 2.4 MW goes to the steam and 1.5 MW leaves the shell. The stripper misses a further 0.142 kg/s of hydrocarbon, so 3.905 kg/s of coke at 7.25 % hydrogen reaches the regenerator and releases 130.1 MW burning with 204 t/h of air, which holds the dense

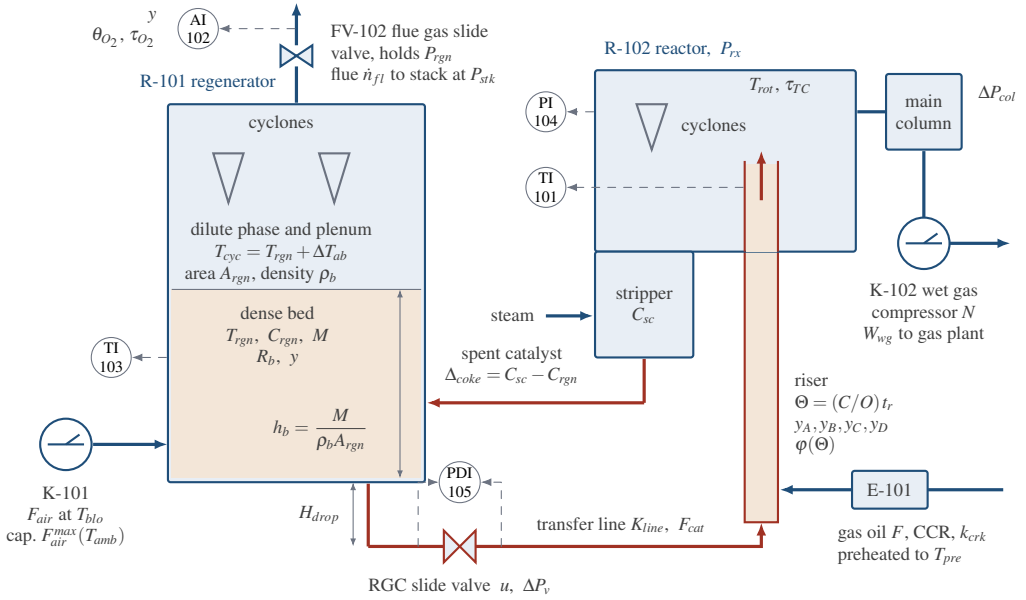


Figure 6.1 Regenerator R-101 and reactor R-102 with one loop of catalyst between them. Gas oil F preheated to T_{pre} in E-101 meets regenerated catalyst F_{cat} at the foot of the riser and cracks along catalyst contact Θ into the four lumps y_A to y_D at activity φ, leaving at T_{rot}. Coked catalyst at C_{sc} is stripped and drops into the dense bed, where air F_{air} from blower K-101 burns it at rate R_b leaving oxygen mole fraction y, so the bed holds M kilograms of catalyst at T_{rgn} with C_{rgn} carbon left on it. Carbon monoxide that escapes the bed burns above it and raises the plenum to T_{cyc}. Circulation returns down a standpipe of elevation H_{drop} through the regenerated catalyst slide valve at opening u and differential ΔP_v, and the transfer line takes K_{line} of differential as the square of the flow. FV-102 holds the regenerator pressure and K-102 holds the reactor pressure P_{rx}, against which ΔP_v is measured.

bed at 700 °C with 2.00 % oxygen by volume left in the flue gas. Carbon on the regenerated catalyst is 0.050 % by weight, delta coke is 0.827 %, the slide valve is 45.0 % open with 57.0 kPa across it, and the gross margin is $15,348 per hour. The vessels, the catalyst loop, the instrument tags and every symbol used below are drawn in Figure 6.1.

Table 6.1 lists everything the control system touches. The clock is in minutes, flows are in kg/s and displayed in t/h with the factor of 3.6 written out, duties are in kW, temperatures in °C, pressures in kPa absolute, inventories in kg and coke as a mass fraction of catalyst. Every derivative below therefore carries a factor of 60 where a per-second quantity enters a per-minute rate. Fifteen quantities are integrated. Ten of them are the plant and five are first-order actuator positions.

6.3.1 The Riser, Which Puts Itself Out

Cracking is marched along the riser against catalyst contact Θ, the product of the catalyst-to-oil ratio and the vapor residence time, because that product and not the length is what an operator changes. Four lumps are carried, gas oil y_A, gasoline y_B, light gas and LPG y_C, and coke y_D, with second-order

Table 6.1 Classification of Variables. Nominal Values Are Those of the Solved Design Point at Which the Basic Scenarios Open.

Symbol	Role	Nominal	Range
T_{rot} riser outlet	CV, measured	525.0	495–555 °C SP
y flue gas oxygen	CV, measured	2.00	0.6–5.0 vol % SP
P_{rx} reactor	CV, measured	310	250–390 kPa a SP
U nearest fence	CV, computed	26.1	60–100 % SP
U_{rgn} bed approach	CV, computed	−0.1	60–100 % SP
u RGC slide valve	MV	45.0	0–100 %
F_{air} air blower	MV	204	100–250 t/h
N compressor speed	MV	85.0	55–100 %
F feed rate	MV, selector	250	150–340 t/h
T_{pre} feed preheat	MV, override	220	175–300 °C
catalyst addition	MV, operator	2.0	0.5–8 kg/min
CCR carbon residue	DV, measured	0.60	0–1.8 wt %
k_{crk} crackability	DV, unmeasured	1.00	0.75–1.15
T_{amb} ambient	DV, measured	15	−5–42 °C
γ CO promoter	DV, unmeasured	1.00	0.02–1.50
stripper condition	DV, unmeasured	1.00	0.40–1.20
a_{er} standpipe aeration	DV, unmeasured	1.00	0.60–1.10
feed metals	DV, unmeasured	1.00	0.2–6.0
catalyst losses	DV, measured	2.0	0–30 kg/min
T_{rot} (°C)	state	525.0	300–700
T_{rgn} (°C)	state	700.0	400–1000
C_{rgn} (wt fraction)	state	0.000500	10^{-7}–0.06
T_{cyc} (°C)	state	703.6	400–1100
y duct oxygen (vol %)	state	2.00	0–21
P_{rx}, P_{rgn} (kPa a)	states	310, 350	130–600, 130–650
M inventory (kg)	state	165000	40000–320000
u_{fl} flue gas valve (%)	state	45.6	5–95
a catalyst activity	state	1.000	0.10–1.60
five actuator positions	states	as above	as above

cracking of the gas oil and first-order overcracking of the gasoline, as shown in Equation (6.1).

$$\begin{aligned}
\frac{dy_A}{d\Theta} &= -\varphi\,(a_1 + a_2 + a_3)\,y_A^2 \\
\frac{dy_B}{d\Theta} &= \varphi\,\left[a_1 y_A^2 - (a_4 + a_5)\,y_B\right] \\
\frac{dy_C}{d\Theta} &= \varphi\,\left(a_2 y_A^2 + a_4 y_B\right) \\
\frac{dy_D}{d\Theta} &= \varphi\,\left(a_3 y_A^2 + a_5 y_B\right)
\end{aligned} \tag{6.1}$$

where $a_i = k_i\,\kappa_i \exp\left[-(E_i/R)\left(T_{rot}^{-1} - T_{rot,0}^{-1}\right)\right]$ are the five rate coefficients evaluated at the riser outlet temperature in kelvin against the kinetic reference $T_{rot,0}$, κ_i is the feed crackability multiplier and applies to a_1 and a_2 only, and φ is the catalyst activity. All four lumps are mass fractions of the feed and sum to unity, so both sides of (6.1) are dimensionless per unit of contact.

The activity term is the part worth reading. Activity at any point in the riser is set by the coke that is already on the catalyst at that point, which is what came around from the regenerator plus what

this same reaction has laid down since the feed nozzles, as shown in Equation (6.2).

$$\varphi(\Theta) = a \exp\left[-\alpha \cdot 100 \left(C_{rgn} + \frac{y_D(\Theta)}{C/O}\right)\right] \qquad C/O = F_{cat}/F \tag{6.2}$$

where a is the equilibrium catalyst activity, α is the activity lost per weight percent of coke on catalyst and the factor of 100 converts the mass fraction to weight percent. The riser therefore extinguishes itself as it travels, with φ falling from 0.928 at the feed nozzles to 0.281 at the outlet at the design point, and carbon on the regenerated catalyst is charged twice, once directly in (6.2) and again through every kilogram of coke that the lost activity failed to prevent.

Contact is set by the circulation and by a vapor residence time that shortens as the vapor rate rises and lengthens with pressure, as shown in Equation (6.3).

$$\Theta = \frac{F_{cat}}{F} t_r, \qquad t_r = \operatorname{sat}\left(t_{r,0} \sqrt{\frac{F_0}{F} \cdot \frac{P_{rx}}{P_{rx,0}}},\ 1.0,\ 9.0\right) \tag{6.3}$$

where $t_{r,0}$ is the design residence time in seconds, F_0 is the design feed rate and the saturation bounds are in seconds. The exponent of one half rather than one is deliberate. Pure ideal gas scaling gives one, but the moles expand as the feed cracks and the catalyst slip velocity rises with gas density, and a full exponent puts a positive feedback path through the reactor pressure that walks the unit off its heat balance in manual operation.

Carbon residue in the feed goes to coke essentially quantitatively and is taken out of the unconverted bottoms rather than out of nothing, which is the standard additive coke treatment, as shown in Equation (6.4).

$$c_{add} = \min(0.9\, y_A,\ k_{CCR}\,\mathrm{CCR}) \qquad y_A \leftarrow y_A - c_{add}, \qquad y_D \leftarrow y_D + c_{add} \tag{6.4}$$

where CCR is the feed carbon residue in weight percent and k_{CCR} converts it to a coke mass fraction. Equation (6.1) is marched with the explicit midpoint rule in $n_r = 14$ equal steps of Θ, which was chosen because explicit Euler needed forty steps before the coke yield stopped moving in the third figure and the operating map evaluates roughly a thousand of these per second.

6.3.2 The Riser Energy Balance

The riser and reactor hold one lumped thermal capacity, and the catalyst is the only heat source in it, as shown in Equations (6.5) and (6.6).

$$Q_{dem} = F\left[c_{p,oil}(T_{rot} - T_{pre}) + \Delta H_{vap}\right] + F x \Delta H_{rx} + F_{st} c_{p,st}(T_{rot} - T_{st}) + Q_{loss,rx} \tag{6.5}$$

$$C_{rx} \frac{\mathrm{d}T_{rot}}{\mathrm{d}t} = 60\left[F_{cat} c_{p,cat}(T_{rgn} - T_{rot}) - Q_{dem}\right] \tag{6.6}$$

where $x = 1 - y_A$ is the conversion, $F_{st} = f_{st} F + F_{strip}$ is the dispersion plus stripping steam, $C_{rx} = M_{strip} c_{p,cat} + M_{shell,rx} c_{p,steel}$ is the combined catalyst and steel capacity in kJ/°C, and the factor of 60 turns kW into kJ/min. There is no other term in (6.6). Lowering the preheat grows the first bracket, and with the riser outlet held the circulation has to grow with it, which is the whole of the preheat handle.

6.3.3 The Stripper, and Why Throughput Reaches the Regenerator

Stripping steam is a fixed rate, so the steam-to-catalyst ratio falls as the circulation rises and entrained hydrocarbon rides into the regenerator on the catalyst, as shown in Equation (6.7).

$$C_{st} = c_{st,0} \exp\left[k_{st}\left(\frac{1}{r_{st}} - 1\right)\right], \qquad r_{st} = \frac{F_{strip}\,\eta_{st}/F_{cat}}{F_{strip}/(6.8 F_0)} \tag{6.7}$$

where η_{st} is the stripper condition multiplier and r_{st} is the steam-to-catalyst ratio as a fraction of its design value, which is four grams of steam per kilogram of catalyst. The carryover rate is $R_{st} = F_{cat}\min(C_{st}, 0.004)$ in kg/s. That stream is hydrogen rich, so per kilogram it releases 30 % more heat than coke and needs 12 % more oxygen, and both arrive exactly when the unit is being pushed. Without (6.7) the regenerator temperature would not move with throughput at all.

Catalytic coke and stripper carryover burn in the same bed and are blended into one pseudo-coke, as shown in Equation (6.8).

$$\begin{aligned} R_c &= F y_D + R_{st} \\ \Delta H_c^{mix} &= \frac{F y_D \Delta H_c + R_{st} H_{st}}{R_c} \\ w_C^{mix} &= \frac{F y_D w_C + R_{st} w_{C,st}}{R_c} \end{aligned} \tag{6.8}$$

where R_c is the total coke reaching the regenerator in kg/s and hydrogen on coke, which with delta coke is the pair a unit engineer reads the regenerator by [16], is $100(1 - w_C^{mix})$, or 7.25 % at the design point.

6.3.4 Coke Burn and Excess Oxygen, Solved Together

The dense bed is treated as well mixed, so the oxygen that the combustion sees is the oxygen in the flue gas. The burn rate is first order in carbon on catalyst and first order in oxygen mole fraction, and the oxygen balance closes on the same mole fraction, which makes the pair solvable exactly as shown in Equation (6.9).

$$\begin{aligned} R_b = \beta y, \qquad F_g y = \dot{n}_{O_2,in} - \nu R_b \implies y &= \frac{\dot{n}_{O_2,in}}{F_g + \nu\beta} \\ \beta = k_b M C_{rgn} \frac{P_{rgn}}{P_{rgn,0}} \exp\left[-\frac{E_b}{R}\left(T_{rgn}^{-1} - T_{rgn,0}^{-1} \right) \right] & \end{aligned} \tag{6.9}$$

where $F_g = F_{air}/M_{air}$ is the air in kmol/s, $\dot{n}_{O_2,in} = 0.2095 F_g$, $\nu = w_C^{mix}/12.011 + (1 - w_C^{mix})/(4 \times 1.008)$ is the oxygen demanded per kilogram of coke burned completely, and k_b is the combustion rate constant. Read at a frozen carbon inventory, (6.9) says that more air really is more burn, because y climbs toward 20.95 % and R_b approaches 0.2095β. Read at steady state it says the opposite, because the coke burned has to equal the coke the riser sends and the riser does not know how much air there is. That single line is the whole of the result in Section 6.4.2.

6.3.5 Afterburn Above the Bed

Below about one percent oxygen a growing fraction of the carbon leaves the dense bed as carbon monoxide, and above the bed there is no catalyst to absorb heat, so whatever monoxide meets whatever oxygen up there burns in the cyclones, as shown in Equation (6.10).

$$\begin{aligned} f_{CO} &= \frac{1}{1 + \gamma (y/y_{CO})^{n_{CO}}}, \qquad \iota = \frac{1}{1 + \mathrm{e}^{-(T_{rgn} - T_{ign})/w_{ign}}} \\ \dot{n}_{CO} &= f_{CO} \frac{R_b w_C^{mix}}{12.011}, \qquad \dot{n}_{ab} = \iota \min(\dot{n}_{CO},\, 2 y F_g) \\ \Delta T_{ab} &= \frac{\dot{n}_{ab} \Delta H_{CO}}{(F_{air} + R_b)\, c_{p,fl}} \end{aligned} \tag{6.10}$$

where γ is the carbon monoxide promoter multiplier, y_{CO} is the bed oxygen at which half the carbon still leaves as monoxide, n_{CO} sets the sharpness of the transition, ι is an ignition sigmoid of width w_{ign} about T_{ign}, and ΔH_{CO} is the heat of monoxide combustion in kJ/kmol. The min in (6.10) is what makes the afterburn a peak rather than a slope, because it needs both reactants and each runs out at one end of the sweep. Heat released in the bed is the full combustion heat less whatever walked out as monoxide, $Q_{dense} = R_b \Delta H_c^{mix} - \dot{n}_{CO} \Delta H_{CO}$, so a bed that afterburns is a bed that is cooling.

6.3.6 The Regenerator Energy and Carbon Balances

The dense bed carries the plant's inertia, and the air is a cooling load as well as an oxidant, as shown in Equations (6.11) and (6.12).

$$\begin{aligned} C^{th}_{rgn}\frac{\mathrm{d}T_{rgn}}{\mathrm{d}t} &= 60\,[Q_{dense} - F_{cat}c_{p,cat}(T_{rgn}-T_{rot}) \\ &\quad -F_{air}c_{p,air}(T_{rgn}-T_{blo}) \\ &\quad -R_b c_{p,fl}(T_{rgn}-T_{rot}) - Q_{loss,rgn}] \end{aligned} \tag{6.11}$$

$$\begin{aligned} \frac{\mathrm{d}M}{\mathrm{d}t} &= \dot{m}_{add} - \dot{m}_{loss} \\ \frac{\mathrm{d}C_{rgn}}{\mathrm{d}t} &= \frac{60(R_c - R_b) - C_{rgn}\,\mathrm{d}M/\mathrm{d}t}{M} \end{aligned} \tag{6.12}$$

where $C^{th}_{rgn} = Mc_{p,cat} + M_{shell,rgn}c_{p,steel}$ and $\dot{m}_{add}$ and $\dot{m}_{loss}$ are the fresh catalyst addition and the cyclone losses in kg/min. At the design point the 130.1 MW released in the bed splits into 95.0 MW carried to the riser, 31.8 MW spent heating air from the blower discharge to bed temperature, 0.8 MW into the flue gas and 2.5 MW through the shell. The air term in (6.11) is why raising throughput does not raise the bed temperature. More feed brings proportionally more coke and proportionally more air, and the second of those is a cooling load.

Two first-order states follow the bed, the cyclone and plenum temperature and the flue gas duct, as shown in Equation (6.13).

$$\begin{aligned} \tau_{cyc}\frac{\mathrm{d}T_{cyc}}{\mathrm{d}t} &= T_{rgn} + \Delta T_{ab} - T_{cyc} \\ \tau_{duct}\frac{\mathrm{d}y_{duct}}{\mathrm{d}t} &= 100y - y_{duct} \end{aligned} \tag{6.13}$$

where y_{duct} is the flue gas oxygen reported in volume percent and $\tau_{duct} = 0.25$ min is the duct holdup, which is a different quantity from the analyzer lag τ_{O_2} of Figure 6.1. The lagged T_{cyc} is the cyclone temperature that the application displays, logs and scores, and the afterburn fence constrains it through the cyclone thermocouple. Raising τ_{cyc} from 0.4 to 10 min in scenario 6 leaves the plenum at 704.2 °C one minute after the promoter loss at 240 min, against 729.0 °C at the default.

6.3.7 Pressures and the Wet Gas Compressor

Reactor pressure is one gas holdup between the light gas the reaction makes and the volume the compressor swallows, as shown in Equation (6.14).

$$\begin{aligned} \frac{\mathrm{d}P_{rx}}{\mathrm{d}t} &= \frac{60R(T_{sys}+273.15)}{V_{sys}M_{wg}}\,(W_{wg} - W_{out}) \\ W_{wg} &= Fy_C, \qquad W_{out} = \frac{N}{100}Q_{wgc}\,\frac{(P_{rx}-\Delta P_{col})M_{wg}}{R(T_{suc}+273.15)} \end{aligned} \tag{6.14}$$

where N is the compressor speed in percent, Q_{wgc} is the suction volumetric capacity at full speed, ΔP_{col} is the main column pressure drop to the compressor suction, and R is the gas constant. The machine is a fixed volumetric capacity, so once N reaches 100 % the reactor pressure is set by the reaction rather than by a controller.

Regenerator pressure has its own holdup and its own slide valve, held on the regenerator pressure itself rather than on a differential, as shown in Equation (6.15).

$$\begin{aligned} \frac{\mathrm{d}P_{rgn}}{\mathrm{d}t} &= \frac{60R(T_{rgn}+273.15)}{V_{rgn}} \\ &\quad \times \left(\dot{n}_{fl} - K_{fl}u_{fl}\sqrt{P_{rgn}-P_{stk}}\right) \\ \frac{\mathrm{d}u_{fl}}{\mathrm{d}t} &= K_{flv}\left(P_{rgn} - P^{SP}_{rgn}\right) \end{aligned} \tag{6.15}$$

where $\dot{n}_{fl} = F_g + R_b\left[-(1 - w_C^{mix})/(4 \times 1.008)\right] + \frac{1}{2}\dot{n}_{CO}$ is the flue gas molar rate after the change in moles that combustion produces, and u_{fl} is the flue gas slide valve position in percent. Holding P_{rgn} and not a differential is what lets the compressor limit reach the catalyst circulation, because a reactor pressure that rises against a fixed regenerator pressure squeezes the differential the catalyst slide valve has to work with.

6.3.8 The Slide Valve, Which Is Not a Flow Source

The regenerated catalyst slide valve is a restriction in a circuit, and the transfer line downstream of it takes differential pressure as the square of the flow. Writing $F_{cat} = K_s u\sqrt{\Delta P_v}$ with $\Delta P_v = \Delta P_{av} - K_{line}(F_{cat}/F_{cat,0})^2$ gives an implicit equation that solves in closed form, as shown in Equation (6.16).

$$\begin{aligned} F_{cat} &= \frac{K_s u\sqrt{\Delta P_{av}}}{\sqrt{1 + K_s^2 u^2 K_{line}/F_{cat,0}^2}} \\ \Delta P_{av} &= P_{rgn} - P_{rx} + \frac{580\, a_{er}\, g\,(H_{drop} + h_b)}{1000} \\ h_b &= \frac{M}{\rho_b A_{rgn}} \end{aligned} \tag{6.16}$$

where K_s is the valve coefficient, u is the opening in percent, a_{er} is the standpipe aeration multiplier, $F_{cat,0} = 6.8F_0$ is the design circulation, and the constant 580 is the aerated catalyst density in kg/m^3 at design aeration. The head term is in kPa. The closed form matters, because marching (6.16) by substitution oscillates and then diverges above about seventy percent open, which is exactly the region where the low-differential limit lives and therefore the only region a student examines closely. At the design point the standpipe and vessel differential supply 119.0 kPa, the transfer line takes 62.0 kPa of it, and 57.0 kPa is left across the valve against a 25 kPa minimum and a 14 kPa interlock.

6.3.9 Catalyst Activity, Actuators and Economics

Metals in the feed land on the catalyst and never leave, and the only answer is fresh catalyst, as shown in Equation (6.17).

$$\tau_a \frac{\mathrm{d}a}{\mathrm{d}t} = \mathrm{sat}\left(\frac{a_{max}}{1 + k_{me}\, m/\dot{m}_{add}},\ 0.1,\ 1.6\right) - a \tag{6.17}$$

where m is the feed metals multiplier, a_{max} is the fresh catalyst activity and τ_a is the turnover time, which is set to 90 min here against weeks on a real unit so that the effect can be watched in one session. Five actuators are first-order lags on commands the controller has already clamped, as shown in Equation (6.18).

$$\begin{aligned} \tau_s \frac{\mathrm{d}u}{\mathrm{d}t} &= u_{cmd} - u, & \tau_b \frac{\mathrm{d}F_{air}}{\mathrm{d}t} &= F_{air,cmd} - F_{air} \\ \tau_w \frac{\mathrm{d}N}{\mathrm{d}t} &= N_{cmd} - N, & \tau_f \frac{\mathrm{d}F}{\mathrm{d}t} &= F_{cmd} - F \\ \tau_p \frac{\mathrm{d}T_{pre}}{\mathrm{d}t} &= T_{pre,cmd} - T_{pre} \end{aligned} \tag{6.18}$$

No actuator can therefore walk past its stop, and the slide valve state is also clipped after each step.

Gross margin is products less feed, power and fresh catalyst, as shown in Equation (6.19).

$$\begin{aligned} J &= 3.6F\left(y_B p_{gaso} + y_C p_{light} + y_A p_{btms} - p_{feed}\right) \\ &\quad - \left(W_{blo} + W_{wgc}\right) p_{pow} - 60\,\dot{m}_{add}\, p_{cat} \\ W_{blo} &= \frac{1.005\,F_{air}\,(288.15)}{\eta_{blo}}\left[\left(\frac{P_{rgn}+30}{101.3}\right)^{0.286} - 1\right] \\ W_{wgc} &= \frac{2.0\,W_{wg}\,(T_{suc}+273.15)}{\eta_{wgc}} \\ &\quad \times \left[\left(\frac{P_{dis}}{P_{rx}-\Delta P_{col}}\right)^{0.200} - 1\right] \end{aligned} \tag{6.19}$$

where J is the gross margin in dollars per hour, p_{gaso}, p_{light}, p_{btms} and p_{feed} are in dollars per tonne, p_{pow} is in dollars per kilowatt hour and p_{cat} is in dollars per kilogram. Both machines are priced from an adiabatic head with a fixed efficiency, so 1.005 and 2.0 are the specific heats of air and of wet gas in kJ/(kg K), 288.15 and $T_{suc}+273.15$ are the two suction temperatures in kelvin, and the exponents 0.286 and 0.200 are $(\gamma-1)/\gamma$ for each gas. Each pressure ratio is floored at 1.05 and the compressor suction pressure at 20 kPa, so neither power turns negative when a pressure collapses. Both powers are in kW and the prices appear in Table 6.2.

6.3.10 The Five Fences on One Scale

Each limit is expressed as the fraction of its own approach span that has been used up, so a comparison across four different units means something, as shown in Equation (6.20).

$$U_i = 100\left(1 - \frac{\text{limit}_i - \text{PV}_i}{\text{span}_i}\right) \tag{6.20}$$

The five are the air blower against its capacity, the wet gas compressor against 100 % speed, the cyclone temperature against its metal limit, the slide valve differential against its minimum, and the regenerator dense bed against its metal limit. Blower capacity is not a constant, because a centrifugal machine takes in a volume rather than a mass, as shown in Equation (6.21).

$$F_{air}^{max} = F_{air,0}^{max}\,\frac{288.15}{273.15 + T_{amb}} \tag{6.21}$$

so the fence itself moves with the weather, from 230.4 t/h at 15 °C to 213.4 t/h at 38 °C. The spans were chosen so that each fence the feed rate can reach is worth a similar amount of feed. Measured as local slopes at the design point with the riser outlet and the oxygen held, one span is worth 48.5 t/h of feed on the blower, 49.0 on the compressor and 62.9 on the slide valve differential, which is what lets one gain and one integral time serve four constraint controllers. The cyclone fence barely moves with feed while the oxygen is held, and it is reached through the oxygen falling once the blower runs out. The regenerator span says something different, because thirty degrees is the whole authority of the feed preheat handle.

> **TRY IT IN THE APP** **11 · The fence moves**
>
> Load the case, set the speed to 20× and press Run. Nothing else has to be set. At 90 min the afternoon warms from 15 to 38 °C. Watch the blower capacity fall from 230.4 to 213.4 t/h and the feed rate follow it down from 284.5 to 261.5 t/h. In the evening the feed walks back up to 287.8 t/h.

6.3.11 Parameters, Initialization and Numerics

The 105 constants are listed in Table 6.2. Every one is shown and adjustable in the application's model panel except the 31 marked hidden, which are neither displayed nor adjustable and are the quantities

a student has to infer from a response. Two further constants are fixed in code and never exposed, the gas constant $R = 8.314\,\mathrm{kJ/(kmol\,K)}$ and the molecular weight of air $M_{air} = 28.85\,\mathrm{kg/kmol}$.

Each scenario opens at a solved steady state rather than at a guess. A manual scenario is solved with the operator's four handles given and both temperatures free. An automatic scenario is solved with the riser outlet temperature held, the oxygen on setpoint and the reactor pressure on setpoint, and the manual stations are then moved to match, so that manual operation is bumpless from the first second and no scenario spends its opening ten minutes walking to its own setpoint. The riser loop's integral is seeded net of the feedforward bias of (6.24), so a case that opens away from the design feed rate starts with the slide valve where the solution put it. A constraint scenario asks for more feed than exists on purpose, so the requested throughput is bisected downward until the solved point is inside every fence with room to spare, which is taken as a slide valve below 88 % open and a nearest fence below 88 % of its span.

Three balances are solved simultaneously in either case, the riser energy balance, the regenerator energy balance and the coke balance, with a damped Newton on a numerical Jacobian. Successive substitution around the same circle would contract only by the loop gain of about one half per pass, while the Newton finishes in six or seven iterations at every line-up used here. Two details of that solver are not decoration. The Jacobian column for the carbon on the regenerated catalyst is perturbed relatively rather than absolutely, because five ten-thousandths perturbed like a temperature falls below the noise floor of the riser march. Trust-region step limits and, in valve mode, an eight-step continuation from the design line-up keep the solve on the hot branch, because this plant also has a cold, coked-up solution that satisfies all three balances equally well [2, 3]. With the oxygen held at 2.0 % and the reactor pressure on setpoint, a Newton iteration started from a cold guess converges to riser outlet temperatures of 442 to 445 °C with 0.50 to 0.66 % carbon on the regenerated catalyst for slide valve positions from 42 to 55 %.

Integration is classical fourth-order Runge-Kutta at a fixed 0.02 min (1.2 s) step, the controller executes every 0.1 min (6 s), and the historian records ten samples per simulated minute, which a run exports as a CSV file of 129 columns. The engine, its seed (20260913 here) and its speed settings are those of Section 1.6.1. Speeds up to 20×, twenty simulated minutes per wall-clock second, are what the ninety-minute regenerator tail needs.

6.3.12 Assumptions

Model detail is concentrated on the heat balance and the constraint set. Each simplification below is stated with the conditions under which it fails.

Four lumps and one activity term. The scheme reproduces conversion, the gasoline maximum against severity and the coke yield, which is what the heat balance needs. It says nothing about gasoline octane, LPG olefinicity or cycle oil quality, none of which is a function of conversion alone, and real yield models carry ten to twenty lumps for exactly that reason [10, 11].

Carbon residue goes to coke quantitatively and comes out of the bottoms. That is the standard additive coke treatment. Real resid coke also poisons the catalyst and changes the metals loading, and neither is coupled here.

Catalyst deactivation is compressed by a factor of several hundred. The activity turnover in (6.17) is 90 min against weeks on a real unit. The direction and the trade are right and the clock is not, so no interval reported for the metals scenario predicts a real catalyst addition schedule.

One well-mixed dense phase at one temperature and one oxygen concentration. Real beds have a bubble phase that bypasses the catalyst, radial gradients of twenty degrees or more, and an air distributor whose condition is most of the difference between a unit that afterburns and one that does not [8].

Afterburn is one lumped dilute-phase reaction with an ignition sigmoid. The shape is right and the peak is in the right place, but (6.10) is a correlation rather than a freeboard model [12] and it cannot represent the local hot spots that actually destroy cyclones.

Symbol	Meaning	Value	Unit	Vis.
Feed and steam				
F_0	design feed rate	69.444	kg/s	H
$c_{p,oil}$	feed sensible heat capacity	2.45	kJ/kg °C	A
ΔH_{vap}	feed heat of vaporization	300	kJ/kg	A
ΔH_{rx}	heat of cracking per kg converted	354	kJ/kg	A
f_{st}	dispersion steam, fraction of feed	0.03	—	A
F_{strip}	stripping steam	2.0	kg/s	A
$c_{p,st}$	steam heat capacity	2.10	kJ/kg °C	H
T_{st}	steam supply temperature	250	°C	H
Riser kinetics				
$T_{rot,0}$	kinetic reference temperature	525	°C	H
$t_{r,0}$	vapor residence time at design	3.0	s	A
k_1	gas oil to gasoline rate constant	0.233302	—	A
k_2	gas oil to light gas rate constant	0.091886	—	A
k_3	gas oil to coke rate constant	0.018851	—	A
k_4	gasoline overcracking rate constant	0.01200	—	A
k_5	gasoline to coke rate constant	0.00260	—	A
E_1	activation energy, gasoline	68	kJ/mol	A
E_2	activation energy, light gas	89	kJ/mol	A
E_3	activation energy, coke	48	kJ/mol	A
E_4	activation energy, overcracking	105	kJ/mol	A
E_5	activation energy, gasoline to coke	48	kJ/mol	A
α	activity lost per wt % coke on catalyst	1.50	per wt %	A
k_{CCR}	additive coke from feed carbon residue	0.0090	per wt %	A
n_r	riser integration steps	14	—	H
Stripper				
$c_{st,0}$	hydrocarbon left on spent catalyst	0.00030	—	A
k_{st}	stripping sensitivity to steam ratio	1.20	—	A
H_{st}	heat of combustion, unstripped oil	43000	kJ/kg	A
$w_{C,st}$	carbon fraction, unstripped oil	0.860	—	A
Regenerator				
$T_{rgn,0}$	combustion reference temperature	700	°C	H
k_b	coke combustion rate constant	2.36699	—	A
E_b	coke combustion activation energy	150	kJ/mol	A
w_C	carbon fraction of catalytic coke	0.930	—	A
ΔH_c	heat of complete coke combustion	33014	kJ/kg	A
—	oxygen mole fraction in air	0.2095	—	H
ΔH_{CO}	heat of CO combustion	283000	kJ/kmol	H
y_{CO}	bed oxygen at half CO selectivity	0.0030	—	A
n_{CO}	sharpness of the CO transition	3.0	—	A
T_{ign}	dilute-phase CO ignition temperature	665	°C	A
w_{ign}	ignition transition width	18	°C	H
$c_{p,air}$	air heat capacity, blower to bed	1.10	kJ/kg °C	H
$c_{p,fl}$	flue gas heat capacity	1.15	kJ/kg °C	H
$c_{p,cat}$	catalyst heat capacity	1.15	kJ/kg °C	A
T_{blo}	blower discharge temperature	190	°C	A
$Q_{loss,rgn}$	regenerator shell heat loss	2500	kW	A
$Q_{loss,rx}$	reactor and riser heat loss	1500	kW	A
A_{rgn}	regenerator cross-section	78.5	m^2	A
ρ_b	dense bed density	480	kg/m^3	A
Thermal inertia				
M_{strip}	catalyst in reactor and stripper	25000	kg	A
$M_{shell,rx}$	reactor vessel steel	60000	kg	A
$M_{shell,rgn}$	regenerator steel and refractory	250000	kg	A

Symbol	Meaning	Value	Unit	Vis.
Pressure and wet gas				
$P_{rx,0}$	design reactor pressure	310	kPa a	H
$P_{rgn,0}$	design regenerator pressure	350	kPa a	H
ΔP_{col}	column drop to compressor suction	100	kPa	A
M_{wg}	wet gas molecular weight	30	kg/kmol	A
T_{suc}	compressor suction temperature	40	°C	A
T_{sys}	reactor and column vapor temperature	127	°C	H
V_{sys}	reactor, column and receiver volume	270	m^3	A
Q_{wgc}	compressor suction capacity at 100 %	7.9613	m^3/s	A
V_{rgn}	regenerator freeboard volume	400	m^3	H
P_{stk}	flue gas stack pressure	105	kPa a	H
K_{fl}	flue gas slide valve coefficient	0.002651475	—	H
P^{SP}_{rgn}	regenerator pressure setpoint	350	kPa a	A
K_{flv}	flue gas valve controller gain	0.15	%/min kPa	H
Catalyst loop				
H_{drop}	bed floor to slide valve elevation	9.5	m	A
K_{line}	transfer line loss at design flow	62	kPa	A
K_s	RGC slide valve coefficient	1.39032	—	A
ΔP_{min}	differential at loss of metering	25	kPa	A
ΔP_{trip}	low-differential interlock	14	kPa	A
M_0	design catalyst inventory	165000	kg	A
τ_a	catalyst activity turnover time	90	min	A
k_{me}	activity lost per metals ratio	0.60	—	A
a_{max}	fresh catalyst activity	1.30	—	H
Actuators				
τ_s	slide valve travel	0.12	min	A
τ_b	air blower response	0.35	min	A
r	blower rate limit	2.5	kg/s min	A
τ_w	compressor speed response	0.20	min	A
τ_f	charge pump and feed flow loop	0.60	min	A
τ_p	feed preheat train response	6.0	min	A
Instruments				
τ_{TC}	riser outlet thermocouple	0.25	min	A
τ_{O_2}	oxygen analyzer response	0.50	min	A
θ_{O_2}	oxygen sample line delay	0.60	min	A
τ_{Pm}	pressure transmitter	0.05	min	H
τ_{bed}	regenerator bed thermocouple	0.40	min	H
t_{opt}	feasible optimum search interval	5.0	min	H
Constraints				
$F^{max}_{air,0}$	blower capacity at 15 °C	64.0	kg/s	A
T^{max}_{rgn}	dense bed metallurgical limit	730	°C	A
T^{max}_{cyc}	cyclone and plenum limit	760	°C	A
—	regenerator high-temperature trip	750	°C	A
—	cyclone high-temperature trip	790	°C	A
s_{air}	blower approach span	10.0	kg/s	H
s_{wgc}	compressor approach span	14	%	H
s_{rgn}	dense bed approach span	30	°C	H
s_{cyc}	cyclone approach span	45	°C	H
$s_{\Delta P}$	slide valve approach span	32	kPa	H
Economics				
p_{gaso}	gasoline value	780	$/t	A
p_{light}	light gas and LPG value	520	$/t	A
p_{btms}	cycle oil and slurry value	430	$/t	A
p_{feed}	gas oil feed cost	520	$/t	A

Table 6.3 The Slide Valve Step of Scenario 1, 45.0 to 48.0 % in Manual With Air, Feed and Preheat Held. Percentages Are of the Settled Change.

Time after the step (min)	2	22	41	83
ΔT_{rot} (°C)	2.11	4.53	5.73	6.79
of settled 7.17 °C (%)	29	63	80	95
ΔT_{rgn} (°C)	−0.41	1.70	3.31	4.74
of settled 5.25 °C (%)	−8	32	63	90

The stripper is one exponential in the steam-to-catalyst ratio. That gives hydrogen on coke the right sensitivity to circulation, which is what makes the regenerator temperature move with throughput at all, but a real stripper is a tray or packing efficiency and it fails in ways (6.7) cannot show.

There is no standpipe fluidization model. The differential in (6.16) falls to the interlock and stops there. The flow reversal and the pressure oscillation that a real low-differential event produces are absent, and they are the reason such an event is dangerous rather than merely expensive.

There is no fractionator. Reactor pressure is one gas holdup swallowed by one compressor characteristic and the main column is a fixed pressure drop, so nothing here knows about gasoline endpoint, slurry circulation, or the wet gas molecular weight changing with severity.

No catalyst cooler, no flue gas expander, no CO boiler, no feed recycle and no torch oil. Each is a handle on a real unit, and the catalyst cooler in particular is what a resid unit installs to answer exactly the hot regenerator problem that the heavy feed scenario poses [16].

Perfect repeatability is a teaching convenience. A fixed step and a fixed scan produce one trajectory on every machine, which no distributed control system does.

6.4 What the Heat Balance Does

6.4.1 Measuring the Loop Gain

The experiment is one slide valve move in manual operation with the air rate, the feed rate and the preheat all held. Opening the regenerated catalyst valve from 45.0 to 48.0 % raises the circulation from 1700 to 1751.5 t/h. Solving the riser energy balance of (6.5) and (6.6) alone for that circulation, with the regenerator frozen where it was, gives a riser outlet temperature of 528.38 °C against 524.96 °C, a rise of 3.42 °C. What the unit actually settles at is 532.12 °C, a rise of 7.17 °C, and the dense bed rises 5.25 °C from 699.9 to 705.2 °C without anyone asking it to.

The ratio of those two numbers is the amplification, and the loop gain follows from it, as shown in Equation (6.22).

$$A = \frac{\Delta T_{rot}}{\Delta T_{rot}^{open}} = \frac{7.17}{3.42} = 2.094, \qquad L = 1 - \frac{1}{A} = 0.523 \tag{6.22}$$

where L is the fraction of any disturbance that the plant returns to itself through the coke it makes. A loop gain of 0.523 does two things at once. It leaves the plant stable, because the sum converges, and it makes the plant slow, because each pass around the loop has to wait for the regenerator. The recorded trajectory in Table 6.3 shows both. Push severity up and L climbs, from 0.534 at 525 °C to 0.550 at 535 °C for a one percent valve step at line-ups that hold 2.0 % oxygen, which is why a unit that has always been steady can turn twitchy the week somebody raises the riser outlet setpoint, and why the multiplicity that appears once the gain reaches unity is a property of the same arithmetic [2, 3].

Two features of Table 6.3 are worth separating. The dense bed initially falls, by 0.46 °C at three minutes, because the extra catalyst carries heat out of the regenerator before the extra coke arrives to pay for it, and it crosses back through zero at 8.8 min. The carbon on the regenerated catalyst is a third integrator inside the same loop and it overshoots, rising from 0.0494 to 0.0577 % by weight 13 min after the step before settling at 0.0545. Three coupled integrators with one of them slow is why nobody tunes an FCC main loop aggressively.

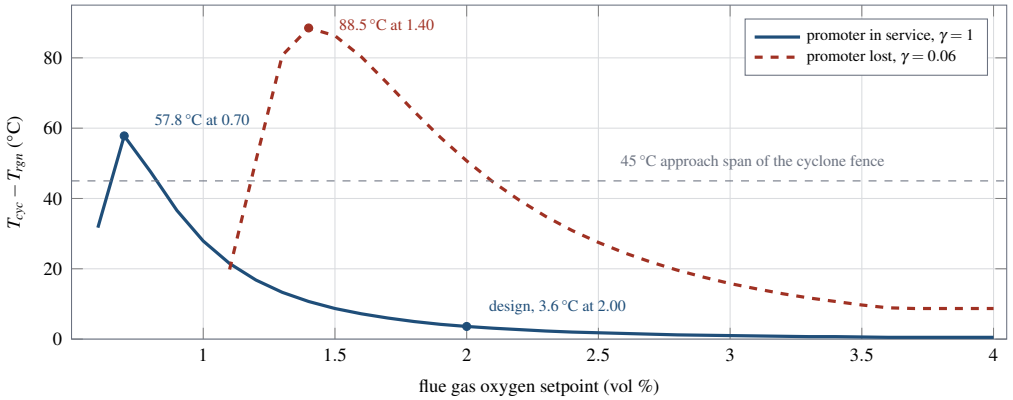


Figure 6.2 Afterburn against the excess oxygen setpoint at settled operating points, swept on the shipped model at 250 t/h and 525 °C. The peak is between partial burn and full burn rather than at either end, because the reaction needs carbon monoxide and oxygen together. The solid line is the promoter in service and the dashed line the promoter lost. Losing the promoter moves the peak right and up and puts 50.7 °C of afterburn at the design oxygen. Both curves flatten above about 3.6 % because the blower runs out of capacity and the oxygen stops at 3.71 and 3.62 %, and the promoter-lost curve is drawn only above 1.1 %, below which the steady-state solve leaves the hot branch.

Pause. Try Exercise 6.1 before reading on.

6.4.2 More Air Is Not More Burn

The two halves of this claim are not the same statement, and (6.9) contains both. At a frozen carbon inventory of 0.050 % by weight and a frozen bed temperature, raising the air from 56.7 to 66.3 kg/s, an increase of 16.9 %, raises the instantaneous burn rate by 15 %, because the rate is first order in oxygen mole fraction. At steady state the same move is worth almost nothing. Scenario 2 takes the air from 204 to 223 t/h, and 75 min later the burn rate has gone from 3.906 to 3.920 kg/s, a change of 0.3 %. It cannot do anything else, because the regenerator burns exactly the coke the riser sends it and the riser does not know how much air there is.

What the extra air buys is a lower carbon inventory. The excess oxygen rises from 2.03 to 3.59 % by volume and the carbon on the regenerated catalyst falls by a third, from 0.0494 to 0.0333 % by weight. What it costs is heat, because the air term in (6.11) is a cooling load, so the bed settles at 691 °C instead of 700 and the riser follows it down from 525 to 519 °C. Taking the air the other way, down to 162 t/h, takes the unit apart. By $t = 300$ min carbon has built to 0.46 % by weight on the regenerated catalyst, the activity in (6.2) has gone with it, conversion has fallen from 75.0 to 55.5 % and the margin from \$15,347 to \$7,019 per hour.

6.4.3 The Afterburn Peak

Sweeping the oxygen setpoint across settled operating points rather than frozen ones is the only honest way to draw the afterburn characteristic, because the carbon inventory moves with the oxygen and it is the pair together that decides how much monoxide leaves the bed. The result is in Figure 6.2.

The peak with the promoter working is 57.8 °C above the bed at 0.70 % oxygen by volume, against 3.6 °C at the design 2.00 % and 1.0 °C at 3.00 %. The road between partial burn and full burn therefore goes through the worst of it in both directions, which is the answer to the suggestion that an operator can simply add air. Losing the promoter moves the peak to 88.5 °C at 1.40 % and leaves 50.7 °C at the design oxygen, and the only remaining answer is a great deal more air. Scenario 6 asks for 4.0 % and does not get it, because the blower saturates at 230.4 t/h, the oxygen stops at 3.62 % and 8.7 °C of afterburn is what is left. The dense bed meanwhile gets cooler as the oxygen falls, from

700 to 686.6 °C at 0.85 % and 672.9 °C at the peak, because the heat in the escaping monoxide leaves the bed with it. An operator watching the bed thermocouple alone sees a unit that is behaving well, right up until the cyclones fail.

> **TRY IT IN THE APP** **6 · Afterburn, and why you cannot just change the air**
>
> Load the case and press Run. The oxygen setpoint is taken down to 0.85 % and the dense bed cools as the oxygen falls. Watch the gap between the plenum and the bed, not the bed alone. The promoter is lost at 240 min. It then asks for 4.0 % oxygen, the blower saturates at 230.4 t/h, the oxygen stops at 3.62 % and 8.7 °C of afterburn remains.

6.4.4 The Fence the Feed Rate Cannot Reach

Four of the five limits respond to throughput. The regenerator dense bed does not, and the arithmetic says why. Bed temperature follows delta coke, which is coke per kilogram of circulating catalyst, and at a held riser outlet temperature delta coke barely moves with throughput because more feed brings proportionally more catalyst and proportionally more air. Measured on this model, fifty tonnes an hour of extra feed at 525 °C moves the dense bed by 4 °C and moves it the wrong way, downward. The handle that does reach the bed is the feed preheat, whose whole travel from 175 to 300 °C is worth 30.8 °C of bed temperature, which is where the thirty degree approach span in (6.20) comes from.

Raising the preheat by sixty degrees at a held riser outlet takes the circulation from 1700 down to 1373 t/h, the catalyst-to-oil ratio from 6.80 to 5.49, and the conversion from 75.00 to 69.95 %, and it makes the regenerator hotter by 15.5 °C. Almost everybody predicts cooler, because less coke is burned in absolute terms. Less coke is correct, but it is burned onto a much smaller stream of catalyst, and the bed temperature follows the ratio and not the rate. Cutting the preheat is not free either. At 180 °C of preheat the unit makes 77.6 % conversion with the slide valve 61 % open and 39.6 kPa across it against a 25 kPa minimum, so defending one fence has spent two others.

Pause. Try Exercise 6.2 before reading on.

6.5 Controller Design

6.5.1 The Implemented PI Algorithm

All eight controllers are the velocity-form PID of Section 1.6.2, with the derivative taken on the measurement, and all eight ship with $\tau_D = 0$, so the derivative term is present in the implementation and inactive in every default tuning. One execution is Equation (6.23), which is Equations (1.16) to (1.19) with an output slew limit added.

$$
\begin{aligned}
e_k &= \sigma\,(\mathrm{SP}_k - \mathrm{PV}_k) \\
P_k &= K_c e_k, \qquad D_k = -K_c \tau_D \,\sigma \left.\frac{\mathrm{d}\mathrm{PV}^f}{\mathrm{d}t}\right|_k \\
u_k^{unsat} &= u_{bias} + \Sigma_k^{PD} + I_{k-1} \\
\Sigma_k^{PD} &= \Sigma_{k-1}^{PD} + \Delta P_k + \Delta D_k \\
u_k &= \mathrm{slew}\left(\mathrm{sat}(u_k^{unsat}, u_{min}, u_{max}),\, u_{k-1},\, r\right) \\
I_k &= I_{k-1} + \left[\frac{K_c}{\tau_I} e_k + \frac{u_k - u_k^{unsat}}{T_t}\right] \Delta t_c
\end{aligned}
\tag{6.23}
$$

where $\sigma = \pm 1$ is the action, PV^f is the measurement after an optional derivative filter that no loop here uses, Σ^{PD} accumulates the proportional and derivative increments without limit so that it telescopes to exactly $P + D$ for the gains in force, I_{k-1} is the integral accumulated through the previous scan, r is the output rate limit, which only the oxygen loop uses, and T_t is the back-calculation tracking time,

Table 6.4 The Implemented Loops as Coded, With Default Tunings. Action $\sigma = +1$ Is Reverse Acting and $\sigma = -1$ Is Direct Acting. The Air Ceiling Is the Blower Capacity of Equation (6.21).

CV	MV	K_c	τ_I (min)	σ	Output limits
T_{rot}	slide valve	0.55	3.5	+1	0–100 %
y	air blower	1.6	6.0	+1	28–F_{air}^{max} kg/s
P_{rx}	compressor	0.80	0.8	−1	55–100 %
U_{air}	feed rate	0.32	9.0	+1	41.7–94.4 kg/s
U_{wgc}	feed rate	0.32	9.0	+1	41.7–94.4 kg/s
U_{cyc}	feed rate	0.32	9.0	+1	41.7–94.4 kg/s
$U_{\Delta P}$	feed rate	0.32	9.0	+1	41.7–94.4 kg/s
U_{rgn}	feed preheat	1.00	10	+1	175–300 °C

Table 6.5 Measured Loop Responses in Scenario 4. The Open-Loop Column Is a 1 % Slide Valve Step at a Frozen Air Rate.

Time after the step (min)	1	2	20	settled
ΔT_{rot} (°C per %)	0.47	0.74	1.55	2.57
Closed-loop setpoint steps				
	t_{63} (min)		t_{95} (min)	
T_{rot}, 525 to 532 °C	4.8		13	
P_{rx}, 310 to 330 kPa	0.4		2.1	

which no loop sets and which therefore equals τ_I throughout. Anti-windup is back-calculation on the integral alone, and it runs only when there is reset action to unwind. While the output is clipped or rate limited, $(u_k - u_k^{unsat})/T_t$ drives I toward the value that reproduces the applied output, removing $\Delta t_c/\tau_I$ of the excess per scan, or 1/35 on the riser outlet loop. The clamp acts on the emitted output alone and never on Σ^{PD}, for the reasons given in Section 1.6.2. Manual operation solves the same expression for I, so transfer is bumpless in both directions. The controller period is $\Delta t_c = 0.1$ min throughout, five times the integration step.

6.5.2 Three Regulatory Loops on Three Different Plants

The regulatory layer is three loops, and the striking thing about them is that no two are the same kind of plant. Their pairings, directions and default tunings are in Table 6.4, and their measured responses are in Table 6.5.

The riser outlet loop on the slide valve has a steady gain, for a one percent step, of 2.6 °C per percent of valve at a frozen air rate and 2.4 °C with the oxygen loop closed, and only 0.74 °C of it, 29 %, has arrived at two minutes. A loop tuned on the first two minutes is tuned for a plant 3.5 times weaker and very much faster than the one underneath it, and the error is in the direction that produces an aggressive controller. Closed at the default tuning, a 7 °C setpoint step is 63 % complete in 4.8 min and 95 % complete in 13 min, with two thirds of a degree of overshoot.

The oxygen loop on the air blower is awkward for the opposite reason. Its steady gain is 0.31 % by volume per kg/s of air, and the immediate response is a thirteenth of that, because three kilograms a second delivers 0.07 of its eventual 0.92 % in the first minute, and the rest arrives as the carbon inventory in (6.12) burns down to its new level, 90 % of it by 7.7 min and 95 % by 12.5 min. On top of that the analyzer sits behind 0.60 min of sample line and 0.50 min of its own response, modeled as a transport delay followed by a first-order lag rather than as one lumped filter, because a delay inside a filter is not the same loop and does not detune the same way. The blower is also the one actuator with a rate limit, 2.5 kg/(s min), because its ceiling in (6.21) moves with the weather and an unrestrained controller takes a four kilogram per second jump in the ceiling all at once.

The reactor pressure loop on the compressor is the only ordinary one, at −2.9 kPa per percent of speed and settled in two minutes, and it is in service in every mode including manual. That is not a shortcut. The catalyst slide valve differential in (6.16) is measured against the reactor pressure, so every kilopascal the reactor gains comes off the differential the circulation is metered with. The 20 kPa setpoint step in Table 6.5 takes the compressor from 85.0 to 78.0 % and the slide valve differential from 57 to 38 kPa.

A feed rate feedforward completes the layer. A feed rate change is a load on the riser rather than a disturbance the temperature loop should have to discover, and the constraint layer above knows exactly how much feed it is about to add, as shown in Equation (6.24).

$$u_{bias} = K_{ff}\,(F_{cmd} - F_0) \tag{6.24}$$

where K_{ff} defaults to 1.50 % of valve per kg/s of feed, which is the travel that holds the catalyst-to-oil ratio where it was. It enters (6.23) as the controller bias, so the feedback loop only ever corrects what the feedforward got wrong and the manual tracker accounts for it automatically. Without it the evening feed increase in the graded challenge, which climbs at the 3.6 t/h per minute the selector allows and adds 16 t/h in the five minutes after the ambient falls at 310 min, pulls the riser outlet 4.4 °C off setpoint. With it, the same increase costs 1.3 °C.

6.5.3 The Low Selector and Its External Reset

Four constraint controllers share the feed rate through a low selector that also carries the operator's own throughput target, as shown in Equation (6.25) and drawn in Figure 6.3.

$$F_{SP} = \min(F_{target},\ u_{air},\ u_{wgc},\ u_{cyc},\ u_{\Delta P}) \tag{6.25}$$

Every one of the four holds its own fence at the same constraint target U_{SP}, which defaults to 97 % of the approach span, and every one of the four carries the same gain and the same integral time, which is possible only because (6.20) put them on a common scale. The same structure, a selector whose losing integrators track the command applied, guards the fuel valve of the turbofan in Volume 4, Chapter 3.

External reset is the detail that makes the structure work, and it costs one line. Each constraint controller is given the output the selector actually applied last scan, plus a slack of 0.10 kg/s, as its own upper output limit, as shown in Equation (6.26).

$$u_i^{max} = \mathrm{sat}(F_{sel} + \delta,\ F_{min},\ F_{max}), \qquad \delta = 0.10\,\mathrm{kg/s} \tag{6.26}$$

The back-calculation term in (6.23) then parks each loser's integral at that limit. With $T_t = \tau_I$ the fixed point of the integral recursion is exactly $I = u_i^{max}$ whatever the error, so a controller a hundred percent away from its own fence sits on the limit, 0.10 kg/s above the selected output, with its integral parked rather than wound to the ceiling and its unclamped demand above the limit by its proportional term alone. The instant its own error goes negative, its proportional term takes it below the others and it wins the actuator in one scan. The application verifies this on itself by running the push scenario for 150 min and requiring that at least two idle controllers sit strictly above the selected output and that the worst of them is within 2.0 kg/s of it. The same slack is also a feed ramp limit, because one slack per six-second scan is 3.6 t/h per minute, about as fast as a charge pump, a feed heater and a riser outlet temperature controller will follow together.

🕮 **Pause.** Try Exercise 6.3 before reading on.

6.5.4 The Override on the Second Actuator

The regenerator dense bed cannot be reached by the feed rate, so a fifth controller holds it with the feed preheat, low-selected against the operator's preheat target, as shown in Equation (6.27).

$$\begin{aligned} T_{pre}^{SP} &= \min\!\left(T_{pre}^{target},\ \mathrm{PID}_{rgn}(U_{SP}, U_{rgn})\right) \\ u_{rgn}^{max} &= T_{pre,k-1}^{SP} + \delta_p \end{aligned} \tag{6.27}$$

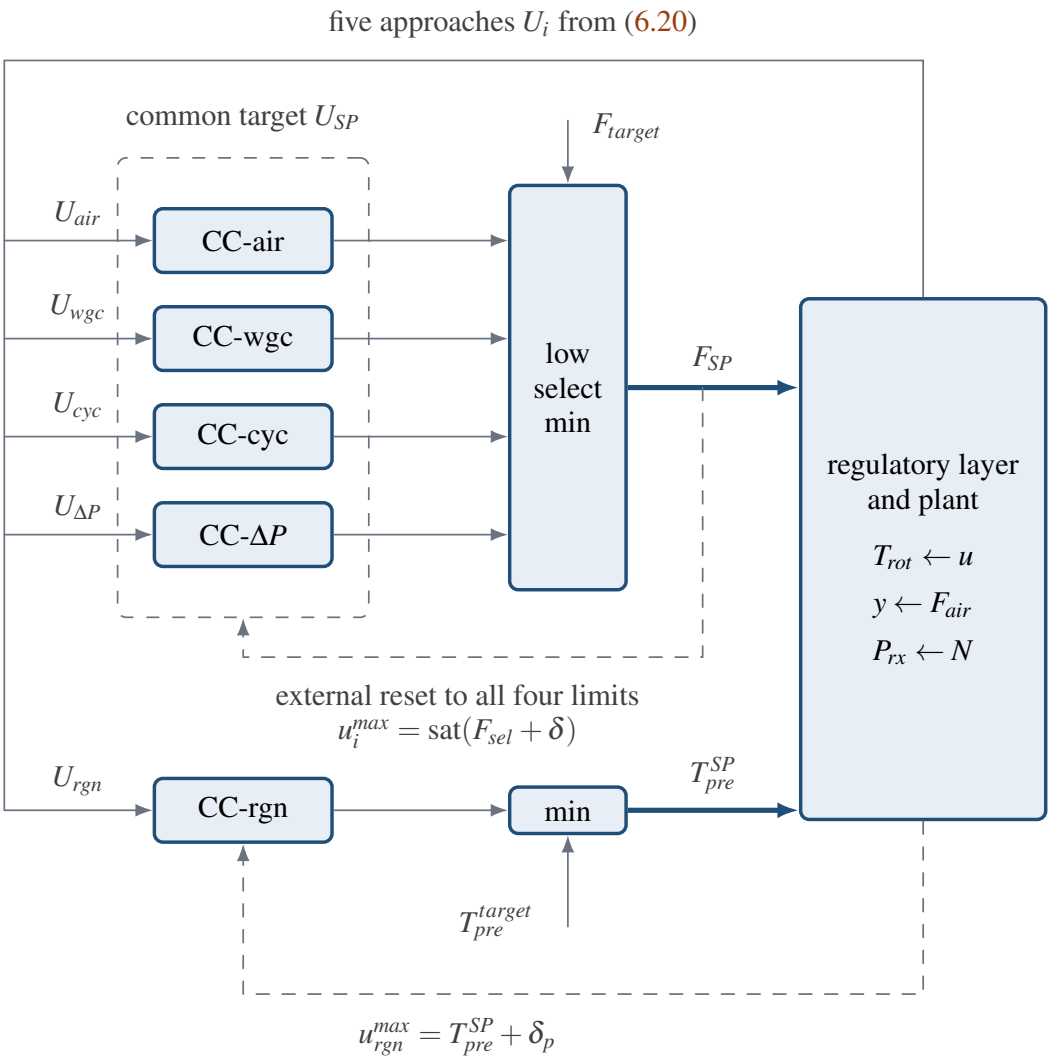


Figure 6.3 The constraint layer. Four controllers hold four fences at one common target U_{SP} and share the feed rate through a low selector that also carries the operator's throughput target, so the selector can only ever take feed away. Each loser is given the selected output plus a slack δ as its own upper output limit, which is external reset and is what parks its integral instead of letting it wind to the ceiling. The regenerator bed does not respond to feed rate, so its controller acts on the feed preheat through a second low select against the operator's preheat target.

where $T_{pre,k-1}^{SP}$ is the preheat setpoint applied on the previous scan, $\delta_p = 3.0\,^{\circ}\text{C}$ is the preheat slack, and the same external reset argument applies. The override is on a second actuator rather than on the shared one for a physical reason and not a structural preference, which is that cutting preheat sends more catalyst around for the same riser outlet temperature and spreads the same coke over more of it. Nothing else available moves delta coke. The cost is that the extra circulation walks straight into the slide valve differential and the blower, so defending the regenerator spends two of the other four fences.

Order inside one control execution never varies. The measurements are formed first, each through its own filter and, for the oxygen, its delay line, then the five fence approaches of (6.20) are evaluated,

then the tuning constants and the weather-dependent blower ceiling are applied, then the feasible optimum marker is refreshed if its own five-minute clock has expired, then the feedforward, then the three regulatory loops or their manual trackers, then the low selector of (6.25), and last the preheat override. The three regulatory loops therefore act on instrument signals, the riser outlet thermocouple, the delayed and lagged analyzer and the pressure transmitter. The constraint controllers read instruments too: the afterburn and bed approaches use the cyclone and bed thermocouples (0.40 min lags), the other three the metered air rate, compressor speed and slide valve differential, with no lag.

6.5.5 Model Predictive Control

A fourth Control option, the model predictive controller (MPC) of Section 1.6.3, replaces the selector and the override with one quadratic program on the four handles constraint control moves, u, F_{air}, F and T_{pre}, leaving the pressure loop in service and the riser outlet setpoint, and so the severity, to the operator [17, 18]. With only u and F_{air} free, a diagnostic copy holds the differential in scenario 10 only by leaving the riser outlet 24 °C under setpoint all run, and leaves scenario 7's bed above 730 °C for 353 of 400 min. Every $T_s = 0.5$ min it plans $N_p = 60$ steps, whose moves are held over eight blocks (1, 1, 2, 2, 4, 6, 10, 34), and applies the first move. Its model is the balances of Section 6.3 at the declared parameters and nominal disturbances, never the run's, linearized once at the design point with the pressure loop closed, with 21 states (the plant's less M and a, the five actuators, the pressure loop and six instrument lags). Its RMS residual over two hours is 1.1 to 16.4 % of the plant's response to slide valve, feed and preheat steps but 44 to 108 % on the riser outlet and bed for air steps, because the oxygen balance curves near 2 %, so ±2 kg/s of air both cool the riser outlet (1.92 and 0.73 °C, the model's ∓1.00). A steady-state Kalman filter at every scan estimates five input disturbances from the three thermocouples, the analyzer, the compressor speed and the applied commands, which makes the plan offset-free [19]. The plan minimizes Equation (6.28).

$$\mathscr{J} = \sum_{k=1}^{N_p} \left[q_T \left(\frac{e_{T,k}}{2} \right)^2 + q_O \left(\frac{e_{y,k}}{0.3} \right)^2 - c_F \frac{F_k}{\bar{F}} - c_P \frac{T_{pre,k}}{\bar{T}_{pre}} \right] + S \sum_{j=1}^{8} \sum_{m=1}^{4} \left(\frac{\Delta v_{m,j}}{s_m} \right)^2 + \mathscr{J}_s \tag{6.28}$$

where e_T and e_y are the predicted riser outlet and oxygen errors, $\bar{F}$ and $\bar{T}_{pre}$ the feed and preheat ranges, $\Delta v_{m,j}$ block j's move of handle m on scales s_m of 10 %, 2.5 kg/s, 20 t/h and 100 °C, $q_T = 1$, $q_O = 0.3$, $c_F = 5$, $c_P = 2$ and $S = 300$, and $\mathscr{J}_s$ prices each soft slack σ, in spans, at $1000(\sigma^2 + \sigma)$, the oxygen floor's at a tenth and quadratically. Hard limits are the shipped clamps and air slew, the selector's 3.6 t/h per minute feed rise, 5 % of valve and 5 °C of preheat per decision, and a valve ceiling set each decision from the measured ΔP_v by the declared law (6.16) so that ΔP_v stays above 25.96 kPa, its target at $U_{SP} = 97$ %. Soft limits are $U_i \le U_{SP}$ on the bed, cyclones, compressor and air, and $y \ge y^{SP} - 0.3$ %, a floor the MPC alone carries, all from the fourth step (2 min) so that an unavoidable early excess buys no sudden move, which hides nothing, since no case crosses a limit in its first 2 min. A dual active-set solver [20] takes 1.32 to 1.37 ms a plan in Chrome (load average 0.7 to 4.1), 0.27 % of a core at 1×. $S = 200$, between the tuning grid's 100 and 300, also meets its rule (IAE 113.4). Lowering S to 10 cuts scenario 13's riser outlet IAE from 127.4 to 55.6 °C min, but below $S = 20$ the challenge rings in a 3 min limit cycle. A rejected plan hands the handles to constraint control, whose seven controllers track the applied commands exactly with no margin, so each moves by one execution of its own law (checked to 10^{-9}, noise off).

Table 6.6 The Fourteen Scenarios and the Mechanism Each Is Built Around

#	Scenario	What it shows
1	The heat balance by hand	loop gain from one valve step
2	Air up a sixth, then down a fifth	the coke balance sets the burn rate
3	Preheat 220 to 280 to 180 °C	the only handle on the bed
4	The three loops, holding still	three plants, one unit
5	Severity to 552 °C	overcracking, then two fences in order
6	Oxygen down to 0.85 %, promoter lost	the afterburn peak
7	A heavier feed slate	the fence feed rate cannot reach
8	Metals to four times normal	perfect control hiding a failure
9	Standpipe aeration fails	the fences are not independent
10	Ask for 330 t/h	the selector and external reset
11	Ambient 15 to 38 to 12 °C	a constraint that moves
12	Severity, then aeration, under the selector	three binding fences, one structure
13	Overrides against a QP	the fences as one program
14	Graded challenge, seven hours	architecture before tuning

Table 6.7 Severity Sweep at 250 t/h, Scenario 5. Yields Are Weight Percent of Feed and Every Figure Is Read at the Settled Point. At the 552 °C Setpoint the Valve Is Fully Open and the Riser Outlet Settles at 551.8 °C.

T_{rot} (°C)	525	535	545	552
Conversion	75.00	78.17	81.01	82.72
Gasoline	46.00	46.27	45.88	45.20
Light gas and LPG	23.58	26.33	29.41	31.69

> ⧖ HISTORY **1978: Predictive Control Starts on a Cat Cracker**
>
> Model predictive control was founded in the late 1970s by two groups working independently, and both had a cat cracker in hand [21]. Richalet and coworkers introduced their method in 1978 with an explicit process model, a receding horizon and constraints on inputs and outputs. One of its three applications was a distillation column of a catalytic cracker, with two manipulated and three controlled variables and a sample time of 3 min. Cutler and Ramaker, at an oil company, used a piecewise linear model to predict the future behavior of a catalytic cracking unit, and reported the method in use there since 1974 [21]. The predictive controller of Section 6.5.5 plans every 0.5 min.

6.6 Results

Fourteen scenarios are arranged from the heat balance outward, listed in Table 6.6. Each fixes the opening mode, the control architecture, the ten tuning constants, the disturbance settings and any timed events, and is run, credited and exported as Section 1.6.4 describes.

6.6.1 Severity, and the Machine That Stops It

Scenario 5 walks the riser outlet setpoint up in three steps with the basic loops in service and the feed rate fixed at 250 t/h. The yields are in Table 6.7.

The gasoline yield peaks at about 535 °C and falls above it, because gasoline is an intermediate lump and a_4 in (6.1) is cracking it to light gas faster than a_1 is making it. Severity is therefore not a number to maximize, which is the practical content of the four-lump scheme [11]. What stops the sweep is the wet gas compressor, since every kilogram of the light gas leaves through one machine and the machine reaches 100 % speed at 540 °C on this feed rate. Past that the reactor pressure climbs because the compressor cannot take the gas away, and the rising reactor pressure eats the slide valve differential in (6.16). The 552 °C setpoint drives the valve fully open, the riser outlet stops at

Table 6.8 Constraint Handover Under the Low Selector at $U_{SP} = 97\,\%$ and a Throughput Target of 330 t/h. Every Figure Is the Settled Value After the Named Change.

	base 15 °C	hot 38 °C	evening 12 °C	severity 540 °C
Binding fence	air	air	air	ΔP
Feed rate (t/h)	284.5	261.5	287.8	245.1
Blower capacity (t/h)	230.4	213.4	232.8	230.4
Slide valve ΔP (kPa)	38.4	51.0	36.5	26.0
Conversion (%)	74.06	74.69	73.97	79.72
Gross margin ($/h)	17235	15986	17409	15746

551.8 °C, and the differential settles at 13.97 kPa, under the 14 kPa interlock, which the application flags at $t = 367$ min. Two fences arrive in order from one setpoint. On settled operating points at this throughput the gross margin peaks at $16,082 per hour at 542.4 °C, 2.5 °C past the 539.9 °C at which the compressor reaches full speed, and falls above it. The compressor limit therefore sits close to the profitable operating point, which is how a unit with its machines sized to that point behaves.

> **+ DID YOU KNOW? Cracking Harder Makes Less Gasoline**
>
> From the simulator (Table 6.7). Raising the riser outlet setpoint from 525 to 552 °C lifts conversion from 75.00 to 82.72 %, yet gasoline peaks at 46.27 % at 535 °C and falls to 45.20 %. The gain goes to light gas, which rises from 23.58 to 31.69 % and has to leave through the compressor.

6.6.2 Constraint Handover

Scenario 10 asks for 330 t/h, which the unit cannot make, and scenarios 11 and 12 then move the binding fence four times without touching a controller. Table 6.8 carries all four settled points.

The first column is the selector doing its job. The feed rate walks up and stops at 284.5 t/h with the air blower held at 97.0 % of its approach span, worth $17,235 per hour against the $15,348 of the design line-up, and the other three constraint controllers sit a fraction above the selected output with their integrals parked.

The second column is the argument for constraint control stated in one number. Warming the afternoon from 15 to 38 °C costs the blower 7.4 % of its mass capacity by (6.21), from 230.4 to 213.4 t/h, and the unit gives up 23 t/h of feed for it without anybody changing a setpoint. In the evening at 12 °C the capacity goes to 232.8 t/h and the feed rate walks back up to 287.8 t/h. Without the layer, somebody has to decide in advance how much feed the unit can take, and because the fence moves with the weather that number has to be safe on the worst afternoon of the year.

The fourth column is the fence changing hands three times. Fifteen degrees of extra severity puts the light gas yield from 23.06 to 27.84 % and takes the feed rate off the blower and onto the slide valve differential at 72 min while the reactor pressure climbs, then onto the wet gas compressor at 96 min, which holds the feed near 248 t/h once the pressures settle. From about 150 min the blower stands at only 47 % of its span, but its controller sits on the shared external reset limit, and on many scans the compressor controller's output reaches the same limit. The selector breaks such an exact tie toward the fence nearest its target, so the label names the compressor on every step from 150 to 200 min, and over the five minutes before the aeration event the feed stays between 247.6 and 248.5 t/h. Then the standpipe aeration deteriorates and the differential takes it for good at 245.1 t/h with 26.0 kPa across the valve. Three machines, three answers, one selector, and no logic changed. The structure does not know or care which constraint is active, which is why the answer to the question of what this unit is limited by is always today, at this severity, on this feed, in this weather. A drill that meets a different binding limit in each soil, in Volume 4, Chapter 7, makes the same point on one machine.

6.6.3 The Override, and Two Failures That Hold Every Setpoint

Scenario 7 switches to a heavier gas oil with the constraint layer in service, raising the carbon residue from 0.60 to 1.25 % by weight and dropping the crackability five percent. Additive coke by (6.4) raises delta coke from 0.827 to 0.99 %, the dense bed peaks at 733.3 °C against a 730 °C limit and spends 36 min above it with the cyclones reaching 742.9 °C, and the override cuts the feed preheat from 220 to 189.8 °C. The bed settles at 729.1 °C, which is exactly what the layer was asked for, the 730 °C limit less three percent of its thirty degree span. Conversion ends at 72.07 % and the margin at $13,008 per hour, and the feed rate never moves at all.

Scenarios 8 and 9 are the two failures a controlled variable cannot show. Putting four times the normal metals in the feed leaves the riser outlet temperature within 1.1 °C, the oxygen within 0.14 % and the reactor pressure within 0.12 kPa of setpoint for the first 220 min, while the slide valve opens from 45 to 72.3 %, the circulation rises from 1700 to 2023 t/h, the differential falls from 57 to 31 kPa, the bed cools from 700 to 670 °C and conversion falls three points to 72.05 % at a cost of $871 per hour. Nothing alarms. The evidence is in the actuator position and the yields, which is an argument for putting conversion, catalyst-to-oil ratio and valve position on the operator's screen as first-class numbers.

Scenario 9 is the mechanical version and it is the clearest demonstration that the fences are not independent. It runs on the basic loops at 285 t/h, and losing 26 % of the standpipe aeration takes the head out of (6.16), so the valve opens from 62.8 to 95.3 % to hold the temperature and the differential falls from 38.2 to 16.7 kPa, below the 25 kPa minimum. Opening the valve makes it worse, because the transfer line takes its share as the square of the flow. Somebody then raises the reactor pressure 25 kPa to give the compressor room, which is a sensible thing to do about a different problem, and the differential is measured against exactly that pressure. The valve pegs at 100 %, the differential settles at 11.5 kPa and stays under the 14 kPa interlock, the riser outlet collapses to 496.4 °C, conversion to 63.09 % and the margin to $12,387 per hour. Less severity cannot recover it, because the riser outlet setpoint at its 495 °C minimum still leaves 12.2 kPa. More preheat can, because 280 °C cuts the circulation from 1699 to 1550 t/h and restores 21.9 kPa at 69.2 % conversion.

6.6.4 What the Structure Is Worth

The graded challenge is seven hours with weak tuning, the feed rate fixed at 215 t/h, nobody watching any of the five limits, and five timed events, a warm afternoon at 40 min, a heavier slate at 110, the promoter falling away at 190, the standpipe deteriorating at 250 and the evening cooling at 310. The application judges the attempt at $t = 420$ min, on the case as given, and credits it when the margin is at least 8 % above the $84,532 the unit makes left alone, the riser outlet has stayed within ±2.5 °C of setpoint over the last third of the shift, the bed has spent under an hour above 730 °C and the cyclones under 5 min above 760 °C, and the slide valve differential has never fallen below 25 kPa. Table 6.9 compares leaving the unit alone against the suggested line-up over that 420 min shift.

The suggested line-up earns 12.8 % more money, 4.8 points clear of the threshold, and 88 more tonnes of gasoline, while the shipped line-up fails the rule on four counts: no margin above the reference, a worst riser outlet error of 2.52 °C, and 296 min each for the bed and the cyclones. The more important rows of Table 6.9 are the ones underneath. Leaving the unit alone puts the plenum 38 °C past the 790 °C trip at its worst and holds it above the 760 °C metal limit for 296 min of the 420 min shift, which on a real unit is a cyclone replacement. Every part of the line-up earns its place. Constraint control means the feed rate follows whichever fence is nearest instead of a number guessed in the morning. A throughput target well above what the unit can make means the selector is always the thing deciding, because a selector never adds feed and a target at or below the unit's capability leaves it nothing to take away. A constraint target of 93 % rather than 100 leaves something for the next disturbance, and there are four of them coming. 2.3 % oxygen rather than 2.0 buys margin against the afterburn of Figure 6.2 at the cost of a little air, which is the cheapest thing in the scenario against the most expensive.

Table 6.9 The Graded Challenge, Both Runs Read at $t = 420$ min. The Suggested Line-Up Is the Constraint Architecture at $U_{SP} = 93\,\%$ With a 330 t/h Target, 530 °C Severity, 2.3 % Oxygen and the Tuned Loops.

	left alone	suggested
Gross margin over the shift ($)	84532	95391
Average margin ($/h)	12076	13627
Gasoline made (t)	678	766
Feed processed (t)	1505	1692
Time above the 730 °C bed limit (min)	296	32
Time above the 760 °C cyclone limit (min)	296	0
Peak plenum temperature (°C)	828.5	757.2
Time below the 25 kPa differential (min)	0	0
Worst riser outlet error, last third (°C)	2.52	1.26
Riser outlet at the end (°C)	532.2	530.0

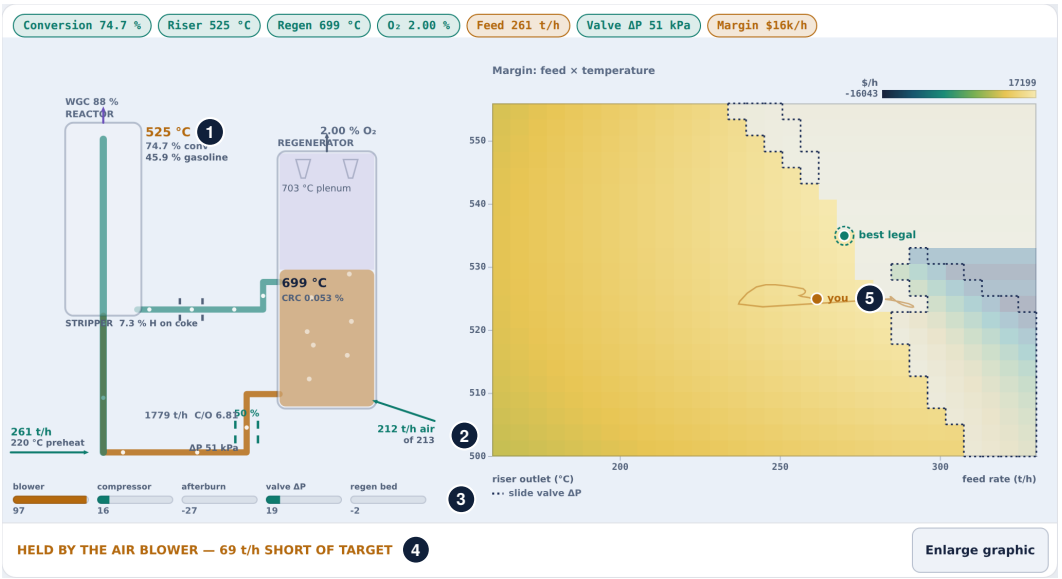


Figure 6.4 Constraint handover in scenario 11 at $t = 250$ min, 160 min after the afternoon warmed from 15 to 38 °C at $t = 90$ min and 10 min before the evening arrives, cropped to the plant view with numbered callouts. The riser outlet (1) reads 525.0 °C and the excess oxygen 2.00 %, both back on setpoint after a brief excursion when the capacity fell, while the blower capacity (2) has stepped down from 230.4 to 213.4 t/h and the feed rate has followed it from 284.5 to 261.5 t/h against an unchanged 330 t/h operator target. The air blower approach (3) sits on its 97 % target with the three other feed fences below it at 16, −27 and 19 % and the regenerator bed at −2, and the banner (4) names the blower and the 69 t/h it costs. The map puts the current point (5) below and left of the ringed best legal one, and the badge reads $16k per hour, the gross margin having fallen from $17,235 to $15,986 over the afternoon. In the full window, outside this crop, the trends cover the last three hours.

6.6.5 Comparison With the Predictive Controller

Table 6.10 runs seven cases under the shipped structure and under the MPC with only Control changed, with the same seeds, events, instruments and durations, no preview, the suggested line-up on the challenge, and 600 to 880 solves per MPC run, none rejected.

Table 6.10 Basic Loops (BL) or Constraint Control (CC) Against the MPC. Floor Is Time Below the MPC's Own Oxygen Floor.

Case (min)		IAE_T °C min	Bed min	ΔP min	IAE_O % min	Floor min	Feed t	Margin $
3 (300)	BL	317.6	0	0	58.95	51.8	1250	74988
	MPC	57.2	0	0	6.45	0.0	1227	74209
6 (440)	BL	272.5	0	0	101.00	128.6	1833	113371
	MPC	38.9	0	0	37.33	8.8	1813	112300
7 (400)	CC	97.5	35.6	0	22.29	17.2	1667	87464
	MPC	55.6	0	0	5.92	6.6	1665	87294
10 (300)	CC	18.4	0	0	3.72	0.0	1424	86198
	MPC	15.9	0	0	20.12	0.0	1430	86461
11 (400)	CC	219.5	0	0	45.49	40.2	1840	111647
	MPC	31.7	0	0	27.79	9.0	1837	111530
12, 13 (420)	CC	471.1	0	37.2	32.55	6.4	1782	113282
	MPC	127.4	0	0	15.79	0.0	1764	112271
14 (420)	CC	140.3	32.2	0	34.93	29.7	1692	95391
	MPC	127.9	0	0	49.41	2.4	1699	95754

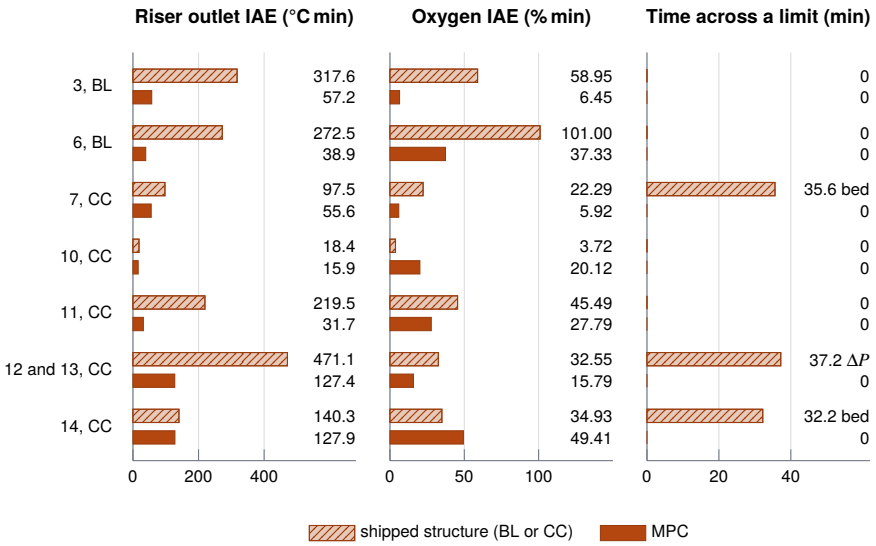


Figure 6.5 Three columns of Table 6.10 as bars, the shipped structure hatched and the MPC solid, for the seven cases in the table's order. The third panel draws the time above the bed limit or below the differential minimum, whichever the case has. Notice that the MPC shortens every bar of the first panel and has none in the third, while its oxygen bar is the longer one in cases 10 and 14. Feed and margin are not drawn.

The MPC lowers the riser outlet IAE in every case and never crosses the bed limit or the differential minimum. On scenario 7 the override cuts the preheat after the bed passes its target (peak 733.3 °C at 52.5 min) and the plan before (peak 729.1 °C at 214.7 min). Oxygen goes both ways. Below the MPC's own floor it spends 8.8 against 128.6 min on scenario 6, at an oxygen IAE of 37.33 against 101.00, but on that scorecard IAE it loses scenario 10 (20.12 against 3.72 % min) and the challenge (49.41 against 34.93), holding the oxygen under setpoint inside its floor to keep the air off its fence and the feed up. Figure 6.5 draws these three comparisons as bars.

It also earns less on scenarios 12 and 13, 1764 against 1782 t and $112,271 against $113,282, with 84 percent of the gap between the severity step and the aeration loss (581.6 against 596.4 t from 60 to 200 min, Figure 6.6). There the selector cuts the feed later, 263 against 248 t/h at 90 min, and its differential is under 25 kPa for 23.7 of its 37.2 min, while its compressor and blower reach their full approach span at 69.0 and 73.4 min without passing it, as the MPC's compressor does at 70.0 min.

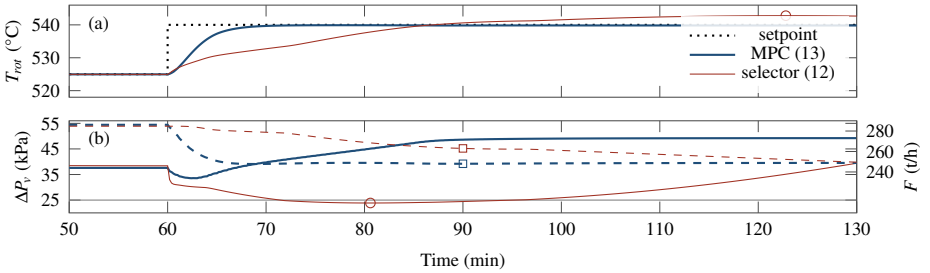


Figure 6.6 Scenarios 12 (selector, thin lines) and 13 (MPC, thick lines). (a) True riser outlet against the dotted setpoint, the selector's peak of 542.8 °C at 122.8 min circled. (b) Slide valve differential (solid, left) with its 25 kPa minimum and the selector's 23.9 kPa at 80.6 min circled, and feed rate (dashed, right), 263 and 248 t/h at the squares.

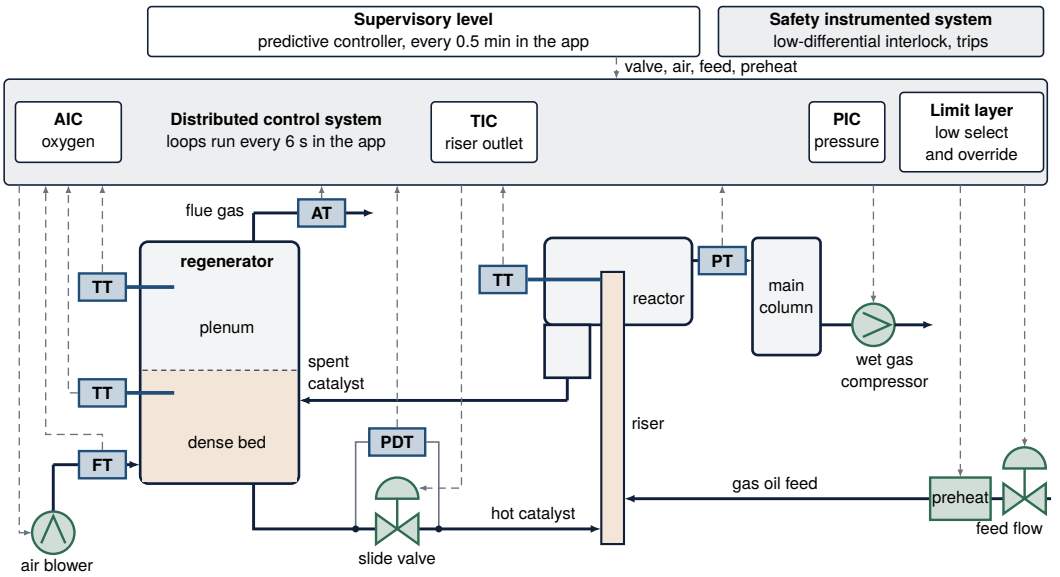


Figure 6.7 The loops of the cat cracker with their hardware. Lettered boxes are sensing devices, namely thermocouples with temperature transmitters (TT), the oxygen analyzer (AT) behind its sample line, the pressure transmitter (PT), the differential pressure transmitter (PDT) across the slide valve and the air flow transmitter (FT). Valve and machine symbols are final elements. Dashed arrows are signals. Notice that the limit layer reads the air flow, the two regenerator thermocouples, the slide valve differential and the compressor speed, and writes only the feed rate and the preheat, so one feed rate answers to whichever machine is nearest its limit. The low-differential interlock sits outside the control system. Not to scale, and not a piping or safety diagram.

On scenarios 3 and 6 it gives up 23 and 20 t by cutting a feed that basic loops never move.

6.7 Technology in Practice

Figure 6.7 draws the loops of Section 6.5 with the hardware a refinery installs.

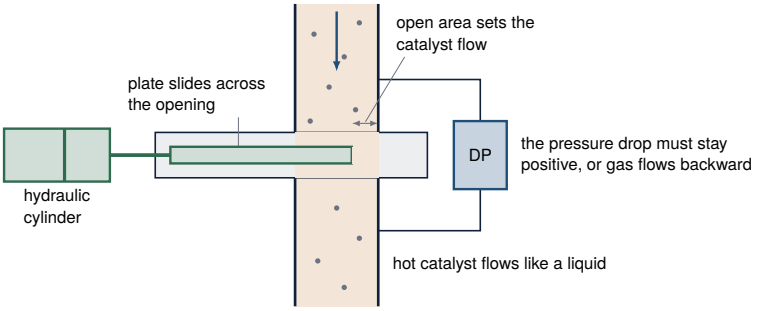


Figure 6.8 Principle of a slide valve in a catalyst standpipe. Notice that the plate only narrows the opening. The flow through it is driven by the pressure difference that the transmitter reads, which is why Equation (6.16) treats the valve as a restriction and not as a flow source.

What Measures It

The riser outlet, the dense bed and the plenum are read by thermocouples. The flue gas oxygen is read by an analyzer. One kind is a zirconium oxide cell, a solid electrolyte between platinum electrodes with flue gas on one side and reference air on the other, whose voltage is converted to an oxygen concentration [22]. An analyzer away from the process needs a sample system, and analyzer problems quite often start there. A plugged line leaves the analyzer reporting stale sample with no way of knowing it [22]. The delay θ_{O_2} and the lag τ_{O_2} of Table 6.2 stand for such a line and its analyzer. Refineries read catalyst levels and densities by differential pressure, because fluidized catalyst exerts a hydraulic pressure like a liquid, and catalyst that is not fluidized exerts none and behaves like fine sand [23]. That is the physical content of the aeration multiplier a_{er} in Equation (6.16).

> **TECHNOLOGY** **Sensor: Thermocouple in a Thermowell**
>
> Two wires of different metals joined at one end produce a voltage roughly proportional to temperature, and that junction is a thermocouple [22]. Thermocouples are less accurate than resistance thermometers but more rugged and wider in range. The Chromel and Alumel type is listed for about 32 to 2300 °F, an upper end near 1260 °C. The sensor sits in a thermowell, a pressure-tight sheath that lets it be replaced under pressure. The mass of the well adds a first-order lag, and an air gap at its tip adds a second [22]. The lags of 0.25 min at the riser outlet and 0.40 min in the bed stand for the well. An open circuit, the most common failure, can read as a wildly varying temperature, so the transmitter is set to drive its reading high or low, whichever is safer [22].

What Moves It

The regenerated catalyst slide valve controls the flow of hot catalyst to the riser [24]. The model carries it as the restriction of Equation (6.16) behind the travel lag τ_s, and Figure 6.8 shows its principle. The other handles are the air blower, the wet gas compressor, the charge pump and the preheat train, each a first-order lag in Equation (6.18).

> **TECHNOLOGY** **Final Element: Catalyst Slide Valve**
>
> Slide valves are a feature of most cat cracker designs. They are gate-type valves that open and close by sliding a disc across an orifice [23]. Their parts must move freely at high temperature, so the design keeps clearances, typically 0.04 to 0.06 inches wide, and a slide valve leaks a little even when new and fully closed [23]. The service is severely erosive, because the flowing catalyst is abrasive and wears the internals [24]. The differential pressure across the valve should always be positive. If it falls too low to sustain circulation, catalyst and gas can reverse and mix air with hydrocarbon [23]. At the refinery of that report the control system began to close both slide valves at 2 psi, which is close to the 14 kPa interlock of this chapter.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisor sends set points to distributed controllers that act through actuators from sensor feedback, with cycle times from milliseconds to minutes [25]. The app executes its controllers every 6 s and its predictive controller every 0.5 min, inside that range. The three regulatory loops, the four constraint controllers, the low selector and the preheat override would sit in the distributed controllers as function blocks. A controller with integral action that a selector has de-selected tends to wind up. One remedy feeds the selected output back to the integral term of each block, which pneumatic practice called external reset [22]. That is Equation (6.26). The low-differential interlock and the temperature trips belong to a safety instrumented system of sensors, logic solvers and final elements, designed to IEC 61511 and often independent of the control system [25].

> **IN PRACTICE** **A Slide Valve Is Not a Block Valve**
>
> What the model's slide valve never does is wear. Two federal investigations of cat cracker explosions, in 2015 and 2018, each found a spent catalyst slide valve eroded after five to six years of service, closed and still passing gas [23, 24]. In the first case the valve had been tested by closing it part way, which proved the mechanism and not the seal, and its position said nothing about the catalyst level above it. The stated lesson was to measure a safeguard directly, here by differential pressure, with a minimum that shuts the unit down [24]. The second report asked that API RP 553, the recommended practice for refinery control valves, cover the body of the slide valve [23].

What Is Modernizing It

Three developments bear on this chapter. The first is predictive control. Its early applications in the petrochemical industry had sampling times of several hours, and today these have largely come down to minutes and seconds. It often runs as a supervisory layer above the PID loops [21], where the app's predictive controller writes its four handles directly. The second is the safety system. Historically it stood alone, and some modern ones are designed to communicate with the control system [25]. The third is guidance. After the explosions of the In practice box the investigators asked in 2022 for a public technical publication on the safe operation of cat crackers, with recommended monitoring during startup, shutdown and standby [23], modes this chapter's simulator does not run.

6.8 Limitations

The application is not a yield model and not a regenerator design tool. With four lumps, one well-mixed bed and a correlated afterburn it cannot support a claim about a real unit's octane, its cyclone temperatures or how much feed it could take tomorrow. The kinetic constants were fitted so that this unit sits at the design point in Section 6.3 and reproduces a gasoline maximum, not against any real unit's yield data. Delta coke and hydrogen on coke behave correctly in direction and roughly in magnitude, which is what the heat balance needs, and nothing beyond that should be read from them. The steady-state solver finds the hot branch by continuation, and it declines to converge on the cold branch below about 1.1 % oxygen with the promoter lost, which is why the dashed curve in Figure 6.2 stops there rather than continuing to zero. That non-convergence is honest about the model and unhelpful as a picture of the plant, since a real unit does have an operating regime down there.

Two omissions matter more than the rest. There is no standpipe fluidization model, so a low-differential event ends at the interlock rather than in the flow reversal that makes it dangerous, and there is no catalyst cooler, which is the piece of equipment a real resid unit installs to answer precisely the problem scenario 7 poses. The individual assumptions and where each one breaks are set out in Section 6.3.12. One linearization at 250 t/h leaves the MPC without the afterburn's turn below 1 % oxygen, the oxygen balance's curvature and the gains that move with throughput.

6.9 Discussion

The economics of the constraint layer can be read directly, because the application searches the throughput and severity plane every five minutes for the best point that breaks no limit and carries the answer beside the achieved margin. At the design conditions that best legal point is 290 t/h at 530 °C, worth $17,982 per hour against the $15,348 the design line-up collects, so the fixed line-up leaves $2,634 per hour on the table. The selective layer of Figure 6.3 recovers $1,887 per hour of that gap in the base case of Table 6.8 without any optimizer in the loop. What remains is the severity dimension, which no controller in this application moves, and that is the natural boundary between a regulatory constraint layer and a real-time optimizer.

The value of a point of conversion sets the scale for everything else. Raising severity from 525 to 528.2 °C at fixed throughput buys 1.06 points of conversion and $248 per hour, about $234 per point, which is close to two million dollars a year on a unit of this size. That is why a cat cracker carries advanced control at all, and it is why the reactor and regenerator were among the first units in a refinery to receive a multivariable constraint controller [15].

The model also supports methods beyond the implemented structures. The fifteen-state plant with five explicit limits and a written objective is the natural form for model predictive control. The predictive controller of Section 6.5.5 already absorbs the selector and the override into one optimization with the constraints written in, but it leaves severity to the operator, and a layer that moves severity and throughput together rather than one at a time is the next step [6]. The choice of which variables to hold and which to let float is the partial control question that this unit posed [9], and the application makes it concrete because holding all five limits with two actuators is not possible and the student can watch which one gives. Moving horizon estimation on the same equations would recover the catalyst activity and the standpipe aeration from the slide valve position and the yields, turning the two invisible failures of scenarios 8 and 9 into soft sensors.

> **OPEN CHALLENGE** **A Feed That Is Too Easy to Crack**
>
> Light feeds strain the heat balance of this chapter from the other side. A 2015 review names the growing use of tight oils as a challenge for cat crackers [1]. These oils are low in carbon residue and aromatics, which makes the coke deposit, and with it the heat balance of the unit, harder to control. In the review's words the feed is in some respects too easy to crack. Its high naphtha and gas yields can also disturb the columns and gas plant downstream and limit throughput. Process changes and added heavier feed answer this only in part [1]. Progress would be a unit that holds its heat balance from such a feed to the heavy slate of scenario 7.

6.10 Conclusion

This chapter studied a fluid catalytic cracking riser and regenerator, a plant that supplies its own reaction heat and then has to be operated against limits it moves itself. The simulator integrates fifteen states, with a four-lump riser marched against catalyst contact, a coke burn and a flue gas oxygen balance solved together in closed form, and a slide valve whose implicit equation is solved rather than iterated. A three percent slide valve move in manual operation raises the riser outlet 7.17 °C against the 3.42 °C the extra catalyst alone accounts for, a heat balance loop gain of 0.523 and an amplification of 2.09, and only 29 % of that answer has arrived at two minutes. The settled coke burn rate moves 0.3 % for a 9.3 % increase in air while the carbon on the regenerated catalyst falls by a third, and the afterburn peaks at 57.8 °C above the bed at 0.70 % oxygen rather than at either end of the sweep. A low selector over four constraint controllers with external reset holds 284.5 t/h against a 330 t/h request and gives back 23 t/h when the afternoon warms 23 degrees, while a fifth controller on the feed preheat parks the dense bed at 729.1 °C on a heavy feed the shared handle cannot help with. Over the 420 min graded shift with five timed events the structure earns $95,391 with no time above the cyclone limit against $84,532 and 296 min above it. A model predictive controller on the same handles crosses no limit in seven cases, at 1.0 % less feed on scenario 13 and a worse oxygen

IAE on two. A fuller model would add a catalyst cooler as a third actuator on the heat balance, a standpipe fluidization model, a fractionator so that the wet gas rate stops being one yield lump, and an extension of the shipped MPC to optimize severity as well as throughput.

KEY POINTS

- A cat cracker has no furnace. The coke it burns is its only heat, so every move comes back to it amplified, 2.09 times here, and slowly.
- Only 29 % of the riser outlet response has arrived at two minutes, so a loop tuned on that reading comes out too aggressive.
- More air lowers the carbon on the catalyst but not the settled burn rate, which the riser sets.
- A low selector with external reset lets four constraint controllers share one feed rate without windup.
- The regenerator bed does not respond to feed rate, so its override acts on the feed preheat.

Where else this applies. A selector whose losing integrators track the applied command guards the fuel valve in Volume 4, Chapter 3, a feed rate pushed against limits that change returns in Volume 4, Chapter 4 and Volume 4, Chapter 7, and amplification by a recycle in Volume 6, Chapter 7.

References

[1] E. T. C. Vogt and B. M. Weckhuysen, "Fluid catalytic cracking: recent developments on the grand old lady of zeolite catalysis," *Chemical Society Reviews*, vol. 44, no. 20, pp. 7342–7370, 2015.

[2] W. M. Edwards and H. N. Kim, "Multiple steady states in FCC unit operations," *Chemical Engineering Science*, vol. 43, no. 8, pp. 1825–1830, 1988.

[3] A. Arbel, I. H. Rinard, R. Shinnar, and A. V. Sapre, "Dynamics and control of fluidized catalytic crackers. 2. multiple steady states and instabilities," *Industrial & Engineering Chemistry Research*, vol. 34, no. 9, pp. 3014–3026, 1995.

[4] F. G. Shinskey, *Process Control Systems: Application, Design, and Tuning*, 4th ed. New York: McGraw-Hill, 1996.

[5] A. Reyes-Lúa and S. Skogestad, "Systematic design of active constraint switching using classical advanced control structures," *Industrial & Engineering Chemistry Research*, vol. 59, no. 6, pp. 2229–2241, 2020.

[6] E. M. B. Aske, S. Strand, and S. Skogestad, "Coordinator MPC for maximizing plant throughput," *Computers & Chemical Engineering*, vol. 32, no. 1-2, pp. 195–204, 2008.

[7] A. A. Avidan and R. Shinnar, "Development of catalytic cracking technology. a lesson in chemical reactor design," *Industrial & Engineering Chemistry Research*, vol. 29, no. 6, pp. 931–942, 1990.

[8] A. Arbel, Z. Huang, I. H. Rinard, R. Shinnar, and A. V. Sapre, "Dynamic and control of fluidized catalytic crackers. 1. modeling of the current generation of FCC's," *Industrial & Engineering Chemistry Research*, vol. 34, no. 4, pp. 1228–1243, 1995.

[9] A. Arbel, I. H. Rinard, and R. Shinnar, "Dynamics and control of fluidized catalytic crackers. 3. designing the control system: Choice of manipulated and measured variables for partial control," *Industrial & Engineering Chemistry Research*, vol. 35, no. 7, pp. 2215–2233, 1996.

[10] V. W. Weekman and D. M. Nace, "Kinetics of catalytic cracking selectivity in fixed, moving, and fluid bed reactors," *AIChE Journal*, vol. 16, no. 3, pp. 397–404, 1970.

[11] L.-S. Lee, Y.-W. Chen, T.-N. Huang, and W.-Y. Pan, "Four-lump kinetic model for fluid catalytic cracking process," *The Canadian Journal of Chemical Engineering*, vol. 67, no. 4, pp. 615–619, 1989.

[12] H. de Lasa and A. Errazu, "Ignition of a fluidized bed catalytic cracking regenerator: Freeboard region influence," in *Fluidization*. Boston, MA: Springer US, 1980, pp. 563–570.

[13] M. Hovd and S. Skogestad, "Procedure for regulatory control structure selection with application to the FCC process," *AIChE Journal*, vol. 39, no. 12, pp. 1938–1953, 1993.

[14] R. C. McFarlane, R. C. Reineman, J. F. Bartee, and C. Georgakis, "Dynamic simulator for a model IV fluid catalytic cracking unit," *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 275–300, 1993.

[15] P. Grosdidier, A. Mason, A. Aitolahti, P. Heinonen, and V. Vanhamäki, "FCC unit reactor-regenerator control," *Computers & Chemical Engineering*, vol. 17, no. 2, pp. 165–179, 1993.

[16] R. Sadeghbeigi, *Fluid Catalytic Cracking Handbook*, 3rd ed. Oxford, UK: Elsevier, 2012.

[17] S. J. Qin and T. A. Badgwell, “A survey of industrial model predictive control technology,” *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[18] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, “Constrained model predictive control: Stability and optimality,” *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[19] G. Pannocchia and J. B. Rawlings, “Disturbance models for offset-free model-predictive control,” *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[20] D. Goldfarb and A. Idnani, “A numerically stable dual method for solving strictly convex quadratic programs,” *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[21] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, “Review on model predictive control: an engineering perspective,” *The International Journal of Advanced Manufacturing Technology*, vol. 117, no. 5–6, pp. 1327–1349, 2021.

[22] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[23] U.S. Chemical Safety and Hazard Investigation Board, “FCC unit explosion and asphalt fire at Husky Superior refinery, Superior, WI, incident date April 26, 2018,” U.S. Chemical Safety and Hazard Investigation Board, Washington, DC, Investigation Report 2018-02-I-WI, 2022.

[24] ——, “ExxonMobil Torrance refinery electrostatic precipitator explosion, Torrance, California, incident date February 18, 2015,” U.S. Chemical Safety and Hazard Investigation Board, Washington, DC, Investigation Report 2015-02-I-CA, 2017.

[25] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, “Guide to operational technology (OT) security,” National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

6.11 Exercises

6.1 Concept check · 2 min

In scenario 1 the regenerated catalyst slide valve opens from 45 to 48 % in Manual with the air, the feed and the preheat held. The regenerator dense bed ends 5.25 °C hotter, but for the first few minutes it falls, by 0.46 °C at three minutes. Why does it fall first?

(a) The carbon on the regenerated catalyst drops at once, so the bed burns less coke for a few minutes

(b) The bed thermocouple lags the bed by a few minutes and shows the old temperature drifting

(c) The extra catalyst carries heat out of the bed before the extra coke it brings back has arrived to burn

(d) The higher circulation lowers the bed level, so the thermocouple reads the cooler dilute phase

6.2 Concept check · 2 min

The heavier feed of scenario 7 takes the regenerator dense bed over its 730 °C limit while the riser outlet is held at 525 °C. Which handle brings the bed back, and why?

(a) The feed rate cut, because less feed sends less coke to the regenerator to burn

(b) The fresh catalyst addition raised, because a more active catalyst lays down less coke per tonne

(c) The slide valve opened, because more circulation carries more of the bed’s heat to the riser

(d) The feed preheat lowered, because more catalyst must circulate and the coke is spread over more of it

6.3 Concept check · 1 min

In scenario 10 the air blower controller holds the feed rate at 284.5 t/h. What are the other three constraint controllers doing?

(a) Each sits on its upper limit, the selected feed plus 0.10 kg/s, with its integral parked there
(b) Each has wound its integral up to the top of the feed range and will unwind when its fence comes near
(c) Each has been put in manual by the selector and restarts from the selected feed when its fence is reached
(d) Each tracks the selected feed exactly, so all four outputs are equal and the selector picks the first

6.4 Theory and calculation · 7 min

In scenario 1 the slide valve opens from 45.0 to 48.0 % in Manual. Solving the riser energy balance alone, with the regenerator frozen, gives a riser outlet rise of 3.42 °C. The unit settles 7.17 °C higher, and 2.11 °C of that has arrived two minutes after the move.

(a) Find the amplification A and the loop gain L, and the factor by which Equation (1.8) stretches a time constant inside the loop. *Checkpoint:* $A = 2.10$, $L = 0.523$, a factor of 2.10.

(b) For a one percent valve step the loop gain is 0.534 at 525 °C and 0.550 at 535 °C. Find the amplification at each, and the percentage by which the settled gain grows between them. *Checkpoint:* 2.15 and 2.22, a growth of 3.6 %.

(c) An engineer fits a process gain to the first two minutes. Find that gain in °C per percent of valve, the settled gain, the fraction of the answer seen and how many times too weak the fitted plant is. *Checkpoint:* 0.70 against 2.39 °C per percent; 29 %, 3.4 times too weak.

6.5 Theory and calculation · 6 min

The blower capacity follows Equation (6.21) with $F_{air,0}^{max} = 64.0\,\mathrm{kg/s}$ at 15 °C. Its approach span is 10.0 kg/s, and one span of the blower fence is worth 48.5 t/h of feed with the riser outlet and the oxygen held.

(a) Find the capacity in t/h at 15, 38 and 12 °C, and the fraction of mass capacity lost from 15 to 38 °C. *Checkpoint:* 230.4, 213.4 and 232.8 t/h; 7.4 %.

(b) With the blower holding the feed at a fixed approach, estimate the feed given up in the afternoon and gained in the evening, and compare with the measured 284.5, 261.5 and 287.8 t/h of Table 6.8. *Checkpoint:* 22.9 t/h lost and 3.3 gained, against 23.0 and 3.3 measured.

6.6 Theory and calculation · 8 min

Use Equation (6.16) with $K_s = 1.39032$, $K_{line} = 62\,\mathrm{kPa}$, $F_{cat,0} = 6.8F_0$ and $F_0 = 69.444\,\mathrm{kg/s}$, $P_{rgn} - P_{rx} = 40\,\mathrm{kPa}$, $H_{drop} = 9.5\,\mathrm{m}$, $M = 165\,000\,\mathrm{kg}$, $\rho_b = 480\,\mathrm{kg/m^3}$, $A_{rgn} = 78.5\,\mathrm{m^2}$ and $g = 9.81\,\mathrm{m/s^2}$.

(a) Find h_b, ΔP_{av}, the circulation at $u = 45\,\%$ in t/h and the differential ΔP_v across the valve. *Checkpoint:* 4.38 m, 119.0 kPa, 1700 t/h and 57.0 kPa.

(b) The standpipe aeration falls to $a_{er} = 0.74$. Find ΔP_{av}, and the opening and the differential that hold the design circulation. *Checkpoint:* 98.4 kPa; 56.3 % open with 36.4 kPa.

(c) At $a_{er} = 0.74$, find the circulation and the differential at 70 and 100 % open. *Checkpoint:* 1824 t/h with 27.1 kPa; 1967 t/h with 15.4 kPa.

6.7 App challenge · 8 min

Open scenario 1, “The heat balance by hand”, which starts in Manual, with the default seed. Set the speed to 20× and press Run. At about 10 min move Regenerated catalyst slide valve from 45 to 48 %, run on to at least 135 min, and export the CSV. Let t_s be the time of the last row in which `slide_cmd_pct` still reads 45.0. Read the change of each column from the row at t_s.

(a) Read `rot_C` at $t_s + 2$ and `trgn_C` at $t_s + 3$ minutes. *Checkpoint:* $+2.11 \pm 0.10$ and $-0.46\,°\mathrm{C} \pm 0.05$.

(b) Read both at $t_s + 120$ minutes, and the peak of `crc_wtpct` with its time after t_s. *Checkpoint:* $+7.05 \pm 0.05$ and $+5.09\,^{\circ}\mathrm{C} \pm 0.05$; 0.0577 % at 13 min.

(c) Take the settled rise of 7.17 °C from Section 6.4. What fraction of it has arrived at 2 and at 120 minutes, and which state sets the slow part? *Checkpoint:* 29 % and 98 %; the regenerator bed.

6.8 App challenge · 8 min

Open scenario 7, “A heavier feed, and the fence feed rate cannot reach”, with the default seed. It runs on Constraint control, and the heavier feed arrives at 20 min. Set the speed to 20×.

(a) Run to 120 min and export the CSV. Read the peak of `trgn_C` and its time, the lowest `tpre_C` and its time, `trgn_C` and `tpre_C` at 120 min, and the minutes above the bed limit in the last row of the column `metric_over_rgn_min`. *Checkpoint:* 733.3 °C at 52.5 min; 186.5 °C at 81.7 min; 729.0 and 189.2 °C; 35.6 min.

(b) At any time between 120 and 150 min, switch the mode to Manual and set Feed rate to 200 t/h. Run on 60 min and read `trgn_C` and `rot_C` 15 and 60 minutes after the cut. *Checkpoint:* Bed 733.7 to 733.8 °C and riser 555.0 to 555.1 °C at 15 minutes; bed 742.1 to 742.2 °C and riser 561.0 to 561.1 °C at 60.

6.9 App challenge · 8 min

Open scenario 11, “The fence moves”, with the default seed. It runs on Constraint control with a 330 t/h target, the afternoon warms to 38 °C at 90 min and the evening cools to 12 °C at 260 min. Run 400 min at 20× and export the CSV.

(a) Read `feed_tph`, `air_max_tph` and `margin_usd_h` at 85, 250 and 400 min, and the `fence` column at each. *Checkpoint:* 284.5, 230.4, 17235; 261.5, 213.4, 15986; 287.8, 232.8, 17408; air blower each time.

(b) Find the lowest `feed_tph` between 90 and 130 min and its time, and the highest `util_air_pct` in the same window. *Checkpoint:* 236.7 t/h at 92.0 min; 143.7 % at 90.1 min.

(c) Compute the feed given up per tonne of lost air capacity in the afternoon, and explain why the feed falls below its new level first. *Checkpoint:* 1.35 tonnes of feed per tonne of air capacity.

6.10 Application area · 3 min

An operator on a unit with the promoter in service reduces the excess oxygen setpoint from 2.00 to 0.85 % to save blower power, and reports that the regenerator dense bed has cooled from 700 to 686.6 °C.

(a) Is the cooler bed good news? Use Figure 6.2. *Checkpoint:* No: the monoxide leaving the bed burns above it.

(b) The promoter feeder then plugs. Why is adding air no longer a complete answer on this unit? *Checkpoint:* The blower saturates and 8.7 °C of afterburn remains.

6.11 Application area · 4 min

Section 6.9 gives the economics of the unit at its design conditions.

(a) One point of conversion is worth $234 per hour at fixed throughput. What is it worth in a year of 8760 hours on stream? *Checkpoint:* About $2.05 million.

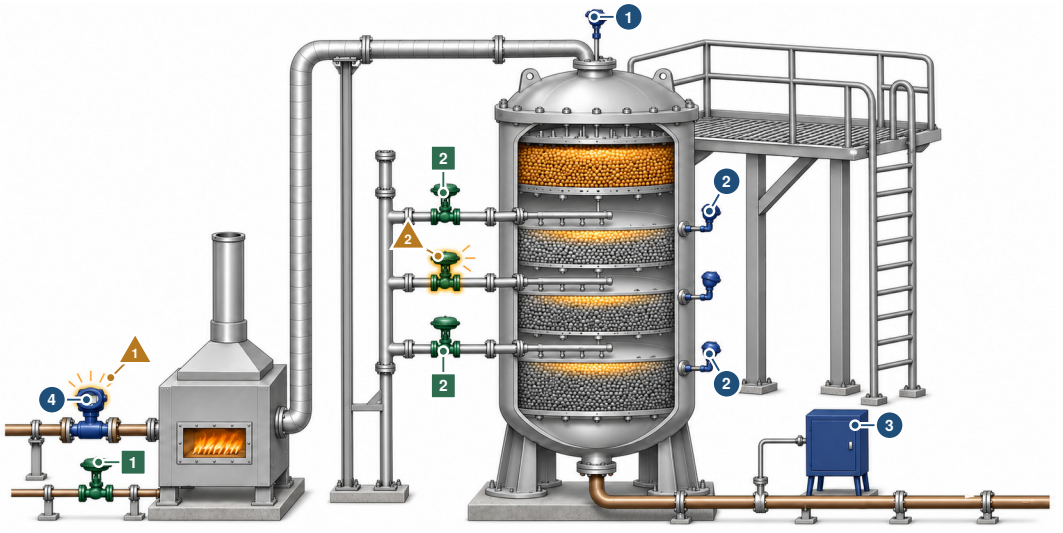
(b) The best legal point is worth $2,634 per hour more than the design line-up, and the selector recovers $1,887 per hour of it. What fraction is recovered, what remains per year, and which layer would have to recover it? *Checkpoint:* 72 %; about $6.5 million a year; an optimizer that moves severity.

6.12 Review the agent · 5 min

An AI assistant is asked to tune the riser outlet loop. It runs scenario 4, “The three basic loops, holding still”, switches to Manual, steps the slide valve from 45.0 to 46.0 % and reads the riser outlet: up 0.47 °C at one minute and 0.74 °C at two, when the trend looks flat. It reports a first-order plant with $K_p = 0.74\,^{\circ}\mathrm{C}$ per percent and $\tau = 1$ min, and proposes lambda tuning, $K_c = \tau/(K_p\lambda)$ with $\lambda = 1$ min and $\tau_I = \tau$, which gives $K_c = 1.35$ percent per °C, “faster than the shipped 0.55”. Find the flaw and say how you would check it.

CHAPTER 7

A Hydrocracker Reactor Simulator for Open-Loop Instability and Layered Protection



The controller moves 1 firing of the feed heater 2 three cold hydrogen valves between the beds

The controller reads 1 first bed inlet temperature 2 inlet temperatures of the three lower beds 3 conversion of the product 4 feed rate

What disturbs it ▲1 a feed cut ▲2 a stuck hydrogen valve ▲ a more active catalyst

The warm first bed at the top has no cold hydrogen valve above it, so only the slow feed heater at the lower left can cool it. Illustration, not to scale.

IN PLAIN TERMS A Reactor That Can Heat Itself Up

The main idea. A hydrocracker breaks heavy oil into lighter fuel as the oil passes down through four beds of solid grains. The breaking gives off heat, and hot oil breaks faster, so a bed can heat itself faster and faster.

The challenge. Cold gas blown in between the beds cools the lower three, but the first has only a slow heater. A guard that turns the heat down can be fooled, because its reading moves the wrong way for minutes after each cut.

Why it matters. Once a bed runs away no control is fast enough, and the reactor must be shut down.

THREE WAYS TO READ THIS CHAPTER

Explore	13 min	This page, the History boxes in Sections 7.2 and 7.3, the Did you know? boxes in Sections 7.1 and 7.6.4, the Try it boxes in Sections 7.5 and 7.6, and Key points in Section 7.10. No equations.
Study	50 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Section 7.6 with Figure 7.4, Section 7.7 and Section 7.8 with its Open challenge box.

The control question An adiabatic catalyst bed with an exponential rate law can be unstable before any controller is added, and one number decides it. How close is the first bed to that boundary, and which layers of control and protection can still act once a feed cut or a hotter catalyst pushes it there?

Learning objectives

- Compute the runaway index and the amplification of a bed from its rise and mean temperature.
- Explain why the interbed quench loops break the chain of amplifying beds, and why bed 1 has no fast handle.
- Explain why a feed cut at a fixed inlet makes the bed hotter.
- Show that an override reading a raw bed rise is positive feedback in the transient, and that one matched lag cures it.
- Compare one inlet setpoint cut by an override with a predictive controller that sets a target per bed.

Prerequisites Sections 1.4.4, 1.4.7, 1.6.2 and 1.6.3.

The app https://apmonitor.com/apps/hydrocracker/

See also Volume 7, Chapter 2 for a stirred tank at an ignition boundary. Volume 4, Chapter 2 for a protective override that rests at its clamp and Volume 4, Chapter 6 for a per-zone safety override that bounds the throughput.

7.1 Introduction

An adiabatic catalyst bed with an exponential rate law is one of the few plants a control course meets that can be unstable without a controller. Write the bed as a lumped stage reacting at the mean of its inlet and outlet temperatures, linearize, and a single dimensionless group appears: $r = \frac{1}{2} K \, \mathrm{d}x/\mathrm{d}T$, where K is the adiabatic temperature rise at complete conversion and x is the fraction converted in the bed. Below $r = 1$ the bed settles by itself and amplifies its inlet by $(1 + r)/(1 - r)$. Above it the operating point is unstable and only a controller can hold the reactor there. Section 1.4.4 derives the index on the running example of this volume, and Volume 7, Chapter 2 meets the same ignition boundary in a jacketed stirred tank.

Computed on the reactor modeled here, the first bed sits at $r = 0.538$ and turns a small inlet disturbance into 3.3 times as much outlet. An inlet 3.8 °C hotter, or eighteen percent more catalyst activity, crosses the boundary. Four beds at a gain of 3.3 in series would turn one degree at the feed heater into more than a hundred at the bottom of the reactor, which is why there is a quench valve between every pair of beds. The quench loops are not trimming. They are what breaks the feedback path three times, and a reactor with one bed of the same total catalyst would be inoperable.

> **DID YOU KNOW? Twenty Degrees Doubles the Rate**
>
> At normal pressure and flow, a rise of 20 °F, about 11 °C, almost doubles the rate of the hydrocracking reaction [1]. To hold it, each reactor has several catalyst beds with cool hydrogen injected between them [1]. The first bed of this chapter rises 23.9 °C from inlet to outlet at design.

The misconception the simulator corrects is that stability is a property of the control loop. On this reactor it is a property of the plant, it is a computable number, and it moves with every catalyst load and every feed cargo. A second misconception follows from the first. Bed 1 has no quench above it, so its only handle is a feed heater upstream with a 2.5 min lag, and once the first bed is past the boundary there is nothing in the control system fast enough. The answer is a trip. An actuator that can act only upstream of a fast instability is not a protection.

The class is general. Parametric sensitivity in fixed bed reactors, the property that a small parameter change produces a disproportionate change in the peak temperature, has a classical treatment from 1970 [2] and a general one later [3], and the safety literature on reactor runaway catalogs the mechanisms across chemistries [4]. This chapter works with an interactive hydrocracker that reports

the runaway index of every bed beside its temperatures. The index can be checked there against an amplification measured by experiment, the three protective layers can be built and audited, and an override turns out to be positive feedback when it reads the wrong measurement. The reactor has four beds with their quench mixing, an instrument chain, an override with a dynamically compensated measurement and a three-way trip, and the chapter gives them in the detail needed to rebuild them. An optional model predictive controller (MPC) of the four inlets is compared on measured runs with the shipped loops, and each of its gains is traced to its mechanism by a counterfactual run. The scenarios include an override that trips the reactor it was installed to protect, and a feed cut that makes the reactor hotter. The lumped bed model, the runaway index and override control are established, and the MPC takes the industrial form [5].

Section 7.2 gives the domain history. Section 7.3 states the model and its assumptions, and Section 7.4 derives the amplification of four beds in series and the transient sign of a bed rise. Section 7.5 describes the inlet loops, the override, the trip, the conversion loop and the MPC. Section 7.6 reports the measured results, Section 7.7 sets the hardware of a real unit beside the model, Section 7.8 gives the limitations, and Sections 7.9 and 7.10 draw the lessons that carry beyond this unit.

7.2 Background

Hydrocracking converts vacuum gas oil into distillate over a bifunctional catalyst at high hydrogen pressure. The cracking reactions are exothermic, the beds are adiabatic, and the heat released is proportional to the conversion, so a hydrocracker is among the highest-pressure, highest-value and most carefully protected units in a refinery.

The reactor is built as several beds in series with cold hydrogen injected between them, and the reason is thermal rather than chemical. Kinetic modeling of the conversion itself is a lumping problem, because the feed is a continuum and the products are a distribution. Discrete lumped kinetic models for vacuum gas oil hydrocracking and the discrimination between them have their own literature [6], and lumping approaches specific to dual bed reactors have been compared directly [7]. None of that is needed to state the control problem, which depends only on how much heat a unit of conversion releases and on how fast conversion rises with temperature.

The theory of what happens when a fixed bed exotherm outruns its cooling is older than the unit, and the History box below traces it. Modern treatments extend it to the maldistribution mechanisms that trigger runaway in trickle beds specifically [8], and the safety literature surveys the whole field [4]. A hydrocracker differs from the classical fixed bed case in one respect that matters for control: the cooling is not a wall, it is an injection between beds, so the protection is a valve rather than a heat transfer coefficient and it can be manipulated. Chapter 8 uses the same interbed hydrogen quench on a diesel hydrotreater.

> ⧖ HISTORY **1970: Runaway Gets a Criterion**
>
> The theory of a reactor that outruns its cooling is older than the modern hydrocracker, a process further developed in the early 1960s [9]. Semenov's balance of the heat an exothermic reaction releases against the heat removed laid the groundwork, and Bilous and Amundson did pioneering work on parametric sensitivity, showing how the peak temperature along a reactor varies with the coolant temperature [4]. In 1970 Van Welsenaere and Froment gave criteria for parametric sensitivity and runaway in fixed bed catalytic reactors [2], drawn from the locus of temperature maxima in the plane of temperature and conversion [4]. Morbidelli and Varma, whose work includes a 1986 paper on the same reactors [3], set the boundary of runaway where the sensitivity itself peaks and generalized the criterion to other parameters [4]. The runaway index of Section 7.3 asks the same question of one lumped bed.

Control developed in layers. The earliest arrangement is manual quench adjustment against bed thermocouples, with the operator reading the temperature rise across each bed as the visible

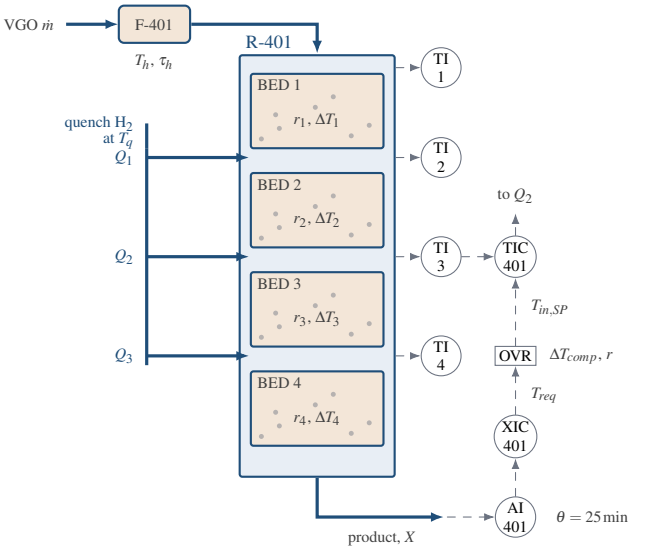


Figure 7.1 Hydrocracker reactor R-401. Vacuum gas oil $\dot{m}$ is heated in F-401 to T_h with lag τ_h and passes down through four adiabatic beds, each with its own runaway index r_i and temperature rise ΔT_i. Cold quench hydrogen at T_q is injected through valves Q_1, Q_2 and Q_3 between the beds. Bed inlet thermocouples TI-1 through TI-4 feed four inlet temperature loops, of which TIC-401, reading TI-3 and moving Q_2, is representative. AI-401 reports conversion X on cycle θ to the supervisory loop XIC-401, whose request T_{req} the override OVR can only lower before it becomes the common inlet setpoint $T_{in,SP}$. Bed 1 has no quench above it, and the trip is not drawn.

form of the reaction rate. Automatic inlet temperature control on the quench valves came next and is universal. A supervisory conversion loop on a product distillation became possible when the measurement did. Above that, the hydrocracker was among the first units in a refinery to receive a large multivariable controller, and the reason was precisely the constrained economics of an exothermic reactor with hard temperature limits: Cutler and Hawkins reported a dynamic matrix controller on a hydrocracker second stage reactor in 1988 [10], one of the earliest published industrial applications of the technology that became standard [5]. A recent study sets the reactor temperatures of an industrial hydrocracker with a feedforward model predictive controller driven by feed measurements, for constant diesel quality [11]. The regulatory layer underneath that controller is the subject of this chapter, together with the override and trip layers above it.

7.3 Process Model

Reactor R-401 processes 250 t/h of vacuum gas oil with recycle hydrogen at about 150 bar through four adiabatic catalyst beds. Every bed inlet is at 381 °C at the design condition, the bed rises are 23.9, 14.0, 11.1 and 9.3 °C for 51.2 % conversion, and the quench flows that return each inlet to 381 °C are 3.38, 1.99 and 1.57 t/h. Figure 7.1 shows the arrangement, and the variables are classified in Table 7.1.

Table 7.1 Classification of Variables at the Design Operating Point

Symbol	Role	Nominal	Range
$T_{in,i}$	CV, continuous	381 °C	340–400 °C SP
X (AI-401)	CV, analyzer	0.512	0.35–0.72 SP
ΔT_{max}	constraint	23.9 °C	alarm 33, trip 42
r_{max}	constraint	0.538	unstable above 1
T_h (F-401)	MV	381 °C	300–420 °C
Q_1, Q_2, Q_3	MV	3.38, 1.99, 1.57 t/h	0–30 t/h
$\dot{m}$	DV, measured	250 t/h	150–340 t/h
activity	DV, unmeasured	1×	0.6–1.6×
stuck valve	DV, unmeasured	off	on/off
$T_{out,i}$	states	profile	—

7.3.1 One Adiabatic Bed

Each bed is a lumped adiabatic stage reacting at the mean of its inlet and outlet temperatures, Equation (7.1), with the fraction converted in bed i given by Equation (7.2).

$$\tau \frac{\mathrm{d}T_{out}}{\mathrm{d}t} = T_{in} + K\,x\!\left(\frac{T_{in}+T_{out}}{2}\right) - T_{out} \tag{7.1}$$

$$x_i(T) = \min\left\{0.9,\; k_{ref}\,\frac{\dot{m}_0}{\dot{m}}\,a\,R_i \times \exp\!\left[-\frac{E}{R}\left(\frac{1}{T}-\frac{1}{T_{ref}}\right)\right]\right\} \tag{7.2}$$

where $\dot{m}_0/\dot{m}$ is the inverse space velocity relative to the design feed, a is the catalyst activity and R_i is the fraction of the feed still unconverted at the inlet of bed i, with $K = 90.0\,^{\circ}\mathrm{C}$ per unit of x, $E/R = 20\,000\,\mathrm{K}$, $T_{ref} = 392.5\,^{\circ}\mathrm{C}$ and $\tau = 3.5\,\mathrm{min}$. The unconverted fraction passes down as $R_{i+1} = R_i(1-x_i)$, and the reactor conversion is $X = \sum_i x_i R_i$. The remaining parameters are given in Table 7.2.

7.3.2 The Runaway Index

Linearizing Equation (7.1) about a steady state and collecting terms gives Equation (7.3), the result of Equation (1.6) with K in place of ΔT_{ad}.

$$\tau \frac{\mathrm{d}\,\delta T_{out}}{\mathrm{d}t} = (r-1)\,\delta T_{out} + (1+r)\,\delta T_{in}, \qquad r = \frac{K}{2}\frac{\mathrm{d}x}{\mathrm{d}T} \tag{7.3}$$

The bed is stable while $r < 1$ and unstable above it, and while stable it is an amplifier with the steady-state gain of Equation (7.4).

$$\frac{\partial T_{out}}{\partial T_{in}} = \frac{1+r}{1-r} \tag{7.4}$$

For an Arrhenius rate below the 0.9 ceiling, $\mathrm{d}x/\mathrm{d}T = xE/(RT^2)$ with T in kelvin, so $r = KxE/(2RT^2)$ and the index rises with conversion and with temperature together.

7.3.3 Quench Mixing

Cold hydrogen at 60 °C mixes with the bed effluent to set the next inlet, Equation (7.5).

$$T_{in,i+1} = \frac{\dot{m}\,c_{p,oil}\,T_{out,i} + \dot{q}\,c_{p,H_2}\,T_q}{\dot{m}\,c_{p,oil} + \dot{q}\,c_{p,H_2}} \tag{7.5}$$

Hydrogen has five and a half times the specific heat of the oil, so a few tonnes an hour moves an inlet temperature a long way. At design, 3.38 t/h takes bed 2 from 405 °C back to 381.

Table 7.2 Model Parameters

Symbol	Meaning	Value	Class
K	local heat-rise factor	90.0 °C	fixed
T_{ref}	reference temperature	392.5 °C	fixed
E/R	activation temperature	20 000 K	adjustable
k_{ref}	conversion per bed at reference	0.2600	adjustable
τ	bed thermal time constant	3.5 min	adjustable
τ_h	feed heater lag	2.5 min	adjustable
τ_q	quench valve lag	0.30 min	adjustable
τ_{TT}	thermocouple filter	0.15 min	adjustable
$c_{p,H_2}/c_{p,oil}$	specific heat ratio	5.5	adjustable
T_q	quench hydrogen temperature	60 °C	adjustable
θ	conversion analyzer cycle	25 min	adjustable
τ_{lab}	conversion analyzer averaging	3 min	adjustable
ΔT alarm	bed rise alarm	33 °C	visible
ΔT trip	bed rise trip	42 °C	visible
T trip	bed temperature trip	440 °C	visible
rate trip	rate of rise trip	28 °C/min	visible
τ_{rate}	rate of rise filter	0.8 min	visible

7.3.4 Why Less Feed Is Hotter

Conversion per bed scales inversely with space velocity, so cutting the feed at a fixed inlet temperature gives more conversion per unit of catalyst and the heat released has less mass to heat. At a fixed inlet of 381 °C, cutting the feed from 250 t/h to 220 t/h takes the steady first bed rise from 23.9 °C to 34.1 °C and the runaway index from 0.538 to 0.756. Below 212 t/h there is no cool steady state at all. A feed cut on a hydrocracker is a heat problem rather than a throughput problem, and it is one of the classic ways units have run away.

🕮 **Pause.** Try Exercise 7.1 before reading on.

7.3.5 Assumptions

Each bed is one lumped adiabatic stage reacting at the mean of its inlet and outlet temperatures. A real bed has an axial profile whose peak is hotter than the outlet, so a real reactor is closer to its limits than its outlet thermocouples suggest.

One lumped cracking reaction carries one apparent activation energy. Real hydrocracking is a network and the hydrotreating reactions in the first bed release heat of their own, so the lumping literature exists for good reasons [6, 7].

The retained thermal surrogate is calibrated on local conversion rather than on whole-feed reaction enthalpy. Bed i releases Kx_i of heat while X credits it with x_iR_i, so the four design rises sum to 58.3 °C while KX is 46.1 °C. Bed 1, where $R_1 = 1$, is unaffected, and beds 2 to 4 run hotter than a constant whole-feed reaction enthalpy would imply. The local correlation preserves the calibrated amplification, fold and protective-layer lessons, but its temperature rises cannot be used as a reaction-duty balance.

Quench mixing is perfect and instantaneous. Maldistribution in a quench zone is one of the commonest causes of real hot spots and this model cannot show it [8].

Hydrogen partial pressure and recycle gas purity are fixed, and there is no catalyst deactivation. A real cycle is twelve months of slowly rising weighted average bed temperature to maintain conversion. The index also depends on conversion: temperature alone at fixed conversion slightly lowers it in this model. The application represents catalyst aging through the activity slider.

The trip is instantaneous. It drops the heater command to its 300 °C minimum and opens all three quench valves to 30 t/h at once, and the feed is not changed. A real emergency depressuring takes several minutes at a controlled rate and the beds keep reacting while it happens.

The conversion analyzer is a sample-and-hold on a 25 min cycle followed by a 3 min averaging lag, so its reading is between 0 and 25 min old and carries no dead time beyond that. Conversion on a real unit is inferred from a distillation and is both slower and less certain.

> ⧖ HISTORY **1997: A Runaway the Quench Could Not Stop**
>
> On January 21, 1997 a hot spot formed in the fourth of five catalyst beds of a hydrocracker reactor at a California refinery and spread to the bed below [1]. One bed outlet reading rose from 628 to 823 °F in 40 seconds. The operators distrusted readings that jumped between normal values and zero. They raised the quench and cut the furnace, and did not depressure the reactor as the written procedure required above 800 °F. The effluent pipe ruptured at a temperature likely above 1400 °F. One operator was killed and 46 people were injured. The investigators calculated that at least three times the maximum design quench flow would have been needed, and that only depressuring could have slowed the reaction. They asked for every bed temperature in the control room, backup indicators and tested emergency systems [1].

7.4 Dynamic Analysis

7.4.1 Four Beds in Series

Equation (7.4) gives the amplification of one bed. If each bed inlet followed the outlet above it one for one, four in series would multiply, Equation (7.6).

$$\frac{\partial T_{out,4}}{\partial T_h} = \prod_{i=1}^{4} \frac{1+r_i}{1-r_i} \tag{7.6}$$

At the design condition the four indices are 0.538, 0.321, 0.254 and 0.215, so the bed gains are 3.33, 1.95, 1.68 and 1.55 and the product is 17, and the first bed's gain alone repeated four times would give 122. With the three quench flows frozen at their design values the chain is weaker than the product, because each quench dilutes the effluent (by factors of 0.931, 0.958 and 0.967) and a hotter bed leaves less feed for the next, and the computed steady-state amplification from heater outlet to bed 4 outlet is 10 °C per °C. Either number is a reactor nobody can operate from a feed heater. A quench loop between each pair drives $\partial T_{in,i+1}/\partial T_{out,i}$ to zero at steady state, which reduces the chain to four independent beds of gain 3.33, 1.95, 1.68 and 1.55. The quench valves are not trimming the profile. They are the reason the vessel has four beds instead of one.

7.4.2 The Transient Sign of a Bed Rise

Take the bed as a first-order lag on the outlet and hold the reaction rise constant for the moment. Then $T_{out} = \mathscr{L}^{-1}\{T_{in}/(\tau s+1)\} + \Delta T_{rxn}$, and the measured rise responds to an inlet step of size a as Equation (7.7).

$$\Delta T_{meas}(t) = T_{out} - T_{in} = \Delta T_{rxn} - a\,e^{-t/\tau} \tag{7.7}$$

which for a negative step, that is a cut in the inlet, starts at $\Delta T_{rxn} + |a|$ and decays to ΔT_{rxn} with time constant τ. The measured rise therefore moves in the opposite direction to the steady-state answer for about $\tau = 3.5$ min after every correction. An override with gain K_o acting on that signal has an effective loop gain of $+K_o$ during that window rather than $-K_o$, which is the positive feedback reported in Section 7.6. Subtracting a lagged inlet, as in Equation (7.12), cancels the transient term in Equation (7.7) exactly and leaves ΔT_{rxn}.

🕮 **Pause.** Try Exercise 7.2 before reading on.

7.5 Controller Design

7.5.1 Velocity Form PID

Every loop uses the shared velocity-form PID of Section 1.6.2 with the derivative on the measurement, restated in this chapter's symbols as Equations (7.8) to (7.10).

$$e_k = a\,(\mathrm{SP}_k - \mathrm{PV}_k) \tag{7.8}$$

$$P_k = K_c e_k, \qquad D_k = -K_c \tau_D\, a\, \frac{\mathrm{PV}^f_k - \mathrm{PV}^f_{k-1}}{\Delta t} \tag{7.9}$$

$$\begin{aligned} s_k &= s_{k-1} + (P_k - P_{k-1}) + (D_k - D_{k-1}) \\ c_k &= s_k + I_{k-1}, \qquad u_k = \mathrm{clamp}(c_k, u_{\min}, u_{\max}) \\ I_k &= I_{k-1} + \left[\frac{K_c}{\tau_I} e_k + \frac{u_k - c_k}{\tau_t}\right]\Delta t \end{aligned} \tag{7.10}$$

where a is $+1$ for direct and -1 for reverse action, PV^f is the measurement through the derivative filter, s_k accumulates the proportional and derivative increments without limit and c_k is the demanded output. The integral is seeded with the initial output, and anti-windup is back-calculation with tracking time $\tau_t = \tau_I$. No output carries a rate limit. In Manual a tracking routine sets I so that c_k equals the operator's value, and in Auto the manual stations follow the controller after every execution. The controller scan is 0.05 min and integration is fixed-step fourth-order Runge-Kutta at 0.01 min.

7.5.2 Three Layers

Regulatory. Four inlet temperature loops. The feed heater holds bed 1, direct acting, with $K_c = 2.5$ and $\tau_I = 3.0$ min. Three quench valves hold beds 2, 3 and 4, reverse acting because more quench is a colder inlet, with $K_c = 0.90$ t/h per degree, $\tau_I = 1.5$ min, $\tau_D = 0.35$ min and a 0.20 min derivative filter. A rate contribution on a loop holding an amplifying plant is worth more here than almost anywhere else. A stuck valve sits after the PID, so the quench loop above it does not see the fault and winds its own output to its 30 t/h limit, which is the signature a stuck valve leaves on a trend.

Override. A one-directional layer pulls the common inlet setpoint down when a bed rise approaches its alarm or the runaway index approaches its limit, Equation (7.11).

$$\begin{aligned} T_{in,SP} &= \mathrm{clamp}\big(T_{req} - \max\left(c_{\Delta T}, c_r\right), T_{\min}, T_{\max}\big) \\ c_{\Delta T} &= K_o\left[\Delta T_{comp} - \Delta T_{lim}\right]^+ \\ c_r &= K_r\left[\hat{r}_{\max} - r_{lim}\right]^+ \end{aligned} \tag{7.11}$$

where T_{req} is the conversion loop output, or the operator's setpoint with that loop in Manual, $T_{\min} = 340\,^\circ$C, $T_{\max} = 400\,^\circ$C, $K_o = 2.5$, $\Delta T_{lim} = 28\,^\circ$C (5 °C below the alarm), $K_r = 60\,^\circ$C per unit and $r_{lim} = 0.75$. The index is read from thermocouples, $\hat{r}_{\max} = \max_i \Delta T_{comp,i}\, E/(2R\bar{T}_i^2)$ with $\bar{T}_i$ the bed's mean temperature and the declared E/R, which is r at a steady state (to 2×10^{-15}) and lags it in a transient. Its 0.75 limit is a rise of 33.1 to 33.8 °C, above the 28 °C where the steeper $c_{\Delta T}$ starts, so it sets the cut in none of 351 runs. The cut is proportional and recomputed every scan. It can only pull down, on purpose. A layer that can also push up is a controller, and the value of an override is that it is not one. The protective override of Volume 4, Chapter 2 is built the same way, resting at its clamp while the machine is healthy. While the override cuts and the conversion loop is in automatic, that loop's integral is also back-calculated against the applied setpoint, each scan adding $(T_{in,SP} - u_k)\,\Delta t/\tau_I$ to I_k, so it does not wind up behind the cut and settles where the cut equals $K_c e$. In scenario 10 the conversion loop ends asking for 385.7 °C against an applied 382.8 °C, and the largest rise holds at 29.2 °C for as long as the run continues.

Trip. The trip fires on a bed rise of 42 °C, a bed temperature of 440 °C, or a rate of rise of 28 °C/min on the hottest bed, filtered with a 0.8 min lag. A rate trip is meant to fire first in a runaway, because

the other two are lagging indicators of an accelerating event, but on this plant it cannot. The bed time constant caps how fast an outlet can move. In scenario 9, run in Manual through an activity step to 1.45 at 10 min, the rise trip fires at 16.6 min with the hottest bed at 423.1 °C while the filtered rate peaks at 2.7 °C/min (4.1 °C/min at an activity of 1.6).

> **TRY IT IN THE APP** **9 · Which protection fires first**
>
> Load the case and press Run. Touch nothing: the reactor is in Manual, and at 10 min the catalyst activity steps to 1.45. Watch the three trip lines. The bed rise trip fires at 16.6 min with the hottest bed at 423.1 °C, while the rate of rise peaks at 2.7 °C/min against its setting of 28 °C/min.

7.5.3 The Rise the Override Reads

A bed rise is an outlet minus an inlet, and the two do not move on the same clock. The inlet answers a quench valve within the valve's 0.30 min lag and the outlet carries the catalyst time constant of 3.5 min. Cut the inlet quickly and the measured rise goes up for several minutes before it comes down, so the signal is inverted by the action meant to fix it, and a proportional override reading it is positive feedback.

The cure is one lag. Since the bed is $\tau\,\mathrm{d}T_{out}/\mathrm{d}t = T_{in} + \Delta T_{rxn} - T_{out}$, the compensated rise of Equation (7.12) equals the reaction rise seen through the same lag, $\Delta T_{rxn}/(\tau s + 1)$, so the thermal transient of the inlet is removed exactly.

$$\Delta T_{comp} = T_{out} - \frac{1}{\tau s + 1} T_{in} \tag{7.12}$$

The override reads the bed temperatures directly, without the 0.15 min thermocouple filter of the inlet loops. This is the same move as compensating a tray temperature for pressure or a flow for density. The measured consequence of omitting it is in Section 7.6.

7.5.4 Supervisory Conversion

XIC-401 holds conversion by setting the common inlet temperature, direct acting, with $K_c = 60\,^{\circ}\mathrm{C}$ per unit and $\tau_I = 120\,\mathrm{min}$. It is slow because the analyzer answers every 25 min through a 3 min average and because the plant beneath it is an amplifier.

7.5.5 Model Predictive Control

The inlet layer can instead be given to the MPC, an option that is never the default, which moves the feed heater and the three quench valves together. Section 1.6.3 describes the option in general, and what follows is specific to this reactor. The conversion loop still sets T_{req}, the override of Equation (7.11) still cuts it to $T_{in,SP}$ on the run's own settings, and the trip and the stuck valve act as they do above the loops, so the MPC takes the place of the four inlet loops and of nothing else. Every $T_s = 0.25\,\mathrm{min}$, five controller scans, it solves a quadratic program over $N = 40$ steps, a horizon of 10 min, for the heater command u_H and the quench commands $u_{Q,j}$ and keeps only its first move, held until the decision five scans later, as constrained receding-horizon control does [12].

The controller's model is the reactor of Section 7.3 written again inside the controller with the declared values of Table 7.2, the design feed $\dot{m}_0 = 250\,\mathrm{t/h}$ and unit activity, never the run's parameters, its feed or its activity, as shown in Equation (7.13).

$$\begin{aligned} \tau_h \frac{\mathrm{d}\hat{T}_h}{\mathrm{d}t} &= u_H + d_H - \hat{T}_h, \quad \tau_q \frac{\mathrm{d}\hat{q}_j}{\mathrm{d}t} = u_{Q,j} + d_{Q,j} - \hat{q}_j \\ \tau \frac{\mathrm{d}\hat{T}_{out,i}}{\mathrm{d}t} &= \hat{T}_{in,i} + K\hat{x}_i - \hat{T}_{out,i}, \quad \tau \frac{\mathrm{d}\hat{f}_i}{\mathrm{d}t} = \hat{T}_{in,i} - \hat{f}_i \end{aligned} \tag{7.13}$$

where the hat denotes the controller's own copy of a quantity, $j = 1, 2$ and 3 counts the quench valves, $\hat{T}_{in,1} = \hat{T}_h$, $\hat{T}_{in,i+1}$ follows from $\hat{T}_{out,i}$ and $\hat{q}_i$ by Equation (7.5) at the design feed, $\hat{x}_i$ is Equation (7.2)

at the design feed with the activity replaced by an apparent activity $\hat{a}_i$ of bed i, and d_H in °C and $d_{Q,j}$ in t/h are offsets on the actuator paths. The state $\hat{f}_i$ is the inlet seen through the bed's own lag, so $\Delta\hat{T}_{c,i} = \hat{T}_{out,i} - \hat{f}_i$ is the model's copy of the compensated rise of Equation (7.12). Four further states carry the inlets through the 0.15 min thermocouple filter, updated once per scan as the loops' filter is, which makes 16 states, and the eight measurements are those four filtered inlets and the four bed exits as the override reads them.

The model is written from the balances rather than identified by a step test, because at the design point bed 1 amplifies its inlet 3.33 times and sits 3.8 °C below its fold, so an open-loop step identifies a plant that is leaving its operating point. Driven open loop by the commands the shipped loops sent through common setpoint steps of ±2 °C, it reproduces the four filtered inlets within 6.6×10^{-6} °C and the four exits within 1.9×10^{-6} °C over 60 min on the declared plant. Linearized once at the design point it would miss the same responses by up to 44 % RMS and by 76 % at its largest error, on the bed 4 exit, so each plan relinearizes the model about the present estimate, by central differences, with an exact zero-order hold over T_s and the present rates as an affine term.

The estimator is a steady-state Kalman filter whose gain is computed once on the model linearized at the design point, sampled at the 0.05 min scan and augmented with the eight disturbances d_H, $d_{Q,j}$ and $\hat{a}_i$ as random walks. At every scan, whichever structure has the valves, the nonlinear model is integrated with the commands the controller sent and the gain corrects the 16 states and the eight disturbances from the eight thermocouple residuals, so selecting the MPC in the middle of a run finds its estimate up to date. Disturbances estimated on the model's inputs in this way are what make a predictive controller offset-free [13]. An output bias, the simpler choice, shifts the predictions of a model that keeps its design index of 0.538 on bed 1. After a feed cut or a more active catalyst it would therefore predict the design amplification of a bed whose true index has risen toward one, and on a bed past its fold a constant offset cannot follow a mode that diverges. The apparent activities carry the higher index into the relinearized model instead. The MPC is given no feed measurement, and the cut of scenario 11 to 220 t/h is read as an apparent activity of 1.136, which is 250/220, on every bed, moving the slowest pole of the planned model from -0.132 to $-0.109\,\mathrm{min}^{-1}$. A stuck valve is not reported to it either, and its unanswered requests appear as a growing quench offset. A bed converting less than 1 % holds its activity estimate, a quench offset is held where the model's flow would turn negative, the disturbances are frozen while the reactor is tripped, and after a reset the states are put back on the thermocouples with the disturbances at the declared plant.

Each decision minimizes the objective of Equation (7.14) over moves held in eight blocks, two of one step, two of two, then 4, 6, 9 and 15 steps, taken as changes from the commands applied now, 32 moves and eight slacks, a 40-variable program on the 12 states that reach a planned output.

$$
\begin{aligned}
\min_{\Delta u,\,\varepsilon} \; & \sum_{k=1}^{N}\sum_{i=1}^{4}\left(\hat{T}_{in,i,k} - r_i\right)^2 + 0.1N\sum_{i=1}^{4}\left(\hat{T}_{out,i,N} - T^{*}_{out,i}\right)^2 \\
& + \lambda\sum_{b=1}^{8}\left(\Delta u_{H,b}^2 + G_Q^2\sum_{j=1}^{3}\Delta u_{Q,j,b}^2\right) + \sum_{\ell=1}^{8}\rho_\ell\left(\varepsilon_\ell^2 + \varepsilon_\ell\right) \\
\text{s.t. } & 300 \le u_H \le 420, \quad 0 \le u_{Q,j} \le 30, \quad \varepsilon_\ell \ge 0 \\
& \hat{T}_{in,i,k} \le T_{in,SP} + \delta + \phi_{i,k} + \varepsilon_i \\
& \Delta\hat{T}_{c,i,k} \le \Delta T_{ceil} + \psi_{i,k} + \varepsilon_{4+i}
\end{aligned}
\tag{7.14}
$$

where k counts the steps and b the move blocks, the first move being taken from the command in force, r_i is the inlet target of bed i and $T^{*}_{out,i}$ the exit it settles to, $\lambda = 0.1$ per °C^2 is the move price, $G_Q = 6.5$ °C per t/h scales a quench move to the bed 2 inlet change it makes at design, the bounds are in °C and t/h, ε_ℓ are the slacks with prices ρ_ℓ, one slack per soft row over the whole horizon, and $\delta = 0.5$ °C. The funnels $\phi_{i,k} = [\hat{T}_{in,i,0} - T_{in,SP} - \delta]^{+}\mathrm{e}^{-kT_s/\tau_h}$ and $\psi_{i,k} = [\Delta\hat{T}_{c,i,0} - \Delta T_{ceil}]^{+}\mathrm{e}^{-kT_s/\tau}$, with $[\cdot]^{+} = \max(0,\cdot)$ and a subscript 0 for the model's present value, bring back a value already above

its limit. Tracking carries 1 per °C^2 at every step. The terminal weight of $0.1N$ per °C^2 on the exits keeps a plan from ending with a slow bed mode still away from its steady value, and the horizon's 10 min exceed the time constant of the slowest mode, $1/0.132 = 7.6$ min. The move price and the horizon were chosen once on scenario 11 and are never retuned per case.

The bounds on the commands are hard and are the limits the inlet loops clamp to. The ceiling is $\Delta T_{ceil} = \Delta T_{lim} - 0.5\,°\mathrm{C}$, where ΔT_{lim} is the override's threshold of Equation (7.11), the run's rise alarm less its override margin, so ΔT_{ceil} is 27.5 °C as shipped, half a degree inside the 28 °C at which the override starts on the same compensated rise. While the MPC holds the ceiling the override has nothing to do and remains the backup. Each rise row is soft, priced 10^3 per °C^2 and 10^3 per °C against 1 per °C^2 of tracking. Each inlet is capped softly at the setpoint after the override plus δ, so a plan never undoes a cut the override made. The caps on beds 2 to 4 carry the prices of the rise rows, and the cap on bed 1 carries 10^5. That separate price came from a restart. At 10^3, a warm-up after a trip reset traded bed 1's inlet for the cold beds below it, which are warmed only through the heater, and took that inlet to 384.0 °C at 220 t/h, 2.1 °C past the 381.87 °C fold of that feed, for 5.3 min, and to 392.6 °C with the stuck valve of scenario 6. At 10^5 the inlet peaks at 381.50 °C in all four restart cases measured, with no time above the fold, where the loops peak at 381.00 °C.

The targets are made consistent with the ceiling before each plan. Bed by bed down the reactor, r_i is the common setpoint after the override or, where lower, the inlet at which the model's steady rise at the present activity estimate equals ΔT_{ceil}, which follows from Equation (7.2) in closed form as Equation (7.15), and $T^{*}_{out,i}$ is the cool steady exit at r_i.

$$r_i = \min\left\{ T_{in,SP},\ \left[\frac{1}{T_{ref}} - \frac{R}{E}\ln\frac{\Delta T_{ceil}}{K k_{ref}\hat{a}_i R_i}\right]^{-1} - \frac{\Delta T_{ceil}}{2} \right\} \tag{7.15}$$

The bracket, in kelvin, is the mean bed temperature at which $K\hat{x}_i = \Delta T_{ceil}$, converted to °C before half the ceiling is subtracted, and R_i is the model's unconverted fraction at the targets of the beds above. After the cut of scenario 11, with $\hat{a}_1 = 1.136$, it gives 379.50 °C for bed 1.

Each condensed program goes to a dual active-set solver [14] with a ridge of $10^{-9}\left(1 + 2(100N + 0.1N + G_Q^2\lambda)\right)$ on its Hessian. At 16 min of scenario 11, with no limit row active, the shipped plan's first move equals a batch least-squares solution of the same problem to a relative 1.0×10^{-14}. A plan is kept only when the solver reports an optimum, its first move, whole plan, predicted outputs, objective value and slacks are all finite, no slack is negative and every move lies within its bound. A plan that fails is marked in the historian, and the four inlet loops, which track the applied commands at every scan with no preload, take the heater and the valves until the next valid plan. With the noise off, two rejections forced at 40 min of scenario 11 move the heater by less than 0.0001 °C and no quench valve by more than 0.0001 t/h on the hand-over scan, and none of the plans in the runs reported here was rejected. The first plan is solved on the priming call, so a run started under the MPC is under its control from $t = 0$, and while the reactor is tripped the MPC is idle. In Chrome a plan takes 0.95 to 1.20 ms, 0.38 to 0.48 % of a single core at the 1 min/s rung and 23 to 29 % at the top speed of 60 min/s, at which the page still draws 59.9 to 60.1 frames a second.

Three models are thus kept apart, the plant of Section 7.3 on the run's parameters, feed, activity and valve state, the instruments the loops also read, and Equation (7.13) on the declared values with eight estimated disturbances standing in for everything it is not told.

7.6 Results

Twelve scenarios run from a manual amplification measurement to a graded challenge. Each carries a completion rule that credits the listed actions and the state the simulation then reaches (Section 1.6.4), and the Exercises at the end of this chapter work through several of them. The challenge is judged on the case as given, so changing a model parameter or a disturbance voids the attempt.

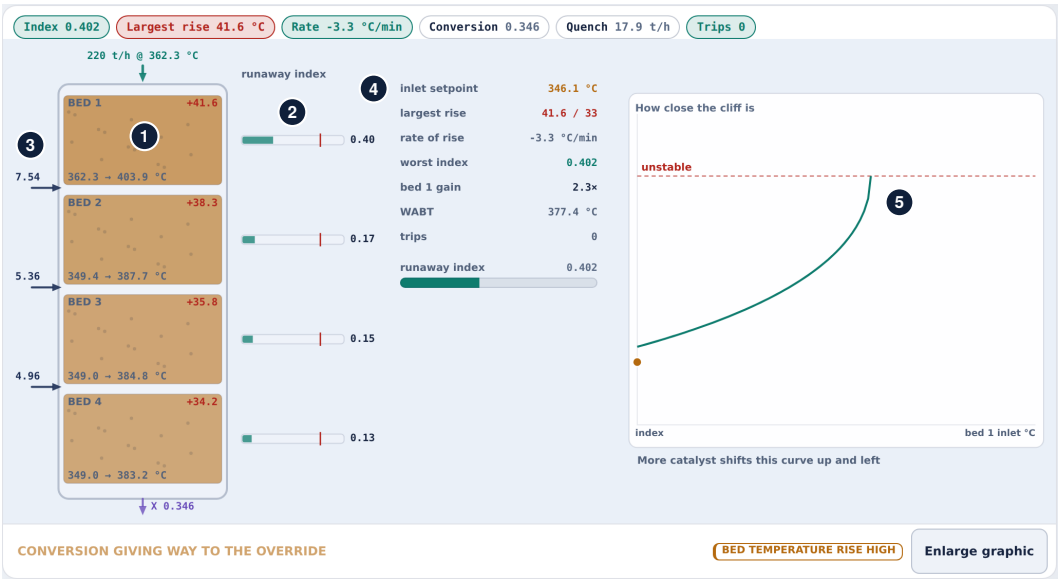


Figure 7.2 The hydrocracker application running scenario 8, the override measurement case, at 28.2 min with the conversion loop in automatic and the override reading the raw bed temperature rise, 8.2 min after a twelve percent feed cut and 0.05 min before the trip, cropped to the plant view with numbered callouts. (1) The reactor is drawn in elevation with its four beds, each shaded by its temperature rise and (2) carrying a runaway index bar against a line at one, and (3) with the quench flows between the beds. (4) The panel shows the inlet setpoint pulled down to 346.1 °C by the override, the largest rise at 41.6 °C against its 33 °C alarm, a rate of rise of −3.3 °C/min and a worst runaway index of 0.402. (5) The inset plots bed 1 runaway index against inlet temperature, solved from the model at the present feed and activity and ending at the fold (381.9 °C at 220 t/h), with the present operating point held at its left edge.

Table 7.3 Bed 1 Runaway Index and Rise Against Inlet Temperature

Inlet (°C)	r	Rise (°C)
380	0.493	21.7
381	0.538	23.9
382	0.592	26.4
383	0.662	29.8
384	0.763	34.7
384.5	0.848	38.9
384.8	0.961	44.4
385	no cool steady state	—

7.6.1 The Index and the Boundary

At the design condition the first bed sits at $r = 0.538$ and amplifies small inlet changes by $(1 + r)/(1 - r) = 3.33$, which the steady-state response to a ±0.5 °C inlet change reproduces to three figures. A whole degree moves it further. Stepping the heater from 381 to 382 °C in scenario 1 moves the bed 1 outlet from 404.86 to 408.45 °C in 120 min, a gain of 3.59, because the index rises with the step. Table 7.3 reports the index on the cool branch against inlet temperature.

The steps are not equal and they accelerate, which is what an exponential rate law does to a linear temperature increment. The last 0.8 °C, from 384 to the fold at 384.8 °C, moves the index

Table 7.4 Override Reading a Raw and a Compensated Bed Rise, on a Feed Cut From 250 t/h to 220 t/h at 20 min

Measurement	Engages (min)	Setpoint low (°C)	Peak rise (°C)	Outcome
Raw ΔT	25.8	344.7	42.2	trips at 28.25 min
Compensated ΔT	26.2	377.4	29.9	holds

almost as far as the three degrees before it (0.198 against 0.225). At the fold the cool steady state and the unstable one merge and vanish, and the only steady state left is an ignited bed held at the 0.9 conversion ceiling with a rise of 81 °C, far past every trip setting. The same fold is reached at a catalyst activity of 1.180 and at a feed rate of 212 t/h.

7.6.2 The Override That Trips the Reactor

Table 7.4 reports scenario 8, a twelve percent feed cut at 20 min with the conversion loop in automatic and the override armed, reading either the raw bed rise or the compensated one. The application gives one run whether the compensation is chosen before Run or carried through a Reset.

On the raw rise the override does nothing until the rise reaches its 28 °C threshold at 25.8 min, then takes the inlet setpoint from 380.2 to 344.7 °C in 2.5 min and trips the reactor on its 42 °C bed rise setting at 28.25 min, 8.25 min after the cut, and the state just before the trip is shown in Figure 7.2. Nothing is broken. The override is correctly signed for the steady state, the gain is reasonable and the measurement is accurate. It fails because the transient sign of the measurement is the opposite of its steady-state sign, so the loop cuts, the number gets worse, and it cuts harder. With the compensated measurement the same override engages at 26.2 min and never cuts more than 2.0 °C, the conversion loop and the override together take the setpoint no lower than 377.4 °C, and the largest rise peaks at 29.9 °C, far from a trip.

> **TRY IT IN THE APP** **8 · The number the override reads**
>
> Load the case and press Run. The feed is cut to 220 t/h at 20 min. Watch the Bed inlet setpoint trace. The override does nothing until 25.8 min, then takes the setpoint from 380.2 toward 344.7 °C in 2.5 min, and the reactor trips at 28.25 min. Press Reset, set Bed rise compensation to Compensated and run again. The rise peaks at 29.9 °C and nothing trips.

7.6.3 Bed 1 Has No Handle

With all four inlet loops in automatic and well tuned and the override Off, as scenario 4 ships, an activity step to 1.30 at 15 min puts bed 1 past the boundary and no inlet loop can save it, because bed 1's only handle is the feed heater with its 2.5 min lag. In scenario 4 the bed 1 index passes one at 26.5 min and peaks at 1.019, and the reactor trips on bed temperature rise at 27.3 min. The quench valves below protect beds 2, 3 and 4 from what bed 1 does, and they do, and with the override Off nothing protects bed 1 from itself. That is the structural reason a hydrocracker trips rather than controlling its way out of a real runaway. With the override On, a setting the scenario does not ship, the same loops do not trip, because the override pulls the common setpoint down to 374.3 °C and bed 1's index peaks at 0.769, at the price of 102.8 min with the compensated rise above 28 °C and a mean conversion of 0.5486. The override's index peaks at 0.692 while the true one passes 0.75 at 16.8 min, so the rise term alone cuts, from 17.25 min. The MPC of Section 7.5.5 holds the same step at a compensated rise of 27.52 °C, as Section 7.6.4 reports.

Pause. Try Exercise 7.3 before reading on.

Table 7.5 Inlet Loops Against the MPC, Default Seed, Full Duration of Each Case

Case	Structure	Rise, comp./raw (°C)	Over 28 (min)	Override (min)	Mean X	IAE_X (min)	Inlet IAE (°C min)	Heater (°C)	Outcome
11, 120 min feed cut	MPC (override On or Off)	27.50/27.50	0.0	0.0	0.5501	4.57	159.0	6.9	480 plans, 0 rejected
	loops and override	28.80/29.92	98.9	99.0	0.5427	3.68	6.7	4.9	—
	loops, override Off	34.05/34.05	98.9	0.0	0.5683	6.75	3.6	0.0	76.0 min over the alarm
	loops, bed 1 at 379.496 °C by hand	27.50/27.50	0.0	0.0	0.5501	4.57	158.2	6.0	—
	hindsight, none over 28[a]	27.72/28.48	0.0	102.3	0.5369	2.99	3.0	2.8	—
	highest X of the 48[b]	29.99/30.44	98.9	96.8	0.5503	4.59	2.5	1.7	—
10, 600 min conversion to 0.600	loops and override	29.28/29.80	512.3	514.3	0.5470	21.27	30.9	36.5	—
	MPC (override On or Off)	27.50/27.53	0.0	0.0	0.5659	9.90	1465.7	9.4	2400 plans, 0 rejected
	loops, MPC's bed 1 target	27.50/27.50	0.0	0.0	0.5659	9.92	—	6.5	—
	hindsight, least IAE_X[c]	31.13/31.33	496.9	415.1	0.5554	16.23	221.9	15.5	—
6, 180 min stuck valve	loops and override	28.37/28.88	149.6	149.8	0.5589	8.44	645.0	2.2	—
	MPC (override On or Off)	27.05/29.12	0.0	0.0	0.5415	5.32	316.7	40.2	720 plans, 0 rejected
	hindsight, least inlet IAE[d]	26.32/28.23	0.0	156.1	0.5374	4.57	359.4	16.0	—
	highest X, none over 28[e]	27.60/27.95	0.0	156.1	0.5463	6.18	524.6	1.2	—
4, 120 min activity 1.30	loops, override Off (shipped)	41.90/41.96	10.0	0.0	0.1362	47.92	4.9	81.0	trip, 27.25 min
	loops, override On	30.67/34.43	102.8	102.8	0.5486	4.39	18.4	17.1	—
	MPC (override On or Off)	27.52/31.30	0.0	0.0	0.5809	8.27	471.0	22.7	480 plans, 0 rejected
	loops, MPC's bed 1 target	29.11/30.65	19.3	0.0	0.5814	8.33	—	16.0	—
7, 300 min feed cut, X loop	loops and override	28.80/29.90	45.2	48.2	0.5244	5.48	78.4	48.4	—
	MPC (override On or Off)	27.32/27.97	0.0	0.0	0.5242	5.80	140.0	41.6	1200 plans, 0 rejected
12, 400 min challenge	loops, suggested	29.84/32.49	33.3	38.5	0.5284	7.09	196.4	414.6	passes, 0.61 °C
	MPC, rest as shipped	28.06/31.72	0.3	0.0	0.5741	24.83	1391.4	346.2	fails, 5.25 °C, 1600 plans
	MPC, suggested	27.55/29.54	0.0	0.0	0.5284	9.48	471.8	235.9	fails, 2.43 °C, 1600 plans

Rise is the largest compensated and raw bed rise before any trip, over 28 is the time the compensated rise spends above the shipped override threshold, IAE_X is the conversion against its setpoint, inlet IAE sums the four inlets against the common setpoint after the override, and heater is the travel of the heater command. The challenge outcome is the worst inlet error in the last third, and no MPC plan was rejected. The 48 hindsight settings are four tunings (cautious, with $K_c = 0.8$ and $\tau_I = 8$ min on the heater and 0.25 t/h per degree and 6 min with no derivative on the quench, shipped, shipped without the quench derivative, and fast, with 4 and 2 min on the heater and 1.5 t/h per degree and 1 min on the quench) by K_o of 1, 2.5, 5 and 8 by margins of 4, 5 and 7 °C. [a]Fast, $K_o = 1$, margin 7. [b]Fast, $K_o = 1$, margin 4. [c]Cautious, $K_o = 1$, margin 4. [d]Fast, $K_o = 8$, margin 7. [e]Cautious, $K_o = 1$, margin 7.

7.6.4 One Bed at Its Limit

Scenario 11 is the feed cut of scenario 5 with the shipped tuning and the override armed, loaded with the MPC on the inlets, and it is compared with the same run repeated after Reset under the inlet loops. Table 7.5 measures scenarios 11, 10, 6, 4 and 7 and the challenge under both structures, the MPC with the settings scenario 11 ships and no retuning, every row sharing one seed, one set of instruments and noise, and one duration. The rows that locate the gain are the override Off, the loops with bed 1's setpoint alone lowered by hand on a copy of the controller, which the shipped structure cannot do, the loops given the MPC's own bed 1 target as it moved, and the best of 48 fixed loop and override settings chosen in hindsight inside the application's ranges. Figure 7.4 draws three columns of the table for the two structures.

Under the MPC the largest compensated rise climbs to the 27.5 °C ceiling and holds there, and the override never acts. Only bed 1 gives up temperature. From 17.5 min its inlet sits at 379.5 °C while beds 2 to 4 return to 381 °C, and the conversion averages 0.5501 over the two hours, 245.9 t converted. Under the loops the override engages at 21.1 min, the compensated rise peaks at 28.80 °C at 24.1 min and settles at 28.49 °C, above the override's own threshold because the cut is proportional, all four inlets are pulled down together to 379.79 °C, and the conversion averages 0.5427 (242.6 t), as shown in Figure 7.3.

The gain is the per-bed target. The rise row is active in none of the 480 plans of scenario 11, because bed 1's target already places its predicted steady rise on the ceiling, and the loops with bed 1's setpoint alone lowered by hand to 379.496 °C at the cut give the MPC's result exactly, a largest rise of 27.50 °C, no override action, a mean conversion of 0.5501 and 245.9 t, with the override On or

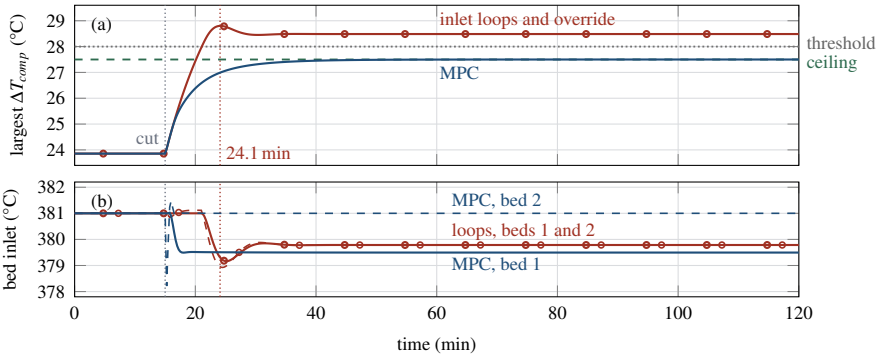


Figure 7.3 Scenario 11, the feed cut from 250 to 220 t/h at 15 min, under the MPC as loaded and, after Reset, under the four inlet loops, default seed. The loops' curves carry open circles and the MPC's carry no marker. (a) The largest compensated bed rise against the MPC's 27.5 °C ceiling and the override's 28 °C threshold. The loops' rise peaks at 28.80 °C at 24.1 min, with the override cutting all four inlets toward 379.0 °C. (b) The bed 1 inlets (solid) and bed 2 inlets (dashed). The MPC holds bed 1 alone at 379.5 °C from 17.5 min while bed 2 returns to 381 °C, and the loops hold all four inlets on the one cut setpoint.

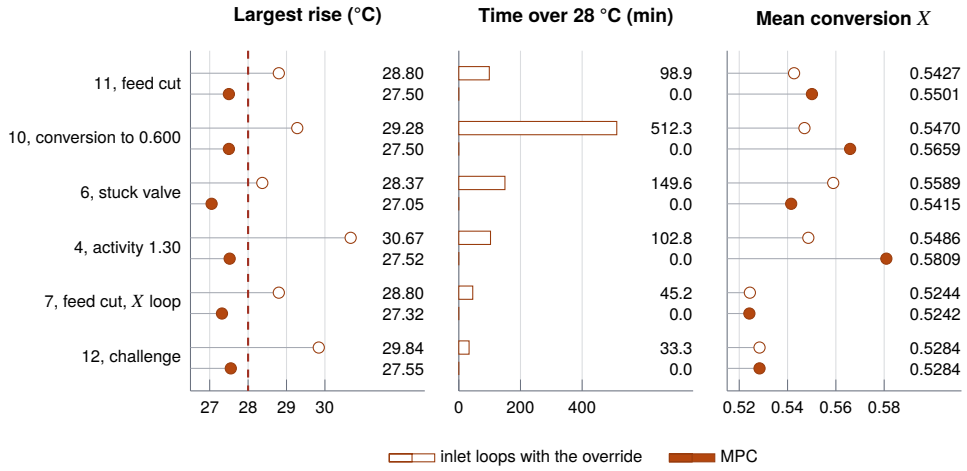


Figure 7.4 Three columns of Table 7.5 for its six cases, the inlet loops with the override as open circles and white bars and the MPC as filled circles and solid bars. For case 4 the loops' row is the one with the override On, and for case 12 both rows use the suggested settings. The dashed line is the override's 28 °C threshold. Notice that every MPC point lies left of that line and spends no time over it, while its conversion is the higher one in cases 11, 10 and 4 and the lower one in case 6.

Off. What the MPC contributes in this case is that it finds that target by itself, from the thermocouples and without the feed, while the shipped structure has one setpoint for four beds and no place to hold one of them lower. The override, with only that setpoint to move, cools all four beds to cool one. No fixed setting of the shipped structure closes the gap. The best of the 48 with no time over 28 °C converts 0.5369, and the one that matches the MPC's conversion, 0.5503, holds the rise above 28 °C for 98.9 min.

Scenario 10 gives the same attribution with the conversion loop above both structures. The MPC more than halves the conversion IAE, 9.90 against 21.27, with no time above 28 °C against 512.3 min, but the loops given the MPC's bed 1 target as it moved reach 9.92, and the best fixed setting reaches

16.23 with 496.9 min above 28 °C.

The plan itself counts on a fast event. Through the activity step of scenario 4 the rise row is active in 10.2 % of the plans, from 15.8 to 27.8 min, and the MPC holds the compensated rise at 27.52 °C with a peak index of 0.716, where the loops following the MPC's own bed 1 target avoid the trip but reach 29.11 °C with 19.3 min above 28 °C.

The losses each have a mechanism. Scenario 6 sticks quench valve 1 at 15 min and raises the activity to 1.10 at 20 min. The MPC is not told that the valve is stuck, so it raises that valve's request at every plan while its estimator reads each unanswered request as a larger offset, and when the request reaches its 30 t/h bound at 59.8 min the plan turns to the heater at once. One decision moves the heater command down 13.08 °C, bed 1's inlet falls to 377.59 °C, and the heater travels 40.2 °C against 2.2 °C under the loops. The MPC halves the inlet IAE, 316.7 against 645.0 °C min, and keeps the rise below 28 °C, but converts 0.5415 against 0.5589, and a fixed setting with no time over 28 °C converts 0.5463. No rise row binds in scenario 6, so the pulse is not the ceiling's doing. The likely cause is the move price of 0.1, chosen on a slow cut and applied with a one-step first block to a heater that has a 2.5 min lag and no rate limit in the model. In scenario 7, with the conversion loop's 25 min analyzer above both structures, the MPC's conversion IAE is 5.80 against 5.48, at a mean conversion of 0.5242 against 0.5244. Its inlet IAE against the common setpoint is larger in every case, 159.0 against 6.7 °C min in scenario 11, because bed 1 is held below that setpoint by design. The cheap move price shows again on a switch to the MPC in the middle of an upset, which with the noise off moves the heater command 11.10 °C on the next scan at 30 min of scenario 11 and 25.59 °C at 70 min of scenario 6.

The challenge's rule is the same under the MPC, and the MPC does not meet its target. With the suggested settings the reactor never trips, the compensated rise stays at or below 27.55 °C and the peak index is 0.669, but the worst inlet error in the last third is 2.43 °C against the ±1 °C band, because the MPC holds bed 1 where its rise meets the ceiling. With everything else as shipped, a 382 °C start with no override and no conversion loop, it again never trips, with a peak index of 0.685 and a worst inlet error of 5.25 °C, where the shipped loops trip at 150.0 min. Neither MPC run breaks an unscored limit, with no trip, no time over the 33 °C alarm and every index below one, and the verdicts hold at half the physics step and at seeds 1 and 777.

> **✦ DID YOU KNOW? The Predictive Controller Fails the Challenge**
>
> From the simulator (Table 7.5). With the suggested settings the four inlet loops pass the graded challenge with a worst inlet error of 0.61 °C. The predictive controller never trips and keeps the compensated rise at or below 27.55 °C, yet it fails at 2.43 °C, because it holds bed 1 below the common setpoint on purpose.

7.6.5 The Graded Challenge

The challenge runs 400 min with catalyst activity steps to 1.08 at 40 min and 1.13 at 260 min and a feed cut to 225 t/h at 140 min, starting a degree above the design inlet temperature with cautious tuning, no override, no conversion loop, 0.15 °C of thermocouple noise and 0.8 % of analyzer noise. It is met when every bed inlet is held within 1 °C of setpoint through the last third of the run, no runaway index reaches one, and the reactor does not trip. As shipped the reactor trips on bed temperature rise at 150.0 min, ten minutes after the feed cut. The suggested configuration arms the override with an index limit of 0.72, closes the conversion loop on the design target and applies the tuning of Section 7.5, and it passes with a worst inlet error of 0.61 °C, inside the band by 0.39 °C, and a peak index of 0.724, whether the settings are applied before Run or carried through a Reset. The override is what decides it, because the same settings with the override Off trip at 145.4 min, while armed it carries even the 382 °C start through with a peak index of 0.718. The heater tuning matters next. With the override armed and the cautious heater tuning left in place, bed 1's inlet is still 1.28 °C off in the last third, and the suggested heater tuning brings it to 0.33 °C. Over five other noise seeds the

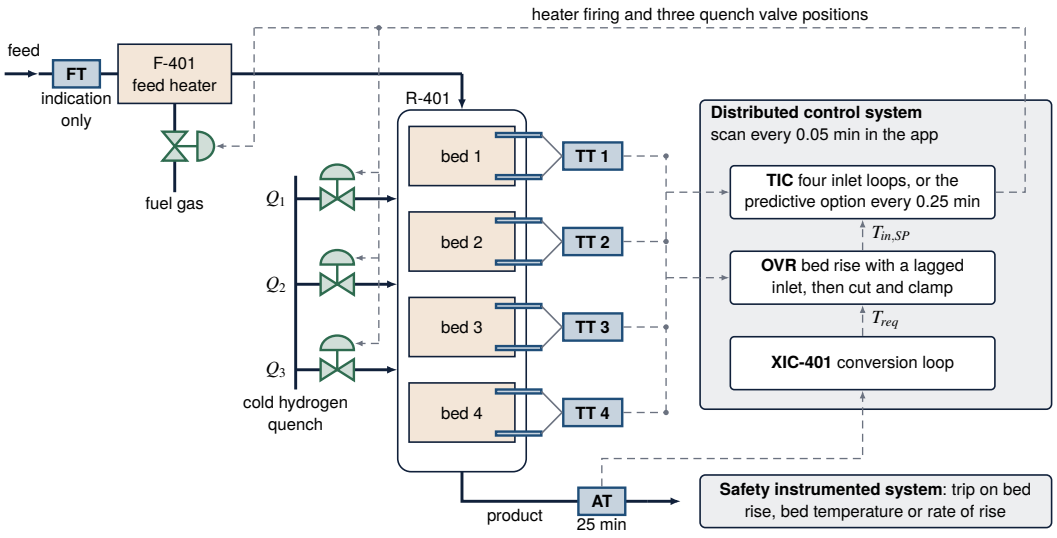


Figure 7.5 The loops of the hydrocracker reactor with their hardware. Lettered boxes are sensing devices: the thermocouples and transmitters at the inlet and outlet of each bed (TT), the conversion analyzer (AT) and the feed flow transmitter (FT). Valve symbols are final elements and dashed arrows are signals. Notice that every bed below the first has a valve of its own, while bed 1 is reached only through the feed heater, and that the override writes nothing but the common setpoint of the inlet loops. The trip sits outside the control system. Not to scale, and not a piping or safety diagram.

shipped case trips between 149.5 and 150.2 min and the suggested one passes with worst errors of 0.50 to 0.65 °C.

7.7 Technology in Practice

Figure 7.5 draws the loops of Section 7.5 with the hardware a refinery installs.

What Measures It

Every bed is read by thermocouples at its inlet and outlet. In the reactor of a 1997 investigation the center inlet thermocouple of each lower bed fed the controller of its quench flow, and the bed outlets were displayed but regulated by nothing [1]. The bed rise that the override reads is a difference of two such sensors, which Equation (7.12) puts on one clock. Conversion on a real unit is inferred from a distillation, for which the analyzer AI-401 stands. A soft sensor, a model that predicts from fast measurements a quality known only at a low sampling rate, is the usual supplement in refinery processes [15].

> **TECHNOLOGY** **Sensor: Thermocouple Array in the Catalyst Beds**
>
> Two wires of different metals joined at one end produce a voltage roughly proportional to temperature. Thermocouples are less accurate than resistance thermometers but more rugged and wider in span [16]. One hydrocracker reactor carried 96 flexible sheathed thermocouples inside its beds, at the inlet, middle and outlet of each, to show axial and radial gradients [1]. The most common failure is an open circuit, which can read as a wildly varying temperature, so the instrument is set to drive its reading high or low, whichever is safer [16]. In that reactor a failed or over-range point was sent as zero, and operators did not understand that zero could mean an extremely high temperature [1]. On a trend such a point drops to zero while its neighbors climb. The 0.15 min thermocouple filter of Table 7.2 is the model's account of this sensor.

What Moves It

Three quench valves and the firing of the feed heater are the handles. In the unit of the investigation the reactor inlet temperature controller set the outlet temperature of a trim furnace, whose controller set the fuel gas pressure [1]. The model folds that cascade into the single lag τ_h. A sudden opening of one quench valve there made the hydrogen flow to the furnaces fluctuate, and both streams came from the recycle gas [1]. The model's quench supply is fixed.

> **TECHNOLOGY** **Final Element: Quench Control Valve**
>
> In a pneumatic control valve the stem position follows the air pressure on a diaphragm or piston against a spring [16]. Friction in the stem packing and the force of the pressure drop on the plug disturb that balance, so a positioner compares the stem position with the control signal and adjusts the air until they agree. With excess friction the stem moves in jerks, or not at all for a small change of signal, which is called stiction. The failure position follows the hazard: a valve that must open on loss of air is built air-to-close [16]. In the hydrocracker of the 1997 report the quench valves leaked when shut, and an alarm warned when one stood more than half open [1]. The model's valve is a 0.30 min lag, tight and free of friction, until scenario 6 sticks one.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisor sends set points to distributed controllers that act through actuators from sensor feedback, with cycle times from milliseconds to minutes [17]. The app scans its loops every 0.05 min and plans its predictive controller every 0.25 min, inside that range. The four inlet loops and the conversion loop are PID blocks. The override is a calculation block, the lag of Equation (7.12) followed by the cut and the clamp of Equation (7.11). Conventional constraint control is built in this way from PID algorithms, lead-lag blocks and select logic, and a predictive controller replaces the combination [5]. The trip belongs to a safety instrumented system of sensors, logic solvers and final elements, designed to IEC 61511 and often independent of the control system [17].

> **IN PRACTICE** **A Limit Is a Shutdown, Not a Target**
>
> What the model's trip leaves out is the decision to use it. The reactor of the 1997 report had written limits, among them a maximum reactor temperature of 800 °F, and two depressuring systems, of 100 and 300 psi per minute [1]. In earlier excursions the operators had answered with more quench, a lower inlet temperature or a feed stop, and most could recall one depressuring. The investigators found that quench is typically sized for minor excursions. Past the point where it can hold a bed, only a lower hydrogen pressure slows the reaction. Their advice was limits that are not flexible, emergency systems that are tested, and emergency alarms that stand apart from operating alarms [1].

What Is Modernizing It

The first development is denser thermometry tied to the safety system. A 2019 account of a hydrotreating reactor revamp describes flexible thermocouples in place of two vertical probes near the wall, 20 to 50 sensors in a reactor for radial and axial coverage, and a trip that opens the emergency depressuring valve when a sensor reads too high. A faulty signal is recognized by its current range, so that it does not trip the unit [18]. The second is the valve. An electronic positioner measures stem position and actuator pressure, and the plot of one against the other, the valve signature, shows excessive packing friction [16]. In scenario 6 the valve sticks with no such warning. The third is control from the feed. A 2022 study sets the reactor temperatures of an industrial hydrocracker by a feedforward predictive controller, because feedback has a long response time [11]. The app's predictive controller is given no feed measurement.

7.8 Limitations

The beds are lumped stages, so there is no axial profile and no hot spot position. A real bed's peak temperature is above its outlet and the difference matters at the limits.

One lumped reaction with one activation energy represents a network whose lumping is a research subject in its own right [6, 7]. The application therefore gets the thermal behavior right and the product slate not at all.

Quench maldistribution is absent, which removes the mechanism most often responsible for real hot spots [8].

Catalyst deactivation over a twelve-month cycle is represented only by an activity slider, so the slow walk of the weighted average bed temperature that dominates real hydrocracker operation cannot be studied.

The trip is instantaneous and does not cut the feed. A real depressuring takes minutes at a controlled rate and the reaction continues while it happens, which is precisely the period a safety analysis is about.

The override is a proportional cut, so it holds the largest rise above its 28 °C threshold by the conversion loop's standing cut divided by K_o, 1.2 °C at the end of scenario 10.

The MPC's model is the simulator's own lumped beds on the declared values, exact on the declared plant, which no real reactor offers, and on a changed plant its accuracy rests on the estimator, with up to 1.63 °C of open-loop error at an E/R of 21 000 K. It is told neither the feed nor a valve's position, its heater has no rate limit, which lets one decision move the heater command by 13.08 °C, and the separate price on bed 1's inlet cap rests on four restart cases.

Finally, the application implements the regulatory, override and trip layers, with the optional MPC in place of the inlet loops, and not the multivariable controller that sits above them on a real unit [5, 10]. Scenario 10 shows the economics running into the override and does not show an optimizer resolving the trade.

> **OPEN CHALLENGE** **Seeing a Hot Spot Before It Spreads**
>
> Each bed of the model has one temperature, and a real runaway can start at one point. In the 1997 accident a single outlet point jumped while the other four points of the same bed stayed normal, and the cause of the poor flow distribution could not be determined [1]. A 2021 review of reactor runaway finds that most studies neglect the flow patterns of the vessel, although these can weigh heavily on how a runaway develops, and that uncertain kinetic parameters affect how early a criterion warns [4]. It states that nothing guarantees a developing runaway is seen in every case. Progress would be a warning that holds under such uncertainty and reads the whole thermocouple array.

7.9 Discussion

Three results generalize past hydrocrackers.

The first is that stability can be a property of the plant and a computable one. The index $r = \frac{1}{2} K \, \mathrm{d}x/\mathrm{d}T$ is available from a heat of reaction and an activation energy before any control system is designed, and it says both whether the plant needs a controller to stand up and how much it amplifies when it does not. Any exothermic reactor with an exponential rate law has such a number and it is worth computing early.

The second is that an override is only as good as the sign of its measurement in the transient. The measured failure here takes the setpoint down 35.5 °C in 2.5 min and trips the unit 8.25 min after a feed cut, on an override that is correctly designed for the steady state. Any difference of two measurements with different dynamics deserves the same check, and the fix is usually one lag rather than a retuning.

The third is structural. Bed 1's only actuator is upstream of it behind a 2.5 min lag, so no tuning makes the control system fast enough once that bed is past its boundary. Asking of any fast instability

where the nearest actuator is, and how long it takes to reach the place the heat is being made, is a question that belongs in the process design rather than in the control design.

7.10 Conclusion

A hydrocracker bed is one of the few plants a control course meets that can be unstable without a controller, and the property is a number rather than an opinion. In the simulated reactor of this chapter the first bed sits at a runaway index of 0.538 and amplifies its inlet by 3.3, the index rises to 0.763 at 384 °C of inlet and the cool steady state vanishes at 384.8 °C, and the same fold is reached at a catalyst activity of 1.180 or a feed rate of 212 t/h. Cutting the feed from 250 t/h to 220 t/h raises the steady first bed rise from 23.9 °C to 34.1 °C, which is the opposite of what a production trend suggests. The override installed to protect the reactor trips it 8.25 min after that feed cut when it reads a raw bed temperature rise, because the inlet answers a quench valve within 0.30 min while the outlet carries a 3.5 min catalyst time constant and the difference is therefore inverted during the correction. One lag on the inlet, matched to the bed, holds the largest rise at 29.9 °C instead. An optional model predictive controller holds the same feed cut at its 27.5 °C ceiling with a mean conversion of 0.5501 against 0.5427, a gain the loops reproduce when bed 1's setpoint alone is lowered to 379.5 °C, and on a stuck quench valve it converts 0.5415 against 0.5589. Further work would add an axial bed profile so that the hot spot position becomes visible, quench maldistribution, a catalyst deactivation model over a full cycle, and a depressuring that takes the minutes a real one takes. The simulator is a place to learn what an override reads.

KEY POINTS

- A catalyst bed that makes its own heat can be unstable with no controller, and one number, the runaway index, says how close it is.
- Below an index of one a bed amplifies its inlet, 3.33 times at design, and the quench loops break the chain of four beds.
- Less feed is hotter: a cut from 250 to 220 t/h raises the first bed rise from 23.9 to 34.1 °C.
- An override is only as good as the sign of its measurement in the transient, and one matched lag cures the raw bed rise.
- Bed 1 has no fast handle, so past its boundary the answer is a trip.

Where else this applies. The same ignition boundary stands in the stirred tank of Volume 7, Chapter 2, a protective override that rests at its clamp guards the machine of Volume 4, Chapter 2, and a safety override for each zone bounds the throughput of Volume 4, Chapter 6.

References

[1] U.S. Environmental Protection Agency, "EPA chemical accident investigation report: Tosco Avon refinery, Martinez, California," U.S. Environmental Protection Agency, Chemical Emergency Preparedness and Prevention Office, Washington, DC, Tech. Rep. EPA 550-R-98-009, 1998.

[2] R. van Welsenaere and G. Froment, "Parametric sensitivity and runaway in fixed bed catalytic reactors," *Chemical Engineering Science*, vol. 25, no. 10, pp. 1503–1516, 1970.

[3] M. Morbidelli and A. Varma, "Parametric sensitivity and runaway in fixed-bed catalytic reactors," *Chemical Engineering Science*, vol. 41, no. 4, pp. 1063–1071, 1986.

[4] A. Kummer and T. Varga, "What do we know already about reactor runaway? a review," *Process Safety and Environmental Protection*, vol. 147, pp. 460–476, 2021.

[5] S. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[6] P. Balasubramanian and S. Pushpavanam, "Model discrimination in hydrocracking of vacuum gas oil using discrete lumped kinetics," *Fuel*, vol. 87, no. 8-9, pp. 1660–1672, 2008.

[7] S. Sadighi, A. Ahmad, and M. Shirvani, "Comparison of lumping approaches to predict the product yield in a dual bed vgo hydrocracker," *International Journal of Chemical Reactor Engineering*, vol. 9, no. 1, 2012.

[8] W. Strasser, "Cfd study of an evaporative trickle bed reactor: Mal-distribution and thermal runaway induced by feed disturbances," *Chemical Engineering Journal*, vol. 161, no. 1-2, pp. 257–268, 2010.

[9] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[10] C. Cutler and R. Hawkins, "Application of a large predictive multivariable controller to a hydrocracker second stage reactor," in *1988 American Control Conference*. IEEE, 1988, pp. 284–291.

[11] E. Iplik, I. Aslanidou, and K. Kyprianidis, "A feedforward model predictive controller for optimal hydrocracker operation," *Processes*, vol. 10, no. 12, p. 2583, 2022.

[12] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[13] K. R. Muske and T. A. Badgwell, "Disturbance modeling for offset-free linear model predictive control," *Journal of Process Control*, vol. 12, no. 5, pp. 617–632, 2002.

[14] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[15] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[16] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[17] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[18] P. K. Wagner-Gillen, "How reactive hazard assessment and new reactor internals improve safety at the Shell Rhineland refinery," *Chemical Engineering Transactions*, vol. 77, pp. 511–516, 2019.

7.11 Exercises

7.1 ⌕ Concept check · 2 min

The reactor runs at the design condition with every bed inlet held at 381 °C. A charge pump loses suction and the feed falls from 250 to 220 t/h, with the inlet setpoint left where it was. What happens to the steady rise and the runaway index of bed 1?

(a) Both fall, because less feed through the reactor releases less heat of reaction

(b) Both rise, because the feed spends longer on the catalyst and its heat warms less mass

(c) The rise is unchanged and the index falls, because the conversion per tonne is fixed

(d) The rise grows and the index falls, because the bed now reacts at a hotter mean temperature

7.2 ⌕ Concept check · 2 min

In scenario 8 the override reads the raw bed rise $T_{out} - T_{in}$, engages 5.8 min after a feed cut and then trips the reactor 2.45 min later on a bed rise it created. Which explanation is right?

(a) The override gain of 2.5 is too high, so the override oscillates and overshoots into the trip

(b) The override is wrongly signed for the steady state, so each cut raises the inlet temperature

(c) After a cut the inlet falls faster than the outlet, so for minutes each cut raises the rise it reads

(d) The 0.15 min thermocouple filter delays the rise, so the override acts too late to help

7.3 ⌕ Concept check · 1 min

In scenario 4 all four inlet loops are in Auto with the shipped tuning, the override is Off, and the catalyst activity steps to 1.30. Bed 1's index passes one and the reactor trips on its bed rise. Why can the loops not hold bed 1?

(a) The quench loops are tuned too slowly, with an integral time of 1.5 min
(b) The heater loop has the wrong action and drives bed 1's inlet up
(c) The conversion loop raises the common inlet setpoint as conversion falls
(d) Bed 1's only handle is the feed heater, upstream and behind a 2.5 min lag

7.4 Theory and calculation · 6 min

At the design condition bed 1 enters at 381.0 °C and rises 23.86 °C. Use $K = 90.0\,^{\circ}\mathrm{C}$ per unit of x, $E/R = 20\,000\,\mathrm{K}$, $\tau = 3.5\,\mathrm{min}$ and the Arrhenius form of the index, $r = KxE/(2RT^2)$ with T the mean bed temperature in kelvin.

(a) Compute the bed's conversion x, its mean temperature in kelvin and its runaway index r. *Checkpoint:* $x = 0.2651$, 666.08 K, $r = 0.538$.

(b) Compute the steady amplification from inlet to outlet, the time constant of the linearized bed and its pole. *Checkpoint:* 3.33, 7.6 min, $-0.132\,\mathrm{min}^{-1}$.

7.5 Theory and calculation · 8 min

Bed 1 leaves at 404.86 °C on 250 t/h of oil. Quench hydrogen enters at $T_q = 60\,^{\circ}\mathrm{C}$ with $c_{p,H_2}/c_{p,oil} = 5.5$, and bed 2's inlet is held at 381.0 °C. Use Equation (7.5).

(a) Compute the quench flow that holds bed 2's inlet at 381.0 °C. *Checkpoint:* 3.38 t/h.

(b) Differentiate Equation (7.5) with respect to $\dot{q}$ at that point. By how much does bed 2's inlet move per t/h of quench, and what extra quench lowers it by 2.0 °C? *Checkpoint:* −6.57 °C per t/h; 0.30 t/h.

(c) Treat that 2.0 °C cut as a step and hold the reaction rise of bed 2 constant. Using Equation (7.7), how far above the reaction rise does the raw rise read just after the cut and 3.5 min later? What does the compensated rise of Equation (7.12) read? *Checkpoint:* 2.0 °C, 0.74 °C; the reaction rise throughout.

7.6 Theory and calculation · 7 min

The MPC is not told the feed. After the cut of scenario 11 from 250 to 220 t/h it estimates an apparent activity $\hat{a}_1$ on bed 1 and sets the bed's inlet target by Equation (7.15), with $\Delta T_{ceil} = 27.5\,^{\circ}\mathrm{C}$, $K = 90.0\,^{\circ}\mathrm{C}$, $k_{ref} = 0.2600$, $R_1 = 1$, $E/R = 20\,000\,\mathrm{K}$ and $T_{ref} = 392.5\,^{\circ}\mathrm{C}$.

(a) Using Equation (7.2), explain why the cut looks to the estimator like a change of activity, and compute $\hat{a}_1$. *Checkpoint:* $250/220 = 1.136$.

(b) Compute the inlet target r_1 for bed 1. *Checkpoint:* 379.50 °C.

(c) Compute the runaway index of bed 1 at that target, where its rise is 27.5 °C, and the bed's amplification. *Checkpoint:* 0.619 and 4.25.

7.7 App challenge · 8 min

Load scenario 1, "A bed is an amplifier", at the default seed and the default speed of 5 (a 120 min run takes 24 s). The reactor is in Manual and lined out. Press Run, and at any time in the first 20 min raise Feed heater outlet from 381.0 to 382.0 °C. Run to 120 min after the step and export the CSV.

(a) In column `bed1_out_C`, read bed 1's outlet just before the step and 120 min after it. What gain per degree does the step give? *Checkpoint:* 404.86 and 408.45 °C (±0.02 °C); 3.59 per degree.

(b) Read the Bed 1 gain tile at the end and compare both numbers with the 3.33 of Equation (7.4) at the design point. Why do the three differ? *Checkpoint:* 3.9; the index rises from 0.538 to 0.592 across the step.

7.8 App challenge · 9 min

Load scenario 8, “The number the override reads”, at the default seed and speed 5. It loads with Bed rise compensation on Raw ΔT, the conversion loop in Auto and a feed cut to 220 t/h at 20 min.

(a) Run 120 min and export the CSV. Find the first rows with `override_active` $= 1$ and with `tripped` $= 1$, the inlet setpoint `tin_sp_C` on the last row before the trip, and `dt_max_C` on the trip row. *Checkpoint:* 25.80 min, 28.25 min, 346.1 °C, 42.2 °C.

(b) Press Reset, set Bed rise compensation to Compensated, run the same 120 min and export again. Find the first override row, the lowest inlet setpoint, the largest `dt_max_C` and the Trips tile. *Checkpoint:* 26.25 min, 377.45 °C, 29.9 °C, 0 trips.

7.9 App challenge · 8 min

Load scenario 11, “One bed at its limit, not four cut”, at the default seed and speed 5. It loads with Inlet control on MPC, the override armed and a feed cut to 220 t/h at 15 min.

(a) Run 120 min. Read the four bed inlets (`bed1_in_C` to `bed4_in_C`) at 120 min, the largest `dt_max_C`, whether `override_active` is ever 1, and the Converted tile. *Checkpoint:* 379.50, 381.00, 381.00, 381.00 °C; 27.50 °C; never; 246 t.

(b) Press Reset, set Inlet control to PID loops, run the same 120 min and read the same quantities and the first override row. *Checkpoint:* all four at 379.79 °C; 29.92 °C; from 21.10 min; 243 t.

(c) Where does the MPC’s extra conversion come from?

7.10 Application area · 4 min

A safety review of this reactor asks whether the rate-of-rise trip, set at 28 °C/min on a 0.8 min filter, is a protective layer. In scenario 9, with the reactor in Manual and the activity stepped to 1.45, the bed rise trip fires at 16.6 min with the hottest bed at 423.1 °C, while the filtered rate peaks at 2.7 °C/min.

(a) Is the rate trip a protective layer on this reactor? Explain from the bed equation.

(b) From what should its setting be derived, and what goes wrong if it is set far lower than 28 °C/min?

7.11 Application area · 4 min

A charge pump loses suction and the feed falls from 250 to 220 t/h while bed 1’s inlet is held at 381 °C.

(a) At 220 t/h the fold of bed 1 is at 381.87 °C. How much margin does the unit have, and what does the steady rise do?

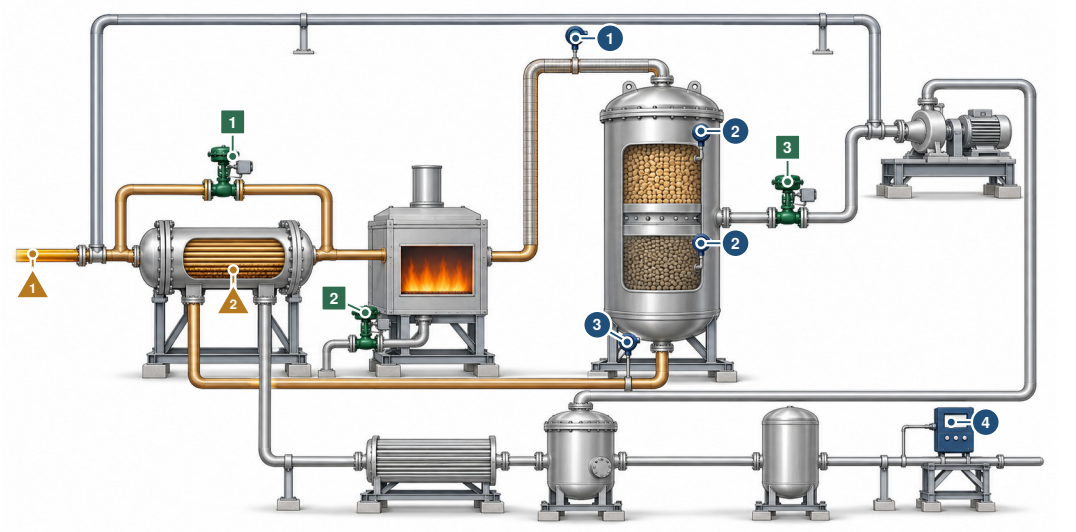
(b) What should the console operator do at once, and what permanent change to the control system would you propose?

7.12 Review the agent · 5 min

An AI assistant reviewing the reactor writes: “Bed 1 amplifies its inlet 3.33 times, and the four beds are in series, so a 1 °C upset at the feed heater reaches the bed 4 outlet as $3.33^4 \approx 123$ °C. The inlet loops act only on the inlets, so they cannot change this chain.” Find the flaw and say how you would check it in scenario 2, “Four inlet loops”.

CHAPTER 8

A Diesel Hydrotreater With Two Inlet Handles and a Half-Hour Sulfur Analyzer



The controller moves 1 valve that lets feed bypass the exchanger 2 fuel valve of the charge heater 3 cold gas valve between the beds

The controller reads 1 reactor inlet temperature 2 temperature of each bed 3 reactor outlet temperature 4 product sulfur, half an hour late

What disturbs it 1 a sourer feed 2 a fouled exchanger ▲ cracked stock in the feed

The hot oil leaving the bottom of the reactor runs back to the exchanger at the left to heat the feed, and only then through the cooler and two drums to the sulfur analyzer at the lower right. Illustration, not to scale.

IN PLAIN TERMS A Reactor Graded Half an Hour Late

The main idea. A diesel hydrotreater removes sulfur from diesel by passing it with hydrogen over hot grains that help the two react. The hot oil leaving the reactor hands most of its heat to the cold oil coming in.

The challenge. The heat goes round in a loop, so a small change in temperature comes back and grows for hours. The instrument that tests the product for sulfur reports thirty minutes late, when the oil it tested has long left.

Why it matters. Road diesel may hold only a few parts of sulfur in a million, and running hotter to be safe wears out the grains sooner.

THREE WAYS TO READ THIS CHAPTER

Explore	13 min	This page, the History boxes in Sections 8.2 and 8.5, the Did you know? boxes in Sections 8.1 and 8.6, the Try it boxes in Sections 8.4 and 8.6, and Key points in Section 8.9. No equations.
Study	50 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 8.6 with Figure 8.3, Section 8.7 and Section 8.8 with its Open challenge box.

The control question A diesel hydrotreater heats its feed with its own effluent and learns its product sulfur from an analyzer half an hour late. Which handle should hold the reactor inlet, and what can a sulfur loop, a Smith predictor or a predictive controller do before the analyzer reports?

Learning objectives

- Compute the loop gain of the feed and effluent exchanger and explain why a heater step keeps moving the inlet for hours.
- Explain the wrong-way rise of the reactor outlet after an inlet cut.
- Design mid-ranging so the fast bypass holds the inlet while the slow heater returns it to mid-range.
- Compare a PI, a Smith predictor and an MPC on a late analyzer, and state what a wrong predictor dead time costs.
- Judge a sulfur setpoint back-off against catalyst life and bed-rise limits.

Prerequisites Sections 1.4.4, 1.4.5, 1.4.6, 1.6.2 and 1.6.3. Chapter 7 for beds with interbed quench.

The app https://apmonitor.com/apps/ulsd-hydrotreater/

See also Volume 4, Chapter 10 and Volume 4, Chapter 9 for mid-ranging and valve-position control. Volume 5, Chapter 5 for a Smith-type predictor and Volume 5, Chapter 3 for a loop on a slow analyzer. Volume 7, Chapter 2 for an exothermic reactor near instability.

8.1 Introduction

A diesel hydrotreater must remove nearly all of the sulfur it is fed. In the generic unit simulated here, 200 t/h of diesel at 0.8 wt% sulfur leaves the reactor with 8.0 wppm (parts per million by weight), so 999 of every thousand sulfur atoms are taken out. Most of the heat the feed needs comes back from the reactor itself, since the feed–effluent exchanger returns 21.6 MW to the charge against the 7.34 MW fired in the charge heater. The product, finally, is judged by an online analyzer whose every report describes oil that left the reactor 30 min earlier.

The first of those facts is set by the regulations that created ultra-low-sulfur diesel (ULSD). In the United States the highway diesel limit fell from the 500 ppm in force since 1993 to 15 ppm for refiners from June 1, 2006 [1], and in the European Union diesel of at most 10 mg/kg had to be marketed from January 1, 2005 and to be the only diesel sold from January 1, 2009 [2]. In the simulator the room between the 8 wppm target and the 10 wppm limit is 2 wppm, and the evidence that the product is inside it arrives 30 min late.

> ✦ DID YOU KNOW? **The Largest Process After Crude Distillation**
>
> On January 1, 2003 the refineries of the United States could hydrotreat 11,247,745 barrels per calendar day, 67.7 % of their crude distillation capacity and more than any other major process downstream of the crude unit [3]. The unit of this chapter treats 200 t/h of diesel.

Three beliefs that a block diagram never challenges are corrected by the simulator. The first is that a fired heater moves the reactor inlet by what it adds. In case 1, 1 MW more in the coil is worth 5.36 K at once, yet the inlet ends 12.83 K higher after 400 min, 2.40 times as much, because the warmer effluent preheats the feed. The second is that cutting the inlet cools the reactor. In case 2 a 10 K cut raises the reactor outlet by 1.10 K for 24.5 min before it falls 13.3 K. The third is that a dead-time compensator removes the analyzer's delay. In case 4 the Smith predictor saves about a fifth of the integrated error against the best proportional-integral (PI) tuning found, 243 against 298 wppm min over 480 min, none of it before the first report, and with its assumed delay 15 min too long it costs 699 wppm min.

Each behavior belongs to a class. An adiabatic reactor preheated by its own effluent closes a positive energy loop whose design and control were analyzed early [4, 5], and whose bypass and furnace give two handles of different speed on one temperature, the pair that dynamic controllability

favors [6]. A packed bed whose composition answers faster than its catalyst temperature shows a transient of the wrong sign when its feed is cooled [7]. A fast valve of small range working beside a slow actuator of large range is the setting of mid-ranging control [8, 9], and a quality loop closed on a late analyzer is the setting of the Smith predictor [10]. A catalyst that deactivates in service [11] adds a far slower time scale. The same shapes recur in any reactor with heat recovery, in any trim valve beside a large actuator, and in any loop whose measurement is a laboratory or an analyzer.

These pieces usually reach a course separately: a heat-integrated reactor as a stability analysis, mid-ranging as a block diagram, dead time as a delay inside a transfer function with a perfectly known output. This chapter studies all three in the plant where they meet, a simulated ULSD hydrotreater in which the bed temperatures, the exchanger, the heater, the hydrogen loop, the sulfur analyzer and the catalyst activity are computed together. It measures how an inlet loop, mid-ranging, a quench override, a bed temperature loop and three sulfur controllers hold a 10 wppm specification. The lumped kinetics are teaching values and the controllers are established practice. Two of the measurements set bounds on the others. The wrong-way rise grows by about 0.4 K with every doubling of the cells and has not converged at 192 cells per bed, and the graded day needs a sulfur setpoint back-off once an expanded noise sweep exposes a missed specification. The model has sixty-four states whose sulfur, hydrogen and energy balances close over disturbed runs. It is the second user of the series' shared packed-bed cell module and sits behind a measurement boundary that keeps the true sulfur away from every controller. Nine cases each carry a completion rule.

Section 8.2 traces hydrotreating from cobalt molybdate to ten parts per million, Section 8.3 builds the model, Section 8.4 examines the plant without feedback, Section 8.5 sets out the controllers, and Section 8.6 reports the nine cases, the graded day among them. Section 8.7 sets the hardware of a real unit beside the model, Section 8.8 states what the model leaves out, and Section 8.9 draws the results together.

8.2 From Cobalt Molybdate to Ten Parts per Million

The catalyst came first. In 1943 Byrns, Bradley and Lee described the catalytic desulfurization of gasolines by a cobalt molybdate process [12], and four decades later the catalysts used for hydrodesulfurization and the related hydrotreating reactions were described as molybdenum, or tungsten, on high-surface-area alumina with cobalt or nickel added as promoters, although no general agreement had yet been reached on the structures present in the active sulfided catalyst, the Co–Mo–S type structures of that review's title, or on the origin of promotion [13]. By 1996 the sales of hydrotreating catalysts, which serve the catalytic hydrogenation processes that saturate unsaturated hydrocarbons and remove sulfur, nitrogen, oxygen and metals from refinery streams, made up close to 10 % of the world market for catalysts [14].

The specification came from engines, and the History box below gives its dates.

> **HISTORY 2001: The Rule That Set Fifteen Parts per Million**
>
> The United States Environmental Protection Agency set fuel quality standards that limited sulfur and aromatics in highway diesel beginning in 1993, at 500 ppm. Its rule of January 2001 tied new heavy-duty engine standards from model year 2007 to exhaust catalysts that sulfur poisons, and so required refiners to produce highway diesel with at most 15 ppm of sulfur beginning June 1, 2006, a reduction of 97 %, with retail stations supplied by September 1, 2006. A temporary option let up to 20 % of highway diesel be produced at 500 ppm through 2009 [1]. Directive 2003/17/EC required each member state of the European Union to market diesel of at most 10 mg/kg by January 1, 2005, and to allow only such diesel from January 1, 2009, a date the directive asked to be confirmed by the end of 2005 [2].

The hardware followed. The agency projected that every refiner could meet 15 ppm by extending the conventional hydrotreating it already ran, which would likely mean a second stage with an additional reactor and supporting hydrogen, sulfur recovery, amine and sour water units, and it

reported that the vendors' projections of hydrogen consumption and reactor volume varied widely [1]. For the deep desulfurization of gas oils, the apparent reaction order and the role of hydrogen sulfide carry consequences for the configuration of an industrial reactor [15]. An order above one for a lumped mixture of first-order species is the subject of Ho and Aris's note on apparent second-order kinetics [16], and the simulator borrows that reading (Section 8.3.1). Hydroprocessing catalysts also deactivate in service [11], and in the simulator the temperature needed for the same product climbs until a limit ends the cycle.

The control of heat-integrated reactors has its own history. The design and control of feed–effluent exchanger and reactor systems was analyzed in 1962 [4], and in 1982 the effect of design on the stability and control of fixed-bed reactors with heat feedback was set out as a set of concepts [5]. Later work on an exchanger, furnace and adiabatic reactor in series noted that the positive feedback of energy can make the process open-loop unstable, that the reactor inlet is usually held by bypassing the exchanger, and that dynamic controllability favors using both the bypass and the furnace, particularly when the reactor gain $\Delta T_{out}/\Delta T_{in}$ is large [6]. A pseudo-homogeneous plug-flow analysis in 1981 named the wrong-way behavior, a transient rise of temperature after a sudden reduction of the feed temperature, and found its size set by the adiabatic temperature rise, the activation energy, the size of the drop and the reactor length [7]. Axial dispersion of heat was later shown to decrease that excursion and prolong it [17], a result that matters here because the cells of the simulator are themselves a dispersion (Section 8.4).

The analyzer's delay has the oldest answer, the predictor that carries the name of Smith, from his 1957 paper on closer control of loops with dead time [10]. Mid-ranging, controlling one variable with a fast and a slow manipulated variable while the slow one returns the fast one toward a target, has its own design analysis [8] and is regarded as well established, with bumpless transfer and saturation the practical issues in a distributed control system [9]. Above those loops, model predictive control became the industrial standard, with refining among its main fields of application [18].

8.3 The Unit as Simulated

The simulated unit is a generic 200 t/h diesel hydrotreater at 60 bar, as shown in Figure 8.1. Diesel at $T_f = 180\,^{\circ}\mathrm{C}$ joins treat gas (recycle gas rich in hydrogen) at 4.0 mol/s per t/h, about 275 $\mathrm{Nm^3/m^3}$ at 850 $\mathrm{kg/m^3}$. The combined feed splits between the tubes of the feed–effluent exchanger and a bypass whose valve opening b sends a fraction $f = b f_{max}$ of it around the exchanger, with $f_{max} = 0.30$. The two streams mix, a fired charge heater raises the mix to the reactor inlet temperature (RIT), and the oil and gas flow down through bed 1 (70 t of catalyst), receive cold recycle gas as quench, and pass bed 2 (130 t). The effluent returns through the exchanger's shell, an air cooler and a separator, whose gas is purged, made up with hydrogen and recompressed as treat gas and quench, and whose liquid is stripped to product. Time in the engine runs in minutes, while bed flows are in mol/s and duties in MW. The engine, its clocks and its seeded noise are those of Section 1.6.1.

The variables are classified in Table 8.1. The state vector holds sixty-four quantities in use, the 48 cell temperatures, eleven other physical states (the heater duty Q, the coil outlet RIT, the exchanger outlet T_X, the bypass and quench valve positions, the recycle purity y, the feed sulfur w_{S0} and cracked-stock share x_{LCO} reaching the reactor, the feed rate F and the two bed activities), and five ledgers that integrate catalyst age, feed, product sulfur, fuel and hydrogen consumed.

8.3.1 Two Lumped Reactions

Hundreds of sulfur compounds and aromatic rings are folded into two reactions with teaching constants. Hydrodesulfurization (HDS), $\mathrm{S} + 3\,\mathrm{H_2} \rightarrow \mathrm{H_2S}$, consumes organic sulfur in the liquid, and the saturation of aromatics, $\mathrm{A} + 3\,\mathrm{H_2} \rightleftharpoons \mathrm{N}$, turns saturable aromatics into naphthenes up to an

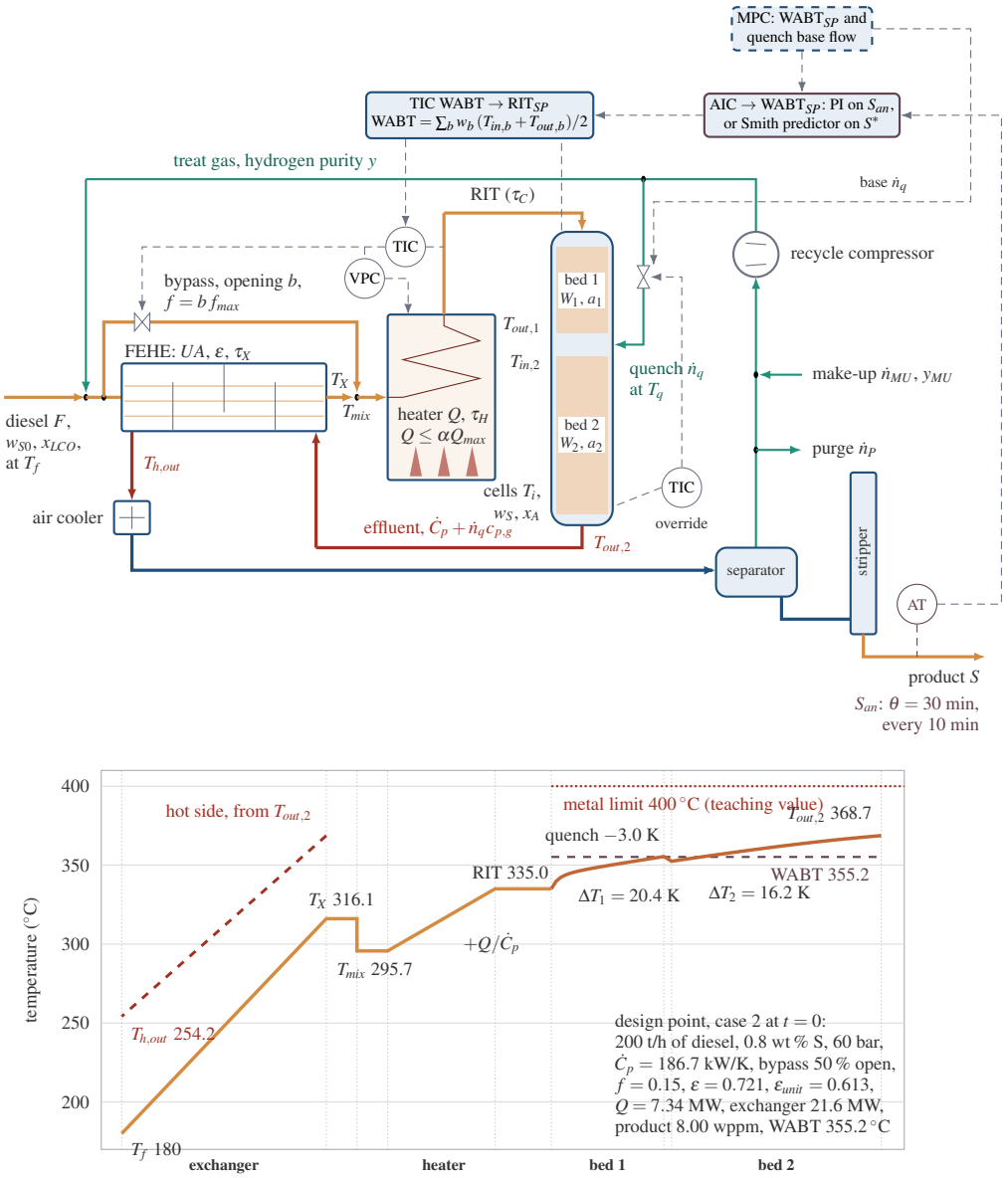


Figure 8.1 ULSD hydrotreater as simulated, drawn after the application's wide Unit view. Above, the loop. Feed at T_f and treat gas of purity y split between the exchanger's tubes and the bypass ($f = b\,f_{max}$), mix at T_{mix}, gain Q in the heater and reach the reactor at RIT. Each bed is 24 cells at T_i with activity a_b, the quench $\dot{n}_q$ at T_q joins between them, and the effluent at $T_{out,2}$ heats the feed and leaves the shell at $T_{h,out}$. The TIC on RIT moves the bypass, the VPC moves the heater to bring the bypass back to its target, the override TIC adds quench on $T_{out,2}$, the WABT controller sets the RIT setpoint, the AIC (a PI on the analyzer reading S_{an} or the Smith predictor on S^*) sets the WABT setpoint, and the MPC (dashed) can set the WABT setpoint and the quench base flow. Below, on the same stations, the design temperatures drawn to scale, with the 48 cell temperatures, the hot side dashed and the 400 °C metal limit dotted. It is a drawing of the model and not of piping or safeguards.

Table 8.1 Hydrotreater Variables, Their Roles and Design Values (Case 2 at $t = 0$)

Variable	Role	Design	Range
S	CV, product sulfur (true)	8.00 wppm	SP 4–15
S_{an}, S^*	measured, estimated sulfur	8.00 wppm	—
WABT	CV and MV, bed temperature	355.2 °C	SP 320–400
RIT	CV and MV, reactor inlet	335.0 °C	SP 300–395
$T_{out,2}$	CV of the override	368.7 °C	SP 360–400
b	MV, bypass valve	50 % open	0–100 %
Q	MV, heater duty	7.34 MW	1–14
$\dot{n}_q$	MV, quench gas	60 mol/s	0–400
F	DV, feed rate	200 t/h	150–240
$w_{S,f}$	DV, feed sulfur (tank)	0.8 wt%	0.3–1.6
x_{LCO}	DV, light cycle oil	0	0–40 %
ϕ	DV, exchanger UA	100 % of clean	40–100 %
α	DV, heater available	100 %	30–100 %
y_{MU}, bias	DV, make-up purity, analyzer	97 mol%, 0	85–100, ±3
T_i	states, 48 cells	Figure 8.1	—
T_X, T_{mix}	exchanger outlet, mix	316.1, 295.7 °C	—
ΔT_1, ΔT_2	bed rises	20.4, 16.2 K	limit 30
y	state, recycle purity	87.3 mol%	—
a_1, a_2	states, activity	1.000, 1.000	—

equilibrium. Equation (8.1) gives both rates on a catalyst-mass basis.

$$\begin{aligned} r_S &= a\,k_S\,\kappa_S\,e^{-\gamma_L x_{LCO}}\,\frac{w_{S0}}{w_r}\left(\frac{w_S}{w_{S0}}\right)^n\frac{p_H/p_r}{1+K_{H_2S}\,p_{H_2S}} \\ r_A &= a\,k_A\,\kappa_A\,\frac{p_H}{p_r}\left(x_A-\frac{x_N}{K_A\,p_H^3}\right) \\ K_A &= \exp\left(\Delta S_A/R-\Delta H_A/RT\right) \end{aligned} \tag{8.1}$$

where $\kappa_j = \exp[-E_j/R\,(1/T - 1/T_r)]$ with $T_r = 350\,°\mathrm{C}$, a is the bed's activity, w_S the sulfur mass fraction of the liquid and w_{S0} its value at the reactor inlet, x_A and x_N the mole fractions of aromatics and naphthenes in the liquid, $w_r = 10^{-3}$, $p_r = 45\,\mathrm{bar}$, and p_H and p_{H_2S} are partial pressures, P times each flow over the gas, which counts hydrogen, H_2S, the light impurities of the recycle and a fixed vaporized share of 0.30 of the oil's moles. Light cycle oil carries the more hindered sulfur, through $e^{-\gamma_L x_{LCO}}$ with $\gamma_L = 2$, and more saturable aromatics, 0.65 against 0.25 mole fraction for straight-run diesel. The same ΔH_A enters the equilibrium and the energy balance.

The order $n = 1.65$ is written on the fraction of the feed's sulfur left, not on the absolute sulfur, and the reason is the reading of an order above one. For a mixture of first-order species whose rate constants k follow a gamma distribution of shape m and mean $\bar{k}$, the fraction left after a contact time τ is as shown in Equation (8.2).

$$\begin{aligned} X(\tau) &= \int_0^\infty g(k)\,e^{-k\tau}\,dk = \left(1+\frac{\bar{k}\tau}{m}\right)^{-m} \\ \frac{dX}{d\tau} &= -\bar{k}\,X^{1+1/m} \end{aligned} \tag{8.2}$$

which is an n-th order lump with $n = 1 + 1/m$, here $m = 1.54$, and an exponential distribution ($m = 1$) gives the apparent second order of Ho and Aris [16]. The fraction left depends on the reactor and not on the feed's level, so a sourer feed makes a proportionally sourer product. With the order written on the absolute sulfur and calibrated to the same rate at 0.8 wt%, a feed 50 % sourer came out cleaner at a fixed 335 °C inlet, 7.32 wppm against 8.00, where the fraction form gives 10.92 wppm and needs a 5.1 K hotter inlet. At design the beds release 6.87 MW, 5.07 MW of it from saturation and 1.80 MW from HDS, and bed 1 removes 99.17 % of the sulfur. The equilibrium saturates 92, 61 and 38 % of

the aromatics at 335, 368 and 385 °C under 45 bar of hydrogen, so a hotter late-cycle bed releases less heat.

8.3.2 Thermal Cells With Algebraic Composition

Each bed is a train of 24 cells of equal catalyst mass from the shared packed-bed module of the series, one train per bed so the quench can add gas and heat capacity between them. The liquid crosses a bed in minutes while the catalyst holds its heat for tens of minutes, so each cell's composition is algebraic and its temperature is a state, as shown in Equation (8.3).

$$\begin{aligned} F_i &= F_{i-1} + \nu^{\mathsf{T}} \xi_i, \qquad \xi_{i,j} = W_c\, r_j(T_i, F_i) \\ C_i \frac{\mathrm{d}T_i}{\mathrm{d}t} &= \dot{C}_{p,b}\,(T_{i-1} - T_i) - \textstyle\sum_j \xi_{i,j}\,\Delta H_j \end{aligned} \tag{8.3}$$

where F holds the molar flows of S, H_2S, A, N and H_2, ν is the stoichiometric matrix, $W_c = W_b/24$, $C_i = W_c c_{cat} + C_{v,b}/24$ the catalyst plus a share of the liquid held up and the vessel, T_0 the bed inlet, and $\dot{C}_{p,b}$ the heat-capacity flow, held along a bed, $\dot{C}_p = \dot{m} c_{p,L} + \dot{n}_g c_{p,g} = 186.7\,\mathrm{kW/K}$ in bed 1 and $\dot{C}_{p,2} = \dot{C}_p + \dot{n}_q c_{p,g} = 188.6\,\mathrm{kW/K}$ in bed 2. The bed heat capacities are 114.2 and 212.2 MJ/K, thermal time constants of 10.19 and 18.75 min. Each cell is solved by Newton's method on the extents with a fixed four iterations from zero extent, converged to 2.6×10^{-9} of a forty-iteration solution from 300 to 395 °C, so the derivative depends on the state alone and every run replays bit for bit.

8.3.3 Quench and the Weighted Average Bed Temperature

The quench is recycle gas at $T_q = 65\,^{\circ}\mathrm{C}$, so bed 2's inlet and the weighted average bed temperature (WABT) are as shown in Equation (8.4).

$$\begin{aligned} T_{in,2} &= \frac{\dot{C}_p\, T_{out,1} + \dot{n}_q c_{p,g} T_q}{\dot{C}_p + \dot{n}_q c_{p,g}} \\ \mathrm{WABT} &= \textstyle\sum_{b=1}^{2} w_b\,(T_{in,b} + T_{out,b})/2 \end{aligned} \tag{8.4}$$

where $w = (0.35, 0.65)$ are the catalyst shares and $T_{in,1} = \mathrm{RIT}$. At design 60 mol/s of quench (4.8 $\mathrm{kNm^3/h}$) cools bed 2's inlet by 2.96 K, and the valve passes at most 400 mol/s.

8.3.4 Exchanger, Bypass and Charge Heater

The exchanger is counter-current, with the cold flow $(1-f)\dot{C}_p$ in the tubes and the effluent $\dot{C}_{p,2}$ in the shell. Its effectiveness, its cold outlet through a lag for the metal and holdup, and the mix are as shown in Equation (8.5).

$$\begin{aligned} \varepsilon &= \frac{1 - e^{-N_t(1-C_r)}}{1 - C_r e^{-N_t(1-C_r)}}, \qquad N_t = \frac{\phi\, UA}{C_{min}} \\ q_X &= \varepsilon\, C_{min}\,(T_{out,2} - T_f) \\ \tau_X \frac{\mathrm{d}T_X}{\mathrm{d}t} &= T_f + \frac{q_X}{(1-f)\dot{C}_p} - T_X \\ T_{mix} &= (1-f)\,T_X + f\,T_f, \qquad T_{h,out} = T_{out,2} - \frac{q_X}{\dot{C}_{p,2}} \end{aligned} \tag{8.5}$$

where $C_r = C_{min}/C_{max}$, $UA = 344\,\mathrm{kW/K}$ clean, ϕ the fouling factor, q_X the duty returned to the feed and $\tau_X = 3\,\mathrm{min}$. At design with the bypass half open ($f = 0.15$), $\varepsilon = 0.7212$, the exchanger returns 21.60 MW, the mix is at 295.67 °C and the effluent leaves at 254.20 °C. A first-order lag carries the heater's command to its absorbed duty, and a second lag, the coil, carries the duty to the inlet, as shown in Equation (8.6).

$$\tau_H \frac{\mathrm{d}Q}{\mathrm{d}t} = \mathrm{sat}(Q_c) - Q, \qquad \tau_C \frac{\mathrm{dRIT}}{\mathrm{d}t} = T_{mix} + \frac{Q}{\dot{C}_p} - \mathrm{RIT} \tag{8.6}$$

where sat clamps the command Q_c between $\min(Q_{min}, \alpha Q_{max})$ and αQ_{max}, with $Q_{max} = 14\,\text{MW}$, $Q_{min} = 1\,\text{MW}$, α the share of burners in service, $\tau_H = 6\,\text{min}$ and $\tau_C = 1\,\text{min}$. The bypass and quench valves follow their commands through lags of 0.1 and 0.15 min. The design duty is 7.343 MW.

8.3.5 Recycle Hydrogen

The hydrogen loop is one well-mixed inventory whose pressure is held, so the make-up replaces every mole that leaves, as shown in Equation (8.7).

$$\begin{aligned} N_L \frac{\mathrm{d}x}{\mathrm{d}t} &= \dot{n}_{MU}(1 - y_{MU}) - \dot{n}_P x, \qquad x = 1 - y \\ \dot{n}_{MU} &= \dot{n}_{H_2,used} + \dot{n}_P \end{aligned} \tag{8.7}$$

where $\dot{n}_P = 36.4\,\text{mol/s}$ is the purge, $N_L/\dot{n}_P = 60\,\text{min}$, and $\dot{n}_{H_2,used}$ the hydrogen the beds consume. At design the beds use 117.6 mol/s, the make-up is 12.4 $\text{kNm}^3\text{/h}$ and the purity 87.3 mol%. The recycle carries no H_2S, as if scrubbed.

8.3.6 Feed Line, Activity and a Compressed Cycle

The feed's sulfur and cracked-stock share reach the reactor through a 20 min surge drum, $\tau_{feed}\,\dot{w}_{S0} = w_{S,f} - w_{S0}$ with $w_{S,f}$ the tank's sulfur, and its rate through a 1 min lag. Each bed loses activity at its mean temperature $\bar{T}_b$, as shown in Equation (8.8).

$$\frac{\mathrm{d}a_b}{\mathrm{d}t} = -k_d \exp\!\left[-\frac{E_d}{R}\left(\frac{1}{\bar{T}_b} - \frac{1}{628.15\,\text{K}}\right)\right](1 + \beta x_{LCO})\,\frac{c}{43\,200} \tag{8.8}$$

per minute, where $k_d = 0.019$ per month at 355 °C, $E_d = 100\,\text{kJ/mol}$, $\beta = 2$, 43 200 min make a month, and c is a declared cycle compression, 1 in every case but the cycle case, where $c = 720$ passes a month of catalyst life every simulated hour. Five kelvin hotter ages the catalyst 16 % faster and 20 % of light cycle oil 40 % faster. The run ends when the hottest catalyst reaches the 400 °C metal limit.

8.3.7 What the Controllers Can Read

All the controllers share one packet of measurements, refreshed at each 10 s execution. The inlet, bed and outlet thermocouples lag 6 s with 0.1 K of noise, the duty, feed and quench meters carry 0.3, 0.2 and 1 %, the recycle purity analyzer lags 1 min, and the assay of the feed tank lined up supplies its sulfur and its cracked-stock share. The product sulfur analyzer draws a sample every 10 min and reports it 30 min after the sample left the reactor, a 20 min product path and sample line followed by a 10 min analysis, with 0.1 wppm of noise and a calibration bias. It is the sampled, delayed analyzer of Section 1.4.6. Out of service, its reading holds and is marked stale after two missed reports. The true sulfur, the cell profiles and the activities never enter the packet, and each execution draws the same normal deviates whatever the settings, so runs differing in a single setting meet identical noise.

8.3.8 Parameters, Initial State and Integration

Table 8.2 lists the parameters. Each case starts lined out, the inlet at its setpoint (or, with a sulfur loop in Auto, at the inlet that holds the sulfur setpoint), the bypass at its 50 % target, the heater at the duty that holds the inlet, the quench at its base flow and the purity at the fixed point of Equation (8.7). A cell profile is lined out bed by bed, each cell bracketed from its inlet and bisected. The state advances by fourth-order Runge–Kutta steps of 5 s, the controllers execute every 10 s, and a row is logged every 30 s. Halving the step leaves the wrong-way rise of case 2 at 1.1035 K and the true sulfur of case 4 at 240 min at 8.056 14 wppm. Over four disturbed runs the sulfur entering the reactor equals the product plus the H_2S made to better than 10^{-12} of itself, the hydrogen loop closes to round-off, and the energy around the exchanger, heater, beds and quench closes to 3×10^{-7} of the heat entering, with the stored heat of the beds, coil and exchanger metal counted.

Table 8.2 Parameters of the Simulated Hydrotreater by Access (A Adjustable in Advanced, H Hidden, F Written Into the Source)

Symbol	Meaning	Value	Unit		Symbol	Meaning	Value	Unit	
P	reactor pressure	60	bar	A	UA	exchanger, clean	344	kW/K	H
W_1, W_2	catalyst per bed	70 000, 130 000	kg	H	f_{max}	bypass share at full open	0.30	—	A
c_{cat}	catalyst heat capacity	1000	J/(kg K)	H	τ_X	exchanger metal lag	3	min	H
$C_{v,1}, C_{v,2}$	holdup and vessel per bed	44.2, 82.2	MJ/K	H	τ_V, τ_q	bypass, quench valve lags	0.1, 0.15	min	H
N_c	cells per bed	24	—	H	Q_{max}, Q_{min}	heater rated, minimum	14, 1	MW	A, H
k_S	HDS at 350 °C	1.104×10^{-3}	mol/(s kg)	H	τ_H, τ_C	firebox, coil	6, 1	min	A, H
E_S, n	HDS activation, order	120, 1.65	kJ/mol, —	H	η_H	heater efficiency	0.85	—	H
K_{H_2S}	H_2S inhibition	0.2	1/bar	H	T_f	combined feed	180	°C	H
γ_L	LCO sulfur refractoriness	2	—	H	T_q	quench gas	65	°C	H
ΔH_S	heat of HDS	−130	kJ/mol S	H	τ_{feed}	surge drum	20	min	A
k_A	saturation at 350 °C	6.364×10^{-4}	mol/(s kg)	H	$\dot{n}_P$, $N_L/\dot{n}_P$	purge, loop time	36.4, 60	mol/s, min	H
$E_A, \Delta H_A$	saturation activation, heat	80, −200	kJ/mol	H	θ, period	analyzer delay, period	30, 10	min	A
ΔS_A	saturation entropy	−403	J/(mol K)	H	k_d	activity lost at 355 °C	0.019	1/month	H
$M_L, c_{p,L}$	diesel molar mass, heat capacity	230, 2900	g/mol, J/(kg K)	H	E_d, β	deactivation activation, LCO	100, 2	kJ/mol, —	H
$c_{p,g}$	recycle gas heat capacity	32	J/(mol K)	H	age, c	starting age, compression	0, 1	months, —	A
—	vaporized oil share	0.30	mol/mol	H	ΔT_{max}	bed rise limit	30	K	A
—	treat gas to oil	4.0	mol/s per t/h	H	T_{metal}	metal limit	400	°C	A
x_A	saturable aromatics, SR, LCO	0.25, 0.65	mol/mol	H	w_b	catalyst shares	0.35, 0.65	—	F
$\dot{n}_{q,max}$	quench valve capacity	400	mol/s	H	T_r, w_r, p_r	reference point	350, 10^{-3}, 45	°C, —, bar	F

8.3.9 Assumptions

Two lumped reactions stand for hundreds, so olefin saturation, denitrogenation and mild cracking are folded into the HDS heat, and the constants are teaching values chosen to give 8 wppm at a 335 °C inlet with rises of 20.4 and 16.2 K, not fitted to a catalyst. Each bed is 24 stirred cells, which act as axial dispersion with a Péclet number near 48 per bed, so the wrong-way rise depends on the cell count (Section 8.4). The liquid and gas hold no inventory, the reactor has one pressure and no pressure drop, the wetting and maldistribution of a trickle bed are absent, and the vaporized share of the oil is fixed. The exchanger has one lag, and its stored heat is not a fixed capacity when the bypass moves. The heater has no tube metal, draft or fuel system, the separator, cooler, scrubber and stripper are not modeled beyond the analyzer's delay, and the surge drum is a fixed 20 min lag whatever the flow, which leaves 0.026 % of the sulfur unbalanced in the drum when the feed rate steps. Activity is one number per bed with no start-of-run decline, metals or nitrogen poisoning. Sulfur in wppm is reckoned on the feed's liquid mass, which the product undercuts by 0.42 %, so the design's 8.00 wppm would read 8.03 on the product's own mass. The bed rise limit of 30 K, the 400 °C metal limit and the 14 MW rating are teaching values, where real limits come from the licensor's operating guidelines, the reactor's design temperature under the ASME pressure vessel code and the heater's design.

8.4 Three Behaviors Before Any Feedback

8.4.1 The Heat That Comes Back

The exchanger and its bypass return $\varepsilon_{unit} = (T_{mix} - T_f)/(T_{out,2} - T_f) = 0.613$ of the effluent's temperature above the feed, and at design a kelvin more at the inlet makes 1.187 K more at the outlet,

so a kelvin added at the inlet returns as 0.728 K and the series converges slowly. That return is the loop gain L of Equation (1.11), and Equation (1.8) gives the amplification $1/(1-L)$ it implies for a small move (Section 1.4.5). The gain of that loop moves with the operating point, 0.899 at a 320 °C inlet and 0.431 at 350 °C, because a hotter bed brings its aromatic saturation nearer equilibrium and releases less extra heat. In case 1 in Manual, 1 MW more firing at 10 min raises the inlet 5.23 K in the next 30 min, 95 % of that within 19 min, then 8.65, 11.69 and 12.83 K at 60, 120 and 400 min after the step. Opening the bypass from 50 to 60 % at 40 min takes the inlet from 340.23 to 337.67 °C in 2 min, 81 % of it in the first minute, because only the coil stands between the mix and the inlet. The two handles act on two clocks, and both are then multiplied by the loop. A unit with a larger exchanger or a hotter exotherm could pass a loop gain of one, where the positive feedback of energy makes the process open-loop unstable [6].

8.4.2 Wrong Way, and Its Size Under Refinement

Case 2 lowers the inlet setpoint from 335 to 325 °C at 10 min. The reactor outlet rises from 368.69 to 369.79 °C, peaks at 31.5 min, stays above its start from 10.5 to 35.0 min, and only then falls, to 355.9 °C at 60 min and 13.3 K below its start at 120 min. The first cells cool and react less, so unsaturated aromatics and sulfur reach catalyst still at its old temperature and release their heat there, while the cooling front moves at the pace of the catalyst's heat capacity. The mechanism was isolated with the beds alone. For an ideal 10 K inlet step at 24 cells the rise is 1.173 K whether the heat capacity is scaled by 1, 0.1 or 0.01, and only its time moves (20.7, 2.07 and 0.21 min), while with the inlet approaching over 2 min the rise shrinks with the heat capacity, 1.082, 0.800, 0.427, 0.159 and 0.057 K at scales of 1, 0.3, 0.1, 0.03 and 0.01. The aromatic saturation carries the effect, 0.685 K with its heat alone against 0.018 K with the HDS heat alone, because bed 1's first cells remove nearly all the sulfur in any case.

> **TRY IT IN THE APP 2 · Cut the inlet 10 °C**
>
> Load the case and press Run. At 10 min lower the RIT setpoint from 335 to 325 °C, after Run and not before. Watch the Reactor outlet trend. The outlet first rises, from 368.69 to 369.79 °C at 31.5 min, and stays above its start until 35.0 min. By 120 min it is 13.3 K below its start.

The size of the rise has not converged with the cells. Case 2 as shipped gives 0.69, 1.10, 1.50, 1.88 and 2.24 K at 12, 24, 48, 96 and 192 cells per bed, about 0.4 K for each doubling, and the peak moves later, from 28.5 to 35.0 min between 12 and 96 cells, as shown in Figure 8.2a. Stirred cells in series are an axial dispersion, and dispersion is known to reduce the excursion [17], so the 24-cell figure is the size this model's dispersion allows, below what a plug-flow bed would show. The sign, the timing within a few minutes and the cause survive every refinement, while the size belongs to the dispersion the cells stand for. A real bed has its own dispersion and interphase transfer, which neither the cells nor the plug-flow limit represent. The converter beds of Volume 13, Chapter 4 show the same dependence of the wrong-way rise on the cell count.

Pause. Try Exercise 8.1 before reading on.

8.4.3 A Sourer Feed Brings Its Own Heat

At a fixed 335 °C inlet a feed of 1.2 wt% sulfur leaves 10.92 wppm, and 340.11 °C at the inlet is needed for 8 wppm. The extra sulfur also releases heat in bed 1. Under the Smith predictor at 240 min of case 4, with the inlet at 340.06 °C, bed 1 rises 27.6 K on 5.09 MW of reaction heat and bed 2 only 12.6 K, the outlet reaches 377.13 °C, the exchanger returns 22.59 MW, and the heater fires 7.30 MW, slightly less than at design, for an inlet 5 K hotter (Figure 8.4). A rise measured as outlet less inlet can also exceed the reaction's share. In case 8 as shipped, at 60 min, bed 2 shows 34.58 K while its reaction heat is worth 15.33 K of its heat-capacity flow, since the collapsing inlet draws the catalyst's stored heat out through the outlet.

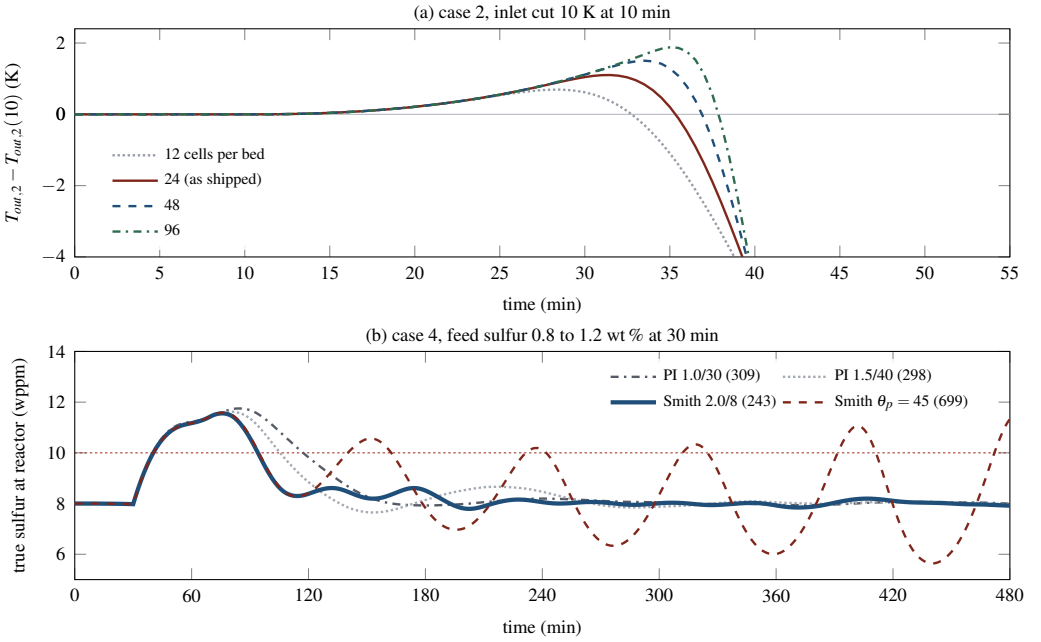


Figure 8.2 Two responses from the application's engine. (a) Case 2, the reactor outlet after a 10 K inlet cut at 10 min, less its value at the cut, at four cell counts. The rise grows with refinement and has not converged. (b) Case 4, the true product sulfur after the feed sulfur steps from 0.8 to 1.2 wt% at 30 min, with the IAE over 480 min in parentheses and the 10 wppm specification dotted.

8.5 The Control Layers in the Code

At each 10 s execution the code first forms the true state, then the instrument readings and the sulfur estimates, and then runs the sulfur layer or the MPC, the WABT loop, the inlet loop, mid-ranging and the quench override from the top down. Only then are the valve and heater commands written. A single Manual and Auto switch in the header covers the bypass, heater and quench stations, and while in Manual each loop follows its station, so switching back to Auto causes no bump.

8.5.1 The Shared PID

Every loop uses the series' velocity-form PID routine of Section 1.6.2, Equations (1.16) to (1.19). The direction $\sigma = \pm 1$ given for each loop below is the action a of those equations. Anti-windup back-calculates the integral only, with τ_I as its tracking time, and any integral time from 600 min upward leaves the loop proportional-only. No loop uses derivative action. In Manual, or when a layer above holds, the routine solves for the integral that makes its output equal the value imposed, which is how every transfer stays bumpless.

8.5.2 The Inlet Loop and Mid-Ranging

The inlet loop moves the bypass between 0 and 100 % at $K_c = 3\,\%/\mathrm{K}$, $\tau_I = 1$ min and $\sigma = -1$, so a hot inlet opens the bypass. Mid-ranging is a valve-position controller on the bypass command against its 50 % target, moving the heater at 0.04 MW per percent with $\tau_I = 20$ min and $\sigma = +1$ between $\min(Q_{min}, \alpha Q_{max})$ and αQ_{max}. With mid-ranging off the heater stays at its manual station, which is how every case but 4 to 7 ships. A power block with three such layers is the subject of Volume 4, Chapter 10.

8.5.3 The Quench Override

A PI on the measured reactor outlet adds quench above the base flow when the outlet passes its override setpoint, 385 °C by default and 374 °C in case 6, at 6 mol/s per kelvin with $\tau_I = 20$ min and $\sigma = -1$, limited to the valve's capacity less the base. In development, a faster tuning oscillated against bed 2's own wrong-way response. The quench loops of Chapter 7 hold each bed's inlet instead.

8.5.4 WABT and Sulfur Layers

A PI on the measured WABT sets the inlet setpoint between 300 and 395 °C at 0.5 K/K with $\tau_I = 10$ min. Above it, the sulfur controller sets the WABT setpoint between 320 and 400 °C with $\sigma = -1$. As a Sulfur PI it acts once per analyzer report on the reading, integrating over the time since the last report. As a Smith predictor the same PI acts every 10 s on the estimate S^* of Equation (8.9). Both layers hold their integrals when the inlet cannot move further, with the bypass shut and the heater fixed or at its limit (or the bypass full open and the heater at its minimum), and the sulfur layer holds while the analyzer is stale or out of service. A WABT schedule, when on, learns a ramp from the sulfur controller's own moves, its rate changing by the trim's last move over a 30 min learning time at each report, and adds the integrated ramp to the WABT setpoint ahead of the trim.

8.5.5 The Smith Predictor

The predictor models the logarithm of the sulfur answering the WABT setpoint that the sulfur controller itself sets, and moves the last report forward to the present, as shown in Equation (8.9).

$$\begin{aligned} z &= \frac{g}{(\tau_1 s + 1)(\tau_2 s + 1)} \left(\mathrm{WABT}_{SP} - \mathrm{WABT}_0\right) \\ S^* &= S_{rep} \exp\left[z(t) - z(t_{rep} - \theta_p)\right] \end{aligned} \tag{8.9}$$

where $g = -0.057$ per kelvin and $\tau_1 = \tau_2 = 10$ min were identified on the commissioned plant from ±5 K setpoint steps, S_{rep} is the last report, t_{rep} the time it arrived, and θ_p the predictor's dead time, a setting of 30 min. Written in logarithms, the correction is exact for a gain that scales with the sulfur. Because the model's input is the controller's own output, no disturbance moves it, so a feed change still arrives one delay late. With θ_p equal to the true delay the delayed model cancels each move when the analyzer catches up, and with θ_p wrong the controller sees its own correction appear and vanish. In RIT only and Manual, where no controller sets the WABT, the estimate shown runs a second model of 2.75 min from the measured WABT. In development, a predictor driven by the measured WABT read a sourer feed's own exotherm as lower sulfur and moved the wrong way, which is why the controller's output drives this one. Volume 5, Chapter 5 uses a predictor of the same family on the delay of a forming line.

🕮 **Pause.** Try Exercise 8.2 before reading on.

8.5.6 Model Predictive Control

The MPC is the series' linear MPC of Section 1.6.3, here on two decisions, the WABT setpoint W and the quench base flow q, every 5 min over 36 steps (3 h) with moves blocked as 1, 1, 2, 2, 4, 6, 8 and 12 steps. Its model is fixed, identified on the commissioned plant with the WABT loop closed and mid-ranging on. The logarithm of the sulfur answers W at −0.0562 per kelvin through lags of 10.5 and 10 min, the outlet answers W at 0.952 K/K and q at −0.0191 K per mol/s, each bed rise answers both, and the feed tank's sulfur assay enters the sulfur one-for-one in logarithm through the surge

drum, a feedforward the PI and the Smith predictor lack. The program is Equation (8.10).

$$\begin{aligned} \min\ & \textstyle\sum_k \left[w_S (\ln S_k - \ln S_{SP})^2 + c_W \mathrm{WABT}_k \right] \\ & \quad + \textstyle\sum_j \Delta u_j^{\mathsf{T}} \Lambda \Delta u_j + \text{slack penalties} \\ \text{s.t. } & 320 \le W \le 400, \quad 0 \le q \le 400 \\ & |\Delta W| \le 4, \quad |\Delta q| \le 100 \\ & T_{out,2,k} \le T_{metal} - 5, \quad \Delta T_{b,k} \le \Delta T_{max} - 1 \end{aligned} \tag{8.10}$$

where $w_S = 100$ on every step and on the last, the catalyst-life cost is $c_W = 0.2$ per kelvin of WABT, $\Lambda = \mathrm{diag}(1, 10^{-4})$, a weight of 10^{-6} holds the quench decision, and the outlet and rise limits are soft, priced at 50 per squared unit and 5 per unit of slack. Each output carries an offset, the sulfur's formed at each report against the model one predictor dead time earlier, as the Smith predictor's is, and the temperatures' filtered over 5 min. If the solver returns a move outside its bounds, the previous decisions stay in force. Over the challenge day the MPC solves 288 times with no fallback. The inlet loop and mid-ranging stay under it, so the bypass keeps its 10 s scan.

> ⧖ HISTORY **Late 1980s: A Hydrotreater Tests Predictive Control**
>
> A hydrotreater was a test case in the history of predictive control. In the late 1980s engineers at a research center in France developed a controller that they described as a bridge between state-space and predictive algorithms, and to motivate it they discussed a hydrotreater with four reactor beds in series [18]. The control system had to hold the average bed temperature at a setpoint, keep the temperature differences between the beds close to a profile and prevent any violation of the maximum temperature in each reactor. It moved the first reactor inlet temperature and the quench flows between the beds. A disturbance in the first bed finally affects every bed downstream, which a model of independent outputs does not carry [18]. The MPC of this section decides the same two kinds of move, a bed temperature setpoint and a quench flow.

8.5.7 Three Models Kept Apart

The plant is Section 8.3. The identified model is whatever a learner estimates from step tests, for example the 5.36 K/MW direct gain of the heater or the gain of −0.057 per kelvin in the logarithm of the sulfur. The controllers' models are the predictor's two lags and the MPC's fixed map, which know neither the activity, the feed's true sulfur in the beds nor the fouling. The end of run reads the true hottest cell and latches.

8.6 Nine Cases Measured

The cases are listed in Table 8.3, each from the defaults of Section 8.5 (RIT only, the inlet at 335 °C, the slow Sulfur PI at 0.4 K/wppm and 90 min ready, mid-ranging off, the override at 385 °C) with only its differences listed. The figures below come from the engine advanced without a display at the default seed, with each learner move at the time the text names. Every case runs identically when loaded first, from the dropdown after another case and after Reset, and the 211 numbers the application's own texts quote reproduce on it. The outcomes are collected in Table 8.4.

8.6.1 Two Handles, Mid-Ranging and a Fouled Exchanger

In case 3 the bypass alone holds an 8 K rise of the inlet setpoint, but it closes fully by 27.5 min and stays shut until 41 min while the cooler effluent of the first minutes comes back, the inlet sagging to 341.7 °C, and then parks at 32.3 %, with 32 % of its travel left for the next upset. Turned on at 120 min, mid-ranging raises the heater from 7.33 to 7.89 MW and returns the bypass within 3 % of its target by 175 min, the inlet within 0.5 K of 343 °C throughout. With mid-ranging on from the start the bypass dips only to 23.3 % and is back by 42.5 min.

Case 8 shows what a bypass with nothing behind it does. Fouled to 60 % at 30 min, the exchanger would return 17.70 MW instead of 21.6 at the design inlet. The bypass shuts by 34.5 min, and the

Table 8.3 The Nine Cases

#	Preset and events	Teaching point
1	Manual, learner raises the heater 1 MW, then opens the bypass to 60 %	two handles, two clocks
2	Learner cuts the RIT setpoint 335 to 325 °C	wrong-way rise
3	Learner raises RIT to 343 °C with mid-ranging off, then on	mid-ranging
4	Sulfur PI 1.0/30 min and mid-ranging, feed 0.8 to 1.2 wt% at 30 min	PI against Smith
5	Mid-ranging, feed 1.2 wt% at 30, 230 t/h at 240 min	feed changes
6	Mid-ranging, override 374 °C, 20 % LCO at 30 min	quench and heat
7	Sulfur PI, mid-ranging, compression 720	catalyst cycle
8	Exchanger fouled to 60 % at 30 min	a bypass out of range
9	Challenge: feed 1.0 wt% at 4, analyzer out 10–12, fouling 60 % at 16 h	the whole problem

Table 8.4 Measured Outcomes by Case, Default Seed

#	Setting	Outcome
1	heater +1 MW	RIT +5.23 in 30 min, +12.83 K in 400 min
	bypass 50 to 60 %	RIT −2.57 K in 2 min
2	RIT SP −10 K	outlet +1.10 K at 31.5 min, −13.3 K at 120
3	mid-ranging off	bypass shut 27.5–41 min, parks at 32.3 %
	on at 120 min	heater 7.33 to 7.89 MW, bypass back by 175 min
4	PI 1.0/30	IAE 309, 77 min over 10 wppm
	Smith 2.0/8	IAE 243, 54 min over
	Smith, $\theta_p = 45, 15$	IAE 699, 399
5	RIT only	440 min over, peaks 11.1 and 15.1 wppm
	Smith	92 min over, 7.91 wppm at 480 min
6	RIT only, override	outlet held at 374.2 °C, 12.3 wppm
	Smith from 240 min	quench to 32.3 kNm^3/h, bed 1 rise 29.5 K, 8.1 wppm
7	schedule off	over 10 wppm from 11.6 months, end of run at 32.8
	schedule on	7.9–8.9 wppm, end of run at 25.6 months
8	mid-ranging off	RIT 282.5 °C, 295 wppm at 240 min
	mid-ranging on	heater 7.35 to 11.26 MW, RIT back at 51 min
	heater at 60 %	RIT 299.8 °C, 90 wppm at 240 min

inlet keeps falling, 319.3, 298.2 and 282.5 °C at 45, 120 and 240 min, the sulfur reaching 295 wppm, because each kelvin lost comes back through the exchanger and a cooler bed releases less heat. With mid-ranging the heater climbs from 7.35 to 11.26 MW, the inlet sags to 329.9 °C at 40 min and is back at 335 °C by 51 min, the bypass at 49.5 %. With the heater limited to 60 % (8.4 MW) the inlet falls to 299.8 °C and the sulfur to 90 wppm. That case is infeasible by design. Lined out with the bypass shut, holding 335 °C at 200 t/h needs 10.39 MW, the limited heater holds only 294.8 °C, and only at 180 t/h does it reach 335.5 °C, so the remedy is the feed rate and a cleaning.

> **TRY IT IN THE APP** **8 · A fouled exchanger**
>
> Load the case and press Run. Touch nothing. At 30 min the exchanger fouls to 60 % of clean. Watch the bypass valve and the reactor inlet. The bypass is shut by 34.5 min and the inlet keeps falling, to 282.5 °C at 240 min. Then turn Mid-ranging On, press Reset and run again. The heater climbs to 11.26 MW and the inlet is back at 335 °C by 51 min.

8.6.2 Sulfur on a Late Analyzer

Case 4 steps the feed sulfur from 0.8 to 1.2 wt% at 30 min, and no controller can act before the analyzer's first reading over 9 wppm, at 70 min in every run, so the true sulfur peaks between 11.48 and 11.76 wppm whatever the controller, as shown in Figure 8.2b. The IAE of the true sulfur against its setpoint over 480 min is 309 for the shipped PI (1.0 K/wppm, 30 min) with 77 min over 10 wppm, and 243 for the Smith predictor at 2.0 K/wppm and 8 min with 54 min over. Both tunings are the best of their own grids, the predictor's of 36 and the PI's of 42 (1.5 K/wppm, 40 min, 298 and 65 min), each at an interior point. Against the best PI the predictor saves 18 % of the IAE and 11 min over

specification, all of it in the recovery. At seeds 1, 2 and 3 the shipped PI gives 304, 299 and 306, the best PI 297, 284 and 291, and the predictor 244, 238 and 241.

The predictor is sensitive to its dead time. With $\theta_p = 45$ min, 50 % long, the IAE is 699, over twice the PI's, and the true sulfur still swings between 5.6 and 11.3 wppm in the last hour, because each move's modeled effect is taken back 15 min after the analyzer has shown it. With $\theta_p = 15$ min it is 399. Both detuned predictors lose to the plain PI at every seed tried (583 to 724 and 379 to 412), and the curve is not symmetric. The predictor run with the PI's own tuning gives 487, so its advantage needs its own tuning.

> **+ DID YOU KNOW? A Wrong Delay Costs More Than No Predictor**
>
> From the simulator (Table 8.4). On a sourer feed the shipped PI scores an integrated error of 309 and the Smith predictor 243. Give the predictor a dead time of 45 min where the analyzer's is 30 min and the same controller scores 699, more than twice the PI. Set 15 min too short, it scores 399.

8.6.3 Feeds and Cracked Stock

Case 5 runs RIT only through a sourer feed at 30 min and 230 t/h at 240 min. The product passes 10 wppm at 40 min, reaches 11.1 and then 15.1 wppm, and spends 440 min over specification. With the Smith predictor the sourer feed takes the true sulfur to 11.56 wppm at 75.5 min, back to 8.15 by 230 min, the rate step, which arrives within a minute instead of through the drum, peaks at 10.70 wppm at 263 min, and the product is at 7.91 wppm by 480 min, 92 min over in all. The WABT climbs from 355.3 to 364.8 °C for the sulfur and to 369.8 °C for the rate, and mid-ranging takes the heater to 10.43 MW. The Sulfur PI with the predictor's gains, acting on the raw analyzer, is unstable, reaching 70.7 wppm before the rate step and 136.9 wppm after it.

Case 6 blends 20 % light cycle oil at 30 min with the inlet fixed. The aromatics of the blend rise from 25 to 33 mol%, the rises from 20.4 and 16.2 to 24.0 and 20.0 K lined out, and the outlet to 375.6 °C at 93.5 min. The override adds quench, 4.8 to 8.2 kNm^3/h by 240 min, and holds the outlet at 374.2 °C while the cooler bed 2 and the harder sulfur leave the product at 12.3 wppm (11.5 wppm with the override off). Chosen at 240 min, the Smith predictor raises the WABT, the outlet overshoots to 382.0 °C at 291 min, the quench reaches its 32.3 kNm^3/h maximum by 556.5 min and brings the outlet back to 375.2 °C, but the WABT loop puts the heat back through the inlet (349.6 °C) and bed 1's rise climbs to 29.5 K. The sulfur returns to 8.1 wppm. Under a WABT loop, quench moves heat from bed 2 to bed 1 rather than removing it.

Pause. Try Exercise 8.3 before reading on.

8.6.4 A Catalyst Cycle in a Day

Case 7 compresses two years into a day with the slow Sulfur PI. Without the schedule the product passes 10 wppm at 11.6 months and stays over for 21.2 months, 26.6 wppm at the end, and the run ends at 32.8 months only because the beds ran cooler than the product needed. With the schedule the sulfur stays between 7.9 and 8.9 wppm, the WABT climbs from 355.2 to 364.0 °C in the first year and to 383.6 °C by month 24, bed 2 falls to an activity of 0.20 against bed 1's 0.43, and bed 1's outlet reaches the 400 °C metal limit at 25.6 months, with the quench at 22.6 kNm^3/h and the heater at 13.1 MW. The compression is a teaching device, and a real unit sees the same climb over years.

8.6.5 A Graded Day and Its Margin

The challenge runs 24 h with a feed of 1.0 wt% sulfur from 4 h through the surge drum, the analyzer out of service from 10 to 12 h and the exchanger fouled to 60 % at 16 h. Its target judges the true sulfur at the reactor every 10 min, analyzer in service or not, at or under 10 wppm at every one of 144 readings and at or under 8.5 wppm over all the day's diesel, both rises at or under 30 K at every step, catalyst life used at or under 1.25 days per day, and the heater never at its limit. The variants are listed in Table 8.5. Figure 8.3 draws three columns of the table against their limits.

Table 8.5 The Challenge Over 24 h, Default Seed. Worst Reading, Readings Over 10 wppm, Mean Sulfur, Peak Rise and Catalyst Life

Settings	Worst (wppm)	Over	Mean (wppm)	Rise (K)	Life (d/d)	Verdict
shipped (slow PI, MR off)	386.97	53	86.27	35.2	0.81	fails
slow PI, MR on	10.14	5	8.83	28.4	1.11	fails
Smith 2.0/8, MR off	386.81	47	85.08	36.5	0.85	fails
PI 1.0/30, MR on	9.88	0	8.10	28.9	1.17	passes
Smith 2.0/8, MR on	9.68	0	8.08	28.8	1.17	passes
Smith 2.0/10, MR on	9.70	0	8.08	28.9	1.17	passes
Smith 2.0/10, SP 7.5 wppm[a]	9.11	0	7.58	29.2	1.21	passes
Smith 2.0/10, SP 6.5 wppm	7.94	0	6.57	30.05	1.30	fails
MPC, MR on	9.26	0	8.18	28.8	1.16	passes

[a]The suggested settings. Other rows use SP 8 wppm, except where stated. MR is mid-ranging.

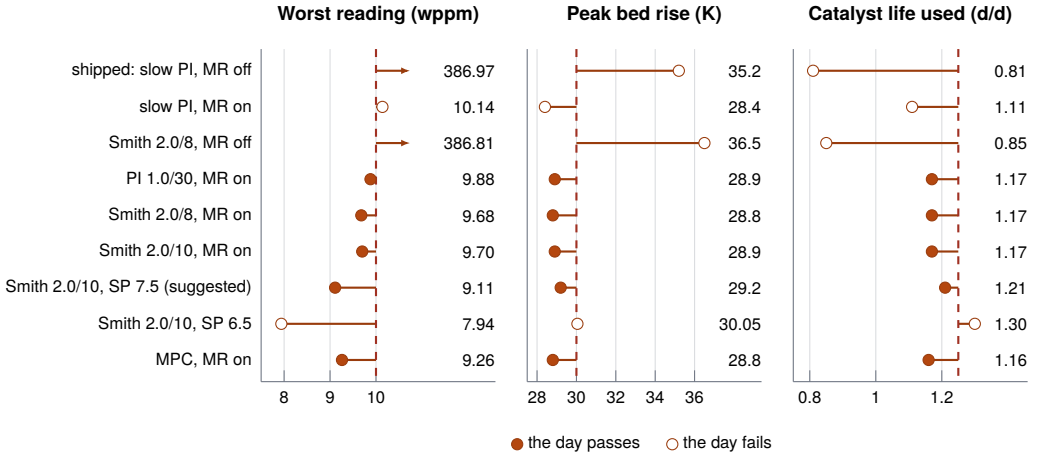


Figure 8.3 Three columns of Table 8.5 against their limits (dashed): the worst of the day's readings against 10 wppm, the peak bed rise against 30 K and the catalyst life used against 1.25 days per day. A filled circle marks a day that passes and an open circle a day that fails. Notice that the two rows without mid-ranging fail by far, that every passing row sits close under the sulfur limit, and that the 6.5 wppm setpoint fails on the two limits at the right and not on sulfur. The worst readings of 386.97 and 386.81 wppm lie off the scale. The mean sulfur and the count of readings over the limit are not drawn.

As shipped, the fouling at 16 h takes the inlet down to 281.8 °C with the heater fixed at 7.34 MW, 53 readings exceed 10 wppm, and the rise peaks at 35.2 K at 16.2 h, the transient of case 8 rather than a runaway. Mid-ranging saves the inlet, and the slow PI then lets the feed switch through, five readings over and 8.83 wppm on average. With mid-ranging, the Smith predictor at 2.0 K/wppm and 10 min passes the default 8 wppm setpoint day with a worst reading of 9.70 wppm, mean 8.08 wppm, rise 28.9 K and catalyst use 1.17 days per day. Expanding the sweep to the default seed and seeds 0–29 exposes seed 26: two readings exceed specification, peaking at 10.05 wppm at both step sizes. The suggested settings therefore use a 7.5 wppm sulfur setpoint. Their default worst reading is 9.11 wppm, mean 7.58 wppm, rise 29.2 K and catalyst use 1.21 days per day. Heater duty peaks at 12.08 MW of 14. The back-off buys quality margin with catalyst life, within the same limits. All 62 trials pass across those 31 seeds and both integration steps. Worst readings span 9.006–9.448 wppm, peak rise stays below 29.736 K and catalyst use below 1.212 days per day.

At the common 8 wppm setpoint, case 4's 8 min integral time passes the default day (9.68 wppm) but fails seed 6 (10.27 wppm, two readings over). Its setpoint-driven predictor cannot see WABT

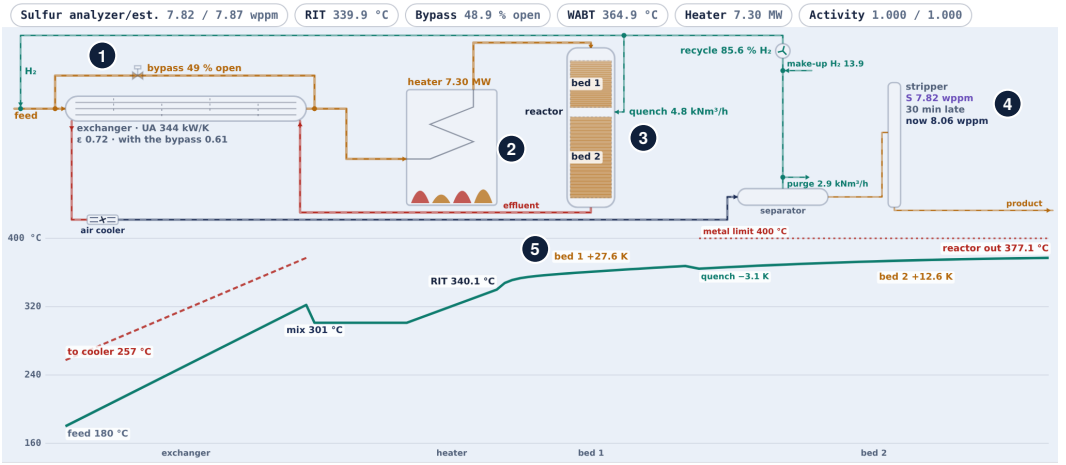


Figure 8.4 Case 4 at $t = 240$ min on the Unit view, cropped to the plant view with numbered callouts, the Smith predictor at 2.0 K/wppm and 8 min chosen before Run. The feed carries 1.2 wt% sulfur since 30 min. The bypass (1) is 48.9 % open and the heater (2) fires 7.30 MW, with the inlet at 339.9 °C and the WABT at 364.9 °C. The quench (3) enters between the beds. At the stripper (4) the analyzer reads 7.82 wppm, 30 min late, the predictor's estimate is 7.87 wppm and the true sulfur at the reactor 8.06 wppm. The temperature ladder (5) reads a mix of 301 °C, an inlet of 340.1 °C, a bed 1 rise of 27.6 K, a quench step of −3.1 K, a bed 2 rise of 12.6 K, a reactor outlet of 377.1 °C and 257 °C to the cooler, and the exchanger's effectiveness is 0.72, or 0.61 with the bypass. The recycle reads 85.6 mol% hydrogen with 13.9 kNm^3/h of make-up. In the full window, outside this crop, the trends show a true peak of 11.56 wppm at 75.5 min that the analyzer reported as 11.38 wppm at 110.5 min, and the scorecard gives a true sulfur IAE of 229 wppm min so far and a run average of 8.93 wppm.

fall short during fouling recovery. The MPC gives 9.25 to 9.45 wppm over seven seeds with the tank assay in its model. Comparing it with the 7.5 wppm suggestion would mix a setpoint change with an architecture change. At 6.5 wppm, the default worst reading falls to 7.94 wppm, but catalyst use 1.30 and rise 30.05 K fail both budgets.

The graded feed switch is 25 %. At 50 %, the original 8 wppm Smith 2.0/10, case 4 predictor, PI 1.0/30 and slow PI with mid-ranging reach 11.40, 11.37, 11.75 and 12.33 wppm before the analyzer reports it, with rises of 31.7–32.6 K. Even MPC reaches 10.08 wppm on two readings. Certification covers the stated day, not an arbitrary sour-feed envelope. The state of case 4 at 240 min under the Smith predictor is shown in Figure 8.4.

8.7 Technology in Practice

Figure 8.5 draws the loops of Section 8.5 with the hardware a refinery installs.

What Measures It

The inlet, the beds and the outlet are read by thermocouples, which are less accurate than resistance thermometers but more rugged and wider in span [19]. Their 6 s lag in Section 8.3 is short beside every other clock of the unit. The measurement that sets the pace is the product sulfur. In one industrial diesel hydrotreater it was known from a laboratory analysis made once a day [20]. The analyzer of this chapter, with a report every 10 min, is the faster instrument that a sulfur loop needs.

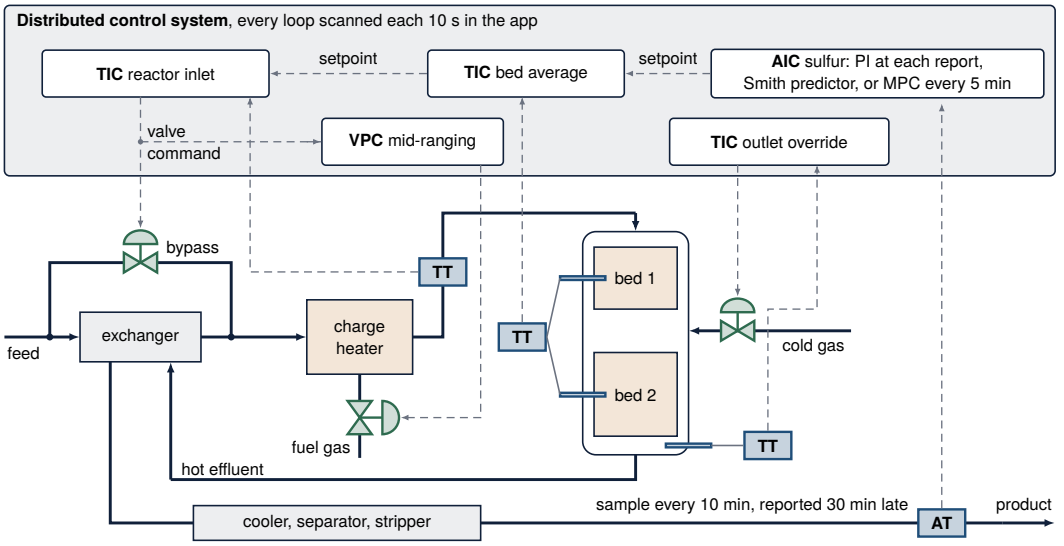


Figure 8.5 The loops of the hydrotreater with their hardware. Lettered boxes are sensing devices, thermocouples with their transmitters (TT) and the sulfur analyzer (AT). Valve symbols are final elements and dashed arrows are signals. Notice that the three controllers of the top row form one cascade from the late analyzer down to the fast bypass valve, and that the heater is moved only by the position of that valve. Not to scale, and not a piping or safety diagram.

> **TECHNOLOGY** **Sensor: Product Sulfur Analyzer**
>
> Product sulfur is defined by laboratory methods. Under the federal fuel rule every batch of ultra-low-sulfur diesel is tested before it leaves the refinery, and the referee procedure is ASTM D2622, wavelength dispersive X-ray fluorescence [21]. ASTM D5453 finds total sulfur by ultraviolet fluorescence [1]. Sulfur dioxide, formed when sulfur burns, fluoresces under ultraviolet light, and a photomultiplier tube measures the light emitted [19]. Online analyzers for product sulfur are mostly expensive, with high maintenance costs [22]. An analyzer away from the process needs a sample system, and analyzer problems quite often stem from it. With a plugged sample line the analyzer measures old, stale sample and has no way of knowing [19]. The 20 min product path and 10 min analysis of Section 8.3 stand for this sensor.

What Moves It

Three final elements carry the control. The bypass valve is the fast one, because only the heater coil stands between the mix and the inlet. The fired heater is the slow one. Such a heater may not be required in normal operation but is always needed for startup [6]. In the audited unit of the box below, every temperature controller of the reactor section had a very slow action and was retuned, one of them to more than three times its former gain [23].

> **TECHNOLOGY** **Final Element: Bypass and Quench Control Valves**
>
> Bypassing the feed and effluent exchanger is the typical way to control a reactor inlet temperature [6], and in a hydrotreater cold gas injected as quench controls the reactor temperature [23]. A control valve with excess stem friction moves in jerks, which is called stiction, and with integral action the controller output then ramps up and down as the valve slips and sticks [19]. An audit of an industrial diesel hydrotreater found this in a quench flow loop. Stiction of about 0.4 % of the valve's travel made the flow swing by 2.5 % of its range, fast enough that the reactors filtered it out of their temperatures [23]. The model's bypass and quench valves are lags of 0.1 and 0.15 min with no friction.

Where the Controller Runs

A refinery runs these loops in a distributed control system. Its loops are sometimes nested, the set point of one being set by another, and their cycle times range from milliseconds to minutes [24]. The app scans every loop each 10 s and solves its MPC every 5 min, inside that range. The sulfur, WABT and inlet controllers are three PID blocks in cascade, and the override is a fourth. Mid-ranging is a well established technique, but in a distributed control system it must handle bumpless transfer, saturation and operation with one of its two handles out of service [9]. Above the loops, a predictive controller replaces a combination of PID, lead-lag and select blocks with one controller [18].

> **IN PRACTICE** **A Leak That No Temperature Fixes**
>
> A real exchanger has a failure that the model does not carry. The feed side stands at the higher pressure, by the pressure drop of the reactor and the exchanger, so a leaking tube passes raw feed into the effluent. At a refining industry panel such leaks were called a big problem for meeting the product specification [25]. One refiner tracks a sulfur species that the reactor converts completely. Found in the product from the exchanger, it shows a leak. A quick field test raises the reactor temperature. Product sulfur that does not answer successive temperature moves points to a problem such as a leak. Tubes are seal welded to the tube sheet and inspected at turnarounds [25].

What Is Modernizing It

The first development is the soft sensor. In one refinery a neural network inferred the product sulfur of a diesel hydrotreater in real time from thirteen online measurements of flow, temperature and pressure, and was meant for the unit's advanced control [20]. The second is the audit of the loops themselves. Software that grades each loop from routine data, with the minimum variance controller as its reference, was applied to the 56 loops of an industrial diesel hydrotreater and found a tuning problem in 27 of them [23]. The third is predictive control. A 2003 survey counted more than 4600 applications, and refining held 67 % of those classified [18]. The app's MPC adds what the PI and the Smith predictor lack, the feed tank's assay as a feedforward.

8.8 Limitations

Every size, rate constant, heat and limit is a generic teaching value for a 200 t/h unit, chosen so the design point gives 8 wppm at 335 °C with rises of 20.4 and 16.2 K, and none was fitted to a catalyst or a plant. The HDS order, its activation energy and the H_2S constant were not checked against a published fit, and the hydrogen consumed, 40 Nm^3/m^3 at design, is a teaching value as well. Two lumped reactions replace the chemistry, so nitrogen, olefins and cracking appear only in the HDS heat. Twenty-four stirred cells set the size of the wrong-way rise, which has not converged, and the trickle bed's wetting, maldistribution and interphase transfer are absent. The exchanger has a single lag, the heater no tube metal, and the separator, scrubber and stripper exist only as the analyzer's delay. The catalyst cycle is compressed 720 times in its case.

The Smith predictor's and the MPC's models are fixed at the commissioned design, so their offsets learn the activity and the feed only through the analyzer, one delay late, and a wrong dead time misleads both. The WABT schedule reads a step as a ramp and overshoots it. The MPC's limits are soft and it holds its last moves on a failed solution. The challenge requires a sulfur setpoint back-off for margin across the expanded noise sweep, and its feed switch was softened to 25 % so that feedback alone can pass. The end of run reads the truth and stops the run rather than modeling a shutdown.

OPEN CHALLENGE **A Sulfur Estimate That Ages With the Catalyst**

An analyzer that reports half an hour late invites an estimate from faster measurements, a soft sensor, and such models have been built for the product sulfur of hydrotreaters [20, 22]. The open problem is keeping one true. A review of soft sensors finds that after launch the prediction usually deteriorates because the process drifts, with catalyst deactivation among the causes, until the model must be retrained or rebuilt. Most soft sensors have no automated mechanism for their own maintenance, and telling a process drift from a sensor drift takes expert knowledge [26]. The predictor and the MPC of this chapter carry fixed models too. Progress would be an estimate that follows the catalyst through its cycle and reports its own error.

8.9 Conclusion

A ULSD hydrotreater is a reactor that heats its own feed and is graded half an hour late, and this simulator computes its 48 cell temperatures, its exchanger, heater and hydrogen loop, its analyzer and its catalyst activity together. At design the exchanger returns 21.6 MW against 7.34 MW of firing, and each kelvin at the inlet comes back as 0.73 K, so 1 MW of heater ends as 12.8 K at the inlet while the bypass moves it within minutes. A 10 K inlet cut first raises the outlet 1.10 K, a size that grows with the cell count while its sign and cause do not. On a sourer feed the Smith predictor saves about a fifth of the IAE against the best PI, 243 against 298 wppm min, and loses to it with a dead time 50 % wrong either way. Mid-ranging keeps a fouled exchanger from taking the inlet from 335 to 282.5 °C, a WABT schedule holds the product within 7.9 to 8.9 wppm across a compressed cycle, and the graded day needs a 7.5 wppm sulfur setpoint to preserve margin over the expanded noise sweep while respecting temperature and catalyst-life budgets. The kinetics are teaching values and the beds are stirred cells. Further work would add a plug-flow bed with axial dispersion fitted to a real one, kinetics fitted to a published ULSD catalyst, a separator and stripper model, and a laboratory sample that corrects the analyzer. The simulator shows that a hydrotreater's inlet, its beds and its sulfur loop have to be designed as one loop.

KEY POINTS

- A hydrotreater heats its own feed, so each kelvin at the reactor inlet comes back as 0.73 K and a heater move grows for hours.
- The bypass moves the inlet within minutes, and mid-ranging lets the slow heater return that valve to mid-range.
- A 10 K inlet cut first raises the outlet 1.10 K, a size that grows with the cell count while its sign and cause do not.
- No controller acts before the analyzer's first report, and a Smith predictor with a dead time 50 % wrong loses to the PI.
- The graded day needs a 7.5 wppm sulfur setpoint, a margin bought with catalyst life.

Where else this applies. Mid-ranging returns in Volume 4, Chapter 10 and Volume 4, Chapter 9, a predictor of the Smith family in Volume 5, Chapter 5, and a loop on a slow analyzer in Volume 5, Chapter 3.

References

[1] U.S. Environmental Protection Agency, "Control of air pollution from new motor vehicles: Heavy-duty engine and vehicle standards and highway diesel fuel sulfur control requirements," *Federal Register*, vol. 66, pp. 5002–5193, Jan. 2001, document 01-2, published January 18, 2001.

[2] European Parliament and Council of the European Union, "Directive 2003/17/EC of 3 March 2003 amending directive 98/70/EC relating to the quality of petrol and diesel fuels," Official Journal L 76, pp. 10–19, Mar. 2003, 22 March 2003.

[3] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[4] J. M. Douglas, J. C. Orcutt, and P. W. Berthiaume, "Design and control of feed-effluent, exchanger-reactor systems," *Industrial & Engineering Chemistry Fundamentals*, vol. 1, no. 4, pp. 253–257, 1962.

[5] J. L. Silverstein and R. Shinnar, "Effect of design on the stability and control of fixed bed catalytic reactors with heat feedback. 1. concepts," *Industrial & Engineering Chemistry Process Design and Development*, vol. 21, no. 2, pp. 241–256, 1982.

[6] F. Reyes and W. L. Luyben, "Steady-state and dynamic effects of design alternatives in heat-exchanger/furnace/reactor processes," *Industrial & Engineering Chemistry Research*, vol. 39, no. 9, pp. 3335–3346, 2000.

[7] P. S. Mehta, W. N. Sams, and D. Luss, "Wrong-way behavior of packed-bed reactors: 1. the pseudo-homogeneous model," *AIChE Journal*, vol. 27, no. 2, pp. 234–246, 1981.

[8] B. J. Allison and A. J. Isaksson, "Design and performance of mid-ranging controllers," *Journal of Process Control*, vol. 8, no. 5-6, pp. 469–474, 1998.

[9] N. Alsop, "Implementing mid ranging in a DCS environment," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 550–555, 2016.

[10] O. J. M. Smith, "Closer control of loops with dead time," *Chemical Engineering Progress*, vol. 53, no. 5, pp. 217–219, 1957.

[11] E. Furimsky, "Deactivation of hydroprocessing catalysts," *Catalysis Today*, vol. 52, no. 4, pp. 381–495, 1999.

[12] A. C. Byrns, W. E. Bradley, and M. W. Lee, "Catalytic desulfurization of gasolines by cobalt molybdate process," *Industrial & Engineering Chemistry*, vol. 35, no. 11, pp. 1160–1167, 1943.

[13] H. Topsøe and B. S. Clausen, "Importance of Co-Mo-S type structures in hydrodesulfurization," *Catalysis Reviews*, vol. 26, no. 3-4, pp. 395–420, 1984.

[14] H. Topsøe, B. S. Clausen, and F. E. Massoth, "Hydrotreating catalysis," in *Catalysis: Science and Technology*, J. R. Anderson and M. Boudart, Eds. Berlin, Heidelberg: Springer, 1996, vol. 11, pp. 1–269.

[15] S. T. Sie, "Reaction order and role of hydrogen sulfide in deep hydrodesulfurization of gas oils: consequences for industrial reactor configuration," *Fuel Processing Technology*, vol. 61, no. 1-2, pp. 149–171, 1999.

[16] T. C. Ho and R. Aris, "On apparent second-order kinetics," *AIChE Journal*, vol. 33, no. 6, pp. 1050–1051, 1987.

[17] V. Pinjala, Y. C. Chen, and D. Luss, "Wrong-way behavior of packed-bed reactors: II. impact of thermal dispersion," *AIChE Journal*, vol. 34, no. 10, pp. 1663–1672, 1988.

[18] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[19] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[20] L. Salvatore, M. B. de Souza Júnior, and M. C. M. M. Campos, "Design and implementation of a neural network based soft sensor to infer sulfur content in a Brazilian diesel hydrotreating unit," *Chemical Engineering Transactions*, vol. 17, pp. 1389–1394, 2009.

[21] U.S. Environmental Protection Agency, "Regulation of fuels, fuel additives, and regulated blendstocks," Code of Federal Regulations, Title 40, Part 1090, sections 1090.95, 1090.305, 1090.1310, 1090.1350 and 1090.1360, edition of July 1, 2024, 2024, govinfo.gov.

[22] S. Shokri, M. T. Sadeghi, M. A. Marvast, and S. Narasimhan, "Improvement of the prediction performance of a soft sensor model based on support vector regression for production of ultra-low sulfur diesel," *Petroleum Science*, vol. 12, no. 1, pp. 177–188, 2015.

[23] L. G. S. Longhi, A. S. Reginato, H. C. G. Teixeira, C. A. C. Cortez, L. P. Lusa, S. S. Gonzalez, and T. D. Fleck, "Control loop performance assessment and improvement of an industrial hydrotreating unit and its economical benefits," *Sba: Controle & Automação Sociedade Brasileira de Automática*, vol. 23, no. 1, pp. 60–77, 2012.

[24] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[25] American Fuel & Petrochemical Manufacturers, "Question 15: ULSD reactor feed/effluent heat exchanger leaks can be a big problem meeting product specifications. what are best practices for detecting and preventing leaks?" AFPM Q&A Database, panel answers by a refiner, a licensor and suppliers, undated, afpm.org, question and answer search.

[26] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

8.10 Exercises

8.1 Concept check · 2 min

In case 2 the inlet loop cuts the reactor inlet from 335 to 325 °C at 10 min. What does the reactor outlet do over the next twenty minutes?

(a) It rises by about 1 K, because the cooler first cells pass unreacted aromatics and sulfur to catalyst still at its old temperature

(b) It falls by about 10 K within minutes, because the oil and gas cross each bed in only a few minutes

(c) It rises by about 1 K, because the inlet loop opens the bypass and the cold feed starves the exchanger of heat

(d) It holds still for twenty minutes, because the catalyst's heat capacity delays every change at the outlet

8.2 Concept check · 2 min

In case 4 the feed sulfur steps from 0.8 to 1.2 wt% at 30 min. The shipped PI, the best PI found and the Smith predictor all let the true sulfur peak between 11.48 and 11.76 wppm. Why does the predictor not lower this peak?

(a) Its gain of 2.0 K/wppm is too low to answer a step of half the feed's sulfur

(b) Its model's input is the controller's own output, so the feed change reaches it only with the analyzer

(c) The surge drum spreads the step over 20 min, longer than the predictor's two 10 min lags

(d) The sulfur layer holds its integral while the analyzer is stale, so the predictor's moves are lost

8.3 Concept check · 1 min

In case 6 the quench override holds the reactor outlet after a light cycle oil blend, and at 240 min the Smith predictor takes over the WABT to bring the sulfur back. Where does the extra heat end up?

(a) In the quench gas, which carries it out of the reactor, so both bed rises fall

(b) In the product, whose sulfur settles above the 10 wppm specification

(c) In bed 1, because the WABT loop raises the inlet as the quench cools bed 2

(d) In the exchanger, which returns it to the feed and lowers the heater duty

8.4 Theory and calculation · 6 min

At design the charge carries $\dot{C}_p = 186.7\,\mathrm{kW/K}$, the exchanger and its bypass return $\varepsilon_{unit} = 0.613$ of the outlet's temperature above the feed, and a kelvin more at the inlet makes 1.187 K more at the outlet.

(a) What inlet rise does 1 MW more in the heater give at once, before the exchanger returns any of it? *Checkpoint:* 5.36 K.

(b) Compute the loop gain L and the amplification $1/(1-L)$ of Equation (1.8). What final inlet rise do they predict for the 1 MW step? *Checkpoint:* $L = 0.728$, 3.67, 19.7 K.

(c) Case 1 ends 12.83 K higher after 400 min. The loop gain is 0.899 at a 320 °C inlet and 0.431 at 350 °C. Explain the gap, and bound the amplification over the range the inlet crosses. *Checkpoint:* 2.40 actual, between 1.76 and 3.67.

8.5 Theory and calculation · 6 min

The HDS rate of Equation (8.1) is written on the fraction $X = w_S/w_{S0}$ of the feed's sulfur left, with order n.

(a) Using Equation (8.2), compute n for a gamma shape $m = 1.54$, and show that $m = 1$ gives a second-order lump. *Checkpoint:* $n = 1.65$.

(b) Hold the bed temperature and the hydrogen and H_2S pressures fixed. Show that X at the outlet does not depend on w_{S0}, and predict the product sulfur for a 1.2 wt% feed if a 0.8 wt% feed gives 8.00 wppm. *Checkpoint:* 12.0 wppm.

(c) At a fixed 335 °C inlet the simulator gives 10.92 wppm, and the order written on the absolute sulfur gave 7.32 wppm. Explain both departures from part (b). *Checkpoint:* The extra sulfur's heat. The absolute form inverts the trend.

8.6 Theory and calculation · 7 min

Bed 1 enters at 335 °C and rises 20.4 K, bed 2 rises 16.2 K, the catalyst shares are 0.35 and 0.65, $\dot{C}_p = 186.7\,\mathrm{kW/K}$, and the quench gas enters at 65 °C with $c_{p,g} = 32$ J/(mol K). Use Equation (8.4).

(a) Compute bed 2's inlet with the design quench of 60 mol/s, and the step the quench makes. *Checkpoint:* 352.44 °C, 2.96 K.

(b) Compute the WABT at design. *Checkpoint:* 355.17 °C.

(c) The quench opens to its 400 mol/s maximum. Holding both rises fixed, compute bed 2's inlet at the design inlet, and the reactor inlet a WABT loop needs to restore the WABT of part (b). *Checkpoint:* 336.77 °C. An inlet of 345.6 °C.

8.7 App challenge · 8 min

Load scenario 1, "Two handles: heater and bypass (Manual)", at the default seed. Before pressing Run, set Heater duty to 8.35 MW. The control moves in steps of 0.05 MW, so this is the setting nearest to 1 MW above the 7.34 MW the case starts from.

(a) Run to 400 min at speed 60 and export the CSV. Read `truth_rit_C` at 0, 30 and 400 min. How far has the inlet risen at 30 min and at 400 min, and what multiple of the heater's direct effect, $\Delta Q/\dot{C}_p$, is the final rise? *Checkpoint:* 335.00, 340.27 and 347.90 °C: 5.27 K and 12.90 K, 2.39 times 5.39 K (±0.02 K).

(b) Press Reset (the heater stays at 8.35 MW). Run at speed 5, pause at 30.0 min, set Bypass valve to 60 % and run to 32 min. Read `truth_rit_C` at 30, 31 and 32 min. How far does the inlet fall in 2 min, and what share of that comes in the first minute? *Checkpoint:* 340.27, 338.18 and 337.71 °C: 2.56 K, about 81 % in the first minute.

8.8 App challenge · 8 min

Load scenario 4, "Sulfur on a late analyzer: PI against Smith", at the default seed and speed 120. Each run is 480 min. After each run, read the Sulfur IAE, true tile and export the CSV.

(a) Run the case as shipped (Sulfur PI, 1.00 K/wppm, 30 min). Then set the Sulfur controller gain to 1.50 K/wppm and its integral time to 40 min, press Reset and run again. Then choose Smith predictor with 2.00 K/wppm and 8 min, press Reset and run a third time. For each run, read the IAE and count the minutes in which `truth_S_reactor_ppm` is above 10 (rows are 0.5 min apart). *Checkpoint:* 309 and 77 min, 298 and 65 min, 243 and 54 min (±2 and ±1 min).

(b) What share of the better PI's IAE does the predictor save? When does the true sulfur first exceed 10 wppm in each run, and why is that time the same? *Checkpoint:* About 18 %. At 40.5 min in all three.

8.9 App challenge · 9 min

Load scenario 9, "Challenge: a day on the hydrotreater", at the default seed and speed 240. The run stops at 24 h. Read the three So far tiles (the sulfur tile's hover text gives the readings over 10 wppm).

(a) Run it as shipped. Which limits fail, and what happened at 16 h? *Checkpoint:* Worst 386.97 wppm, mean 86.27 wppm, life 0.81 days per day, peak rise 35.2 K.

(b) Turn Mid-ranging On, press Reset and run the day again. Which limits fail now? *Checkpoint:* Worst 10.14 wppm, mean 8.83 wppm, life 1.11, rise 28.4 K: five readings over.

(c) Choose Smith predictor, set Sulfur setpoint to 7.5 wppm, the gain to 2.00 K/wppm and the integral time to 10 min, press Reset and run the day again. *Checkpoint:* Worst 9.11 wppm, mean 7.58 wppm, life 1.21, rise 29.2 K: passes.

8.10 Application area · 4 min

A unit's feed and effluent exchanger has fouled to 60 % of its clean heat transfer, and burners out of service limit the charge heater to 60 % of its 14 MW rating. Holding the inlet at 335 °C at 200 t/h with the bypass shut needs 10.39 MW. What should the console do now, and why should exchanger cleaning be planned from the heater's duty rather than from the bypass position?

8.11 Application area · 4 min

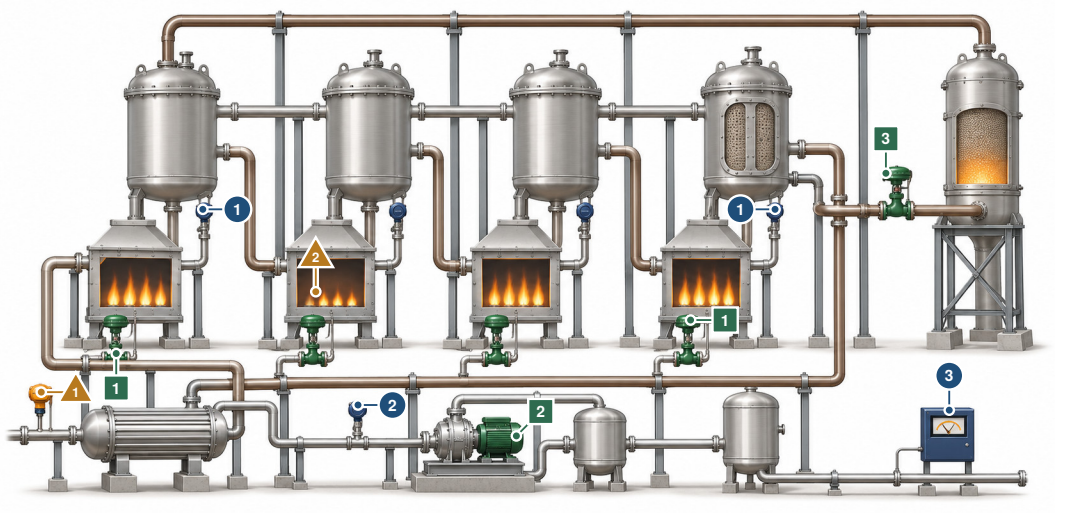
The challenge's suggested settings hold a 7.5 wppm sulfur setpoint against a 10 wppm limit. Using Table 8.5 and the seed sweep, say what the back-off from 8 to 7.5 wppm buys and costs, why 6.5 wppm is not better, and what in the plant sets the size of the back-off.

8.12 Review the agent · 5 min

An AI assistant reviewing the hydrotreater writes: "The sulfur analyzer's delay is 30 min, but sample lines foul and slow down. Set the Smith predictor's dead time to 45 min to be safe. Overestimating a dead time is conservative, just as detuning a PI is, so the worst case is a loop a little slower than the PI." Find the flaw and say how you would check it in the app.

CHAPTER 9

Interheater Coordination and Inferred Octane on a Continuous Catalytic Reformer



The controller moves 1 fuel valve of each of the four heaters 2 drive of the recycle compressor 3 catalyst circulation to the regenerator

The controller reads 1 temperature at each bed inlet 2 recycle gas flow 3 reformate octane analyzer

What disturbs it 1 a leaner naphtha 2 a heater short of burners ▲ an analyzer that reads high

Each heater warms the gas on its way into the reactor above it, and the one octane analyzer stands at the far end of the product line at the lower right, after the compressor and the two drums. Illustration, not to scale.

IN PLAIN TERMS Four Heaters and One Late Octane Test

The main idea. A reformer rebuilds the molecules of a light oil so that it resists knocking in an engine, a quality called octane. Its four beds of hot grains do work that soaks up heat, so a heater must warm the stream before each bed.

The challenge. The first bed swallows most of any extra heat, so a push on one heater fades along the line. The octane is tested late and far downstream, and the tester can read high without anyone knowing.

Why it matters. Running too hot turns good gasoline into gas, and running too cool makes fuel that falls short of its grade.

THREE WAYS TO READ THIS CHAPTER

⊘	**Explore**	13 min	This page, the History boxes in Section 9.2, the Did you know? boxes in Sections 9.1 and 9.6, the Try it boxes in Sections 9.4 and 9.6, and Key points in Section 9.9. No equations.
🕮	**Study**	50 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
⛰	**Practice**	15 min	This page, Section 9.6 with Figure 9.4, Section 9.7 and Section 9.8 with its Open challenge box.

The control question A reformer's beds absorb the heat its interheaters add, and its octane is measured late, rarely or by inference. Which layers of control hold that octane through a leaner naphtha, a heater short of duty and an analyzer that drifts, and at what cost in yield?

Learning objectives

- Compute each bed's reaction heat, each heater's tube metal and the regenerator's burn from the drops, the duties and the catalyst circulation.
- Explain why the first bed absorbs most of a heater step and passes only part of it on.
- Price an octane number in C5+ yield and find the WAIT above which more severity stops paying.
- Compare an octane trim on a late analyzer with one on an inferred octane that validates the analyzer.
- Choose the feedforward, inlet distribution and octane measurement that carry the unit through a graded day of upsets.

Prerequisites Sections 1.4.2, 1.4.4, 1.4.5, 1.4.6, 1.6.2 and 1.6.3, and Chapter 7 for adiabatic beds in series.

The app https://apmonitor.com/apps/catalytic-reformer/

See also Volume 5, Chapter 10 for a soft sensor corrected by a delayed laboratory and Volume 5, Chapter 3 for composition control through a slow analyzer. Volume 9, Chapter 4 for offset-free MPC and Volume 9, Chapter 8 for an optimum that sits on a moving constraint.

9.1 Introduction

A catalytic reformer cools its own feed. In the generic unit simulated here, 150 t/h of naphtha carrying 21 vol% of naphthenes and 20.5 vol% of aromatics, with recycle hydrogen at 3 mol per mole of hydrocarbon, enters the first of four catalyst beds at 500 °C and leaves it 70.4 K colder. The second, third and fourth beds take 45.1, 11.8 and 7.2 K, so four fired heaters must put back 21.6 MW and the gas temperature along the train traces a staircase. Lifting every bed inlet by 10 K raises the research octane number (RON) of the reformate by 2.01 and lowers its liquid yield of molecules with five carbons and more (C5+) by 1.51 points of the feed's mass once the unit lines out.

The unit matters to the whole refinery. Reforming exists to upgrade low-octane naphtha into gasoline blending stock, to make aromatics for petrochemical plants and to supply hydrogen [1], and a commercial reformer is a set of heater and reactor pairs in series working between about 400 and 500 °C, followed by a flash drum whose gas is partly recycled and a stabilizer whose bottoms is the reformate [2]. A catalyst that cokes, a hydrogen ratio that slows the coking, a regenerator that burns the coke off and an octane that the laboratory confirms only a few times a day make the reformer a plant whose operation is governed as much by measurement and constraints as by its chemistry [3, 4].

> ✦ DID YOU KNOW? **A Third of the Gasoline**
>
> On January 1, 2003 the refineries of the United States could reform 3,512,237 barrels per calendar day, 21.1 % of their crude distillation capacity, and the reformer produced around 30 to 40 % of all the gasoline made in the country [5]. The unit of this chapter reforms 150 t/h of naphtha.

The simulator corrects three beliefs that a block diagram leaves intact. The first is that a heater step reaches the bed it feeds. In case 1 heater 1 receives 1.6 MW more, about 10 K at its outlet, and by 120 min its outlet has risen 15.7 K because the warmer effluent comes back through the feed–effluent exchanger, yet bed 1 absorbs 9.3 K of that in a larger drop and hands only 6.3 K to bed 2. The second is that an octane loop on an online analyzer holds the octane being made. The near-infrared (NIR) analyzer here reports the stabilizer bottoms about 20 min after the reactors, and when its calibration shifts by 0.8 RON a trim that trusts it lowers the true octane to 94.24 and keeps it more than 0.3 under target for 3.2 h. The third is that more severity always buys more octane. Lined out at a hydrogen ratio of 3, the regenerator reaches its burn capacity at a weighted average inlet temperature (WAIT)

of 523.5 °C, after which coke returns to the first bed and the octane gained per kelvin falls from 0.23 to about 0.1.

Each of these behaviors has relatives in other plants. A train of adiabatic beds with heating between them is a distributed temperature problem in which each bed's outlet is the next heater's disturbance and the inlet profile, not one temperature, sets the product. Endothermic beds damp what they receive, where the exothermic beds of Chapter 7 amplify it (Section 1.4.4). A quality that is sampled slowly and late while the inputs move fast is the dual-rate problem that inferential control addresses [4], met again in the milk powder moisture of Volume 5, Chapter 10, and a supervisory optimizer that trades a product property against yield inside equipment limits is the refinery setting in which model predictive control is most widely applied [6]. A slowly deactivating catalyst whose activity a regenerator restores adds a slow time scale, months between regenerations in a semiregenerative unit [3, 7] and hours in this compressed simulator, that the regulatory loops never see.

However, the reforming literature treats these pieces apart. Kinetic models of industrial reformers lump the naphtha into paraffins, naphthenes and aromatics, or into a few dozen pseudocomponents, and fit the reactor train to plant data [3, 8], optimization studies follow a semiregenerative unit over its catalyst life [3], and control studies address the heaters, the octane or the optimizer one at a time [4, 9], while a reactor temperature loop usually reaches a student as a transfer function whose output is measured exactly. This chapter works on a dynamic simulator in which the bed temperatures, the coke on each bed, the analyzer, the laboratory and an inferred octane are computed together. It shows how heater loops, a WAIT coordinator, an octane trim with a feed-quality feedforward and a model predictive controller (MPC) hold a reformer's octane through a leaner naphtha, a heater short of duty, a recycle compressor turndown, a stopped catalyst lift and a drifting analyzer, and it records each equation and constant of the reformer model so that the model can be reproduced.

The lumped kinetics and the controllers are established practice, and the octane correlation is an illustrative teaching value. The results are measurements made on a single generic unit with the default random seed, repeated at three further seeds wherever noise could change the outcome. Two of them carry the chapter. One is the burn-capacity ceiling on octane above, and the other is the finding that every strategy on the analyzer fails the graded challenge, because the shifted reading persists for longer than the challenge allows whatever the controller does. The model has forty-three states, and its mass, carbon and hydrogen balances close at the reactor boundary to 10^{-15}. Its bed module carries no chemistry and is shared with the bed reactors of other chapters, among them Chapter 8, and a measurement boundary keeps the true octane, coke and activity away from every controller. Each of the nine cases has a completion rule that grades it (Section 1.6.4).

Section 9.2 follows platinum reforming from its patents to the moving bed. Section 9.3 gives the model, Section 9.4 its behavior before any controller, and Section 9.5 the controllers, the estimator and the optimizer as they are coded. Section 9.6 measures the nine cases, Section 9.7 sets the hardware of a real unit beside the model, Section 9.8 states what the model leaves out, and Section 9.9 draws the results together.

9.2 Platinum, Rhenium and the Moving Bed

The process took its modern form in a patent filed in November 1947 and granted in August 1949, whose content the History box below gives. Two years later Haensel and Sterba compared the new Platforming process with thermal reforming, then in commercial operation at a Pennsylvania refinery, by the relation between yield and octane number, the product properties and the hydrocarbon types [10]. The process crossed the Atlantic quickly, and in 1952 CEPSA became the first European company to license Platforming, for its refinery on Tenerife [11].

> ⧖ HISTORY **1947: Platinum Reforming Is Patented**
>
> Vladimir Haensel filed the patent in November 1947 and received it in August 1949 [12]. It defines reforming as the treatment of gasoline fractions to improve their antiknock rating and lists what the catalyst must do, which is to dehydrogenate naphthenes to aromatics, to close straight-chain paraffins into aromatics and to crack selectively. It gives the high cost of the catalyst as the reason platinum had found limited commercial acceptance and describes catalysts with as little as 0.01 to 1 wt% of platinum on a halogen-containing alumina. It also names the problem a regenerator exists for, coke that deposits on the catalyst, shortens the processing cycle and forces either a burn or a new charge. One test raised a Pennsylvania naphtha of motor octane 41.2 to 64.2 at 872 °F and 500 psi. In 1997 the National Academy of Engineering awarded Haensel the Charles Stark Draper Prize for catalytic reforming [13].

The second catalyst came from Chevron. Kluksdahl's patent, granted in December 1968, claimed platinum with rhenium for reforming sulfur-free naphtha [14]. Its account of the prior art is a compact statement of the trade that runs through the present chapter. Higher octane comes from higher temperature or lower space velocity while the gasoline yield suffers, reforming catalysts are rated by their yield at the desired octane, and hydrocracking, being exothermic where reforming is endothermic, can drive a bed toward runaway. In one of its examples the temperature was adjusted throughout the run to hold a 100-octane product, and between 200 and 845 h on a sulfur-free feed the temperature required rose by about 0.024 °F per hour, while in a simulated life test of a platinum catalyst the yield at the desired octane fell steadily with time. That constant-octane schedule is what case 8 reproduces on a compressed clock.

The regeneration mode separates the generations of reformers, as the History box below tells.

> ⧖ HISTORY **Late 1960s: Regeneration Without a Shutdown**
>
> A semiregenerative unit keeps its catalyst in fixed beds and shuts down for an in-place regeneration every six months to two years. A cyclic unit regenerates on a cycle of a few weeks to a few months and suits poor feeds. A continuous unit regenerates a fixed amount of catalyst at a time, at a rate set by its catalyst circulation [7]. The continuous reformer was introduced in the late 1960s. Its catalyst moves down through the reactors as a column of particles, and small quantities are withdrawn, carried to a regeneration unit and returned [13]. In 1976 this was described as the most recent innovation in catalytic reforming, valued for continuous and stable production of gasoline and hydrogen [15]. Case 8 of this chapter stops the catalyst lift, and the WAIT for constant octane then climbs as it does in a semiregenerative unit.

By 2004 the technology had moved from semiregenerative fixed beds toward continuous units for their efficiency, reliability and operating cost [1]. Optimization of a semiregenerative unit over a whole catalyst life, at constant inlet temperatures or at constant octane, is the problem that continuous regeneration largely removes [3].

Control followed the computers. A reformer controller using nonlinear programming with a penalty-function gradient method for constraints, first run in a process computer, was configured in 1983 in the function blocks of a microprocessor distributed control system [9]. The octane itself resisted automation. The knock engine is slow, and online analyzers infer octane from spectra instead, as near-infrared features were shown to do for gasolines in 1989 [16]. In a pilot study of a platinum–rhenium catalyst the RON of the reformate was linear in its aromatics content [17], the relation the simulator's correlation borrows. As late as 2003 octane quality control on Shell Canada's continuous reformers was typically manual because the RON was measured so seldom, and a dual-rate inferential MPC built for those units cut the octane variance by 40 % [4].

9.3 Four Beds, Four Heaters and a Moving Catalyst

The simulated unit is a generic low-pressure continuous reformer, as shown in Figure 9.1. Naphtha at 150 t/h joins the recycle gas, is heated by the reactor effluent in the feed–effluent exchanger, and

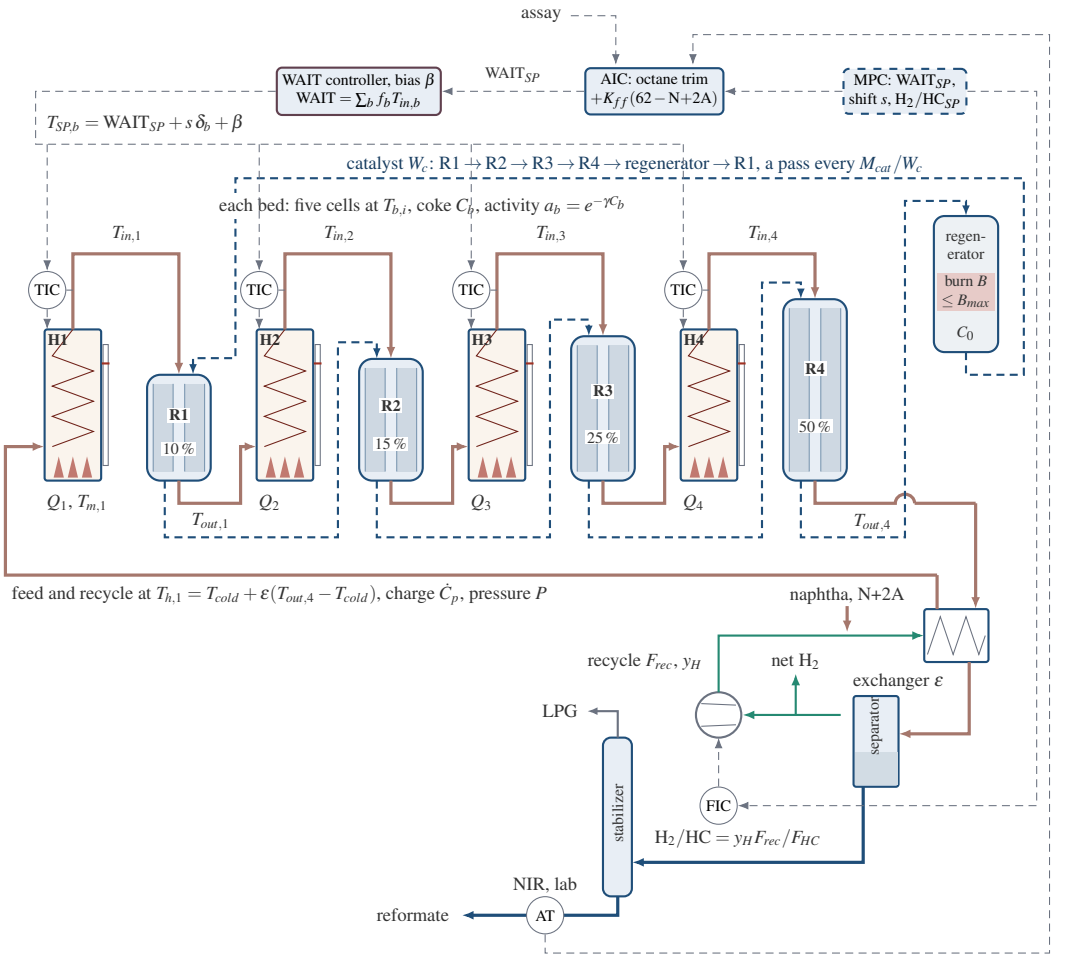


Figure 9.1 Continuous catalytic reformer as simulated, drawn after the wide layout of the application's Unit view, with the product section moved under the last beds. The exchanger delivers feed and recycle at $T_{h,1}$ to heater 1, each heater b with duty Q_b and tube metal $T_{m,b}$ (gauge, with a tick at the 620 °C limit) raises the gas to the bed inlet $T_{in,b}$, and each bed of five cells at $T_{b,i}$ returns it at $T_{out,b}$ to the next heater. Catalyst (dashed) moves from bed to bed at W_c, carrying coke C_b, to a regenerator that burns $B \le B_{max}$ and returns catalyst with coke C_0 to bed 1. The compressor sets the recycle F_{rec} of purity y_H and so H_2/HC against the hydrocarbon feed $F_{HC} = F_P + F_N + F_A$. Four TIC loops hold $T_{SP,b}$ from the WAIT controller, the AIC octane trim on the NIR, laboratory or inferred octane with an N+2A feedforward sets WAIT_{SP}, and the MPC (dashed block) can set WAIT_{SP}, the severity shift and the H_2/HC setpoint. Signals are thin dashed lines. The drawing represents the model, not piping or safety design, and vessel heights follow the cube root of the catalyst.

passes heater 1 (the charge heater), bed 1, heater 2, bed 2, heater 3, bed 3, heater 4 and bed 4 in series at 4.5 bar. Each bed is a radial-flow reactor holding 10, 15, 25 or 50 % of 66 t of catalyst. The effluent gives up its heat in the exchanger and reaches a separator, whose gas the recycle compressor returns to the feed while the net hydrogen leaves, and whose liquid a stabilizer turns into reformate. The catalyst moves down through the four beds in turn at $W_c = 8.25\,\text{t/h}$, is lifted to a regenerator that burns its coke, and returns to bed 1. The clock is the minute, flows inside the beds are moles per second, duties megawatts and coke weight percent of the catalyst.

Each variable in Table 9.1 has a role in the loops. Forty-three states are integrated. Thirty-eight

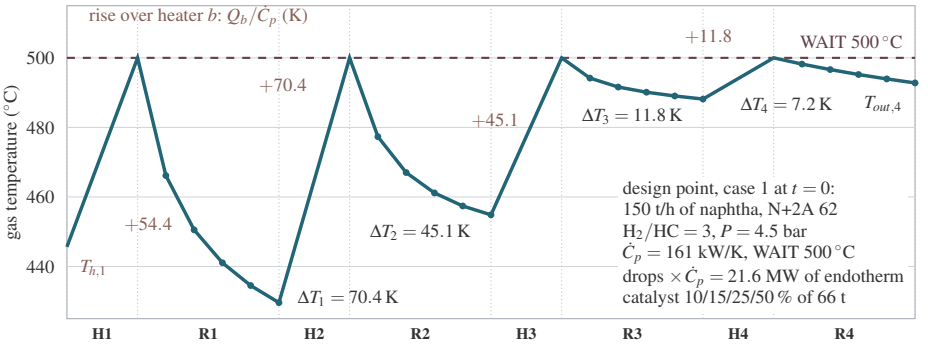


Figure 9.2 The gas temperature along the train at the design point, drawn to scale on the horizontal slots of the heaters and beds of Figure 9.1, with the five cell temperatures of each bed as dots and the WAIT dashed. Each heater raises the gas by $Q_b/\dot{C}_p$ and each bed cools it by its drop ΔT_b.

Table 9.1 Reformer Variables, Their Roles and Design Values (Case 1 at $t = 0$)

Variable	Role	Design	Range
RON	CV, reformate octane	95.0	SP 90–102
WAIT	CV and MV, severity	500.0 °C	SP 470–545 °C
$T_{in,b}$	CV, bed inlets	500 °C each	—
H_2/HC	CV, recycle ratio	3.00 mol/mol	SP 1.5–4.5, floor 2.5
Q_b	MV, heater duties	8.72, 11.30, 7.25, 1.90 MW	0–12, 15, 10, 5
N	MV, compressor speed	84.7 %	0–100 %
W_c	MV, catalyst circulation	8.25 t/h	0–16
N+2A	DV, feed quality	62 (N 21, A 20.5)	48 after a switch
$\dot{m}_f$	DV, feed rate	150 t/h	90–170
a_{H2}	DV, heater 2 available	100 %	30–100 %
N_{cap}	DV, compressor limit	100 %	40–100 %
b_{NIR}	DV, NIR calibration	0	±3 RON
$T_{b,i}$	states, 20 cells	see Figure 9.1	—
C_1–C_4	states, coke	0.21, 0.49, 1.35, 2.99 wt%	—
a_1–a_4	outputs, activity	0.98, 0.94, 0.85, 0.70	—
C_0	state, regenerated coke	0.10 wt%	—
y_H	state, purity	88.9 mol%	—
ΔT_b	outputs, drops	70.4, 45.1, 11.8, 7.2 K	—
$T_{m,b}$	outputs, tube metal	580, 583, 580, 542 °C	limit 620, trip 640
B	output, burn	239 kg/h	capacity 400
Y_{C5+}	output, C5+ yield	94.7 wt% (91.7 vol%)	—
A_{vol}	output, aromatics	44.2 vol%	—

are physical, the temperatures of the twenty cells (five per bed), the four heater duties Q_b, the coke on each bed C_b and on the regenerated catalyst C_0, the compressor speed N, the recycle purity y_H, the naphthenes x_N and aromatics x_A of the feed reaching the reactors, the catalyst circulation W_c, and two lags each for the octane and the yield of the reformate. Five are ledgers that integrate the feed, the C5+ liquid, the net hydrogen, the fuel and the coke burned.

9.3.1 Four Lumped Reactions

The chemistry is Smith's four-reaction lumping on paraffins P, naphthenes N and aromatics A, as reproduced in reformer simulations that represent each group by a single compound [8], with every lump a C7 and a light-ends lump L of equal moles of C1 to C5. The rates per kilogram of catalyst are

as shown in Equation (9.1).

$$\begin{aligned} \mathrm{N} \rightleftharpoons \mathrm{A} + 3\mathrm{H_2} &: r_1 = k_1 a\,(p_N - p_A p_H^3 / K_1) \\ \mathrm{N} + \mathrm{H_2} \rightleftharpoons \mathrm{P} &: r_2 = k_2 a^m (p_N p_H - p_P / K_2) \\ \mathrm{P} + \tfrac{4}{3}\mathrm{H_2} \rightarrow \tfrac{7}{3}\mathrm{L} &: r_3 = k_3 a\, p_P p_H / P \\ \mathrm{N} + \tfrac{7}{3}\mathrm{H_2} \rightarrow \tfrac{7}{3}\mathrm{L} &: r_4 = k_4 a\, p_N p_H / P \end{aligned} \tag{9.1}$$

where $p_i = P F_i / F_t$ is the partial pressure in bar at the reactor pressure P, a the bed's activity and $m = 2$. Dehydrocyclization of a paraffin runs as r_2 backward and then r_1, and ring closure needs both the metal and the acid function, so it alone takes the square of the activity. The rate and equilibrium constants follow Arrhenius and van 't Hoff with the same heats that enter the energy balance, as shown in Equation (9.2).

$$\begin{aligned} k_j &= k_{j,r} \exp\Big[-\frac{E_j}{R}\Big(\frac{1}{T} - \frac{1}{T_r}\Big)\Big] \\ K_j &= \exp\Big(A_j - \frac{\Delta H_j}{RT}\Big) \end{aligned} \tag{9.2}$$

where $T_r = 773.15\,\mathrm{K}$, $E_4 = E_3$ and $\Delta H_4 = \Delta H_2 + \Delta H_3$ by Hess's law. The molar masses are built from toluene ($M_A = 92.14\,\mathrm{g/mol}$) and hydrogen, $M_N = M_A + 3M_H$, $M_P = M_N + M_H$ and $M_L = \frac{3}{7}(M_P + \frac{4}{3}M_H) = 44.10\,\mathrm{g/mol}$, so every reaction conserves mass and atoms exactly. At 500 °C $K_1 = 7.44 \times 10^5\,\mathrm{bar}^3$ and $K_2 = 0.400\,\mathrm{bar}^{-1}$. The constants are teaching values chosen so that the design point gives the bed drops of the plan. Against Smith's constants as reproduced in [8], k_1 is 8.5 times his at 500 °C with his activation energy of about 160 kJ/mol, his equilibrium constants are within a factor of 1.6 of these between 480 and 520 °C, and the two hydrocracking steps are first order in hydrogen where Smith wrote them zero order. At a hydrogen pressure near 3 bar they run at about a fifth of his rate, and in this model more hydrogen cracks more.

9.3.2 Stirred Cells and the Bed Heat Balance

Each bed is five cells of equal catalyst mass W along the radial gas path. The gas crosses a cell in well under a second, so a cell's composition is the stirred-tank solution at its temperature, while its temperature is a state on the heat capacity of the catalyst and a share of the vessel, as shown in Equation (9.3).

$$\begin{aligned} F_{out} &= F_{in} + \nu^{\mathsf{T}} \xi, \qquad \xi_j = W\, r_j(T, F_{out}) \\ C_{cell} \frac{\mathrm{d}T}{\mathrm{d}t} &= \dot{C}_p\,(T_{prev} - T) - \sum_j \xi_j \Delta H_j \end{aligned} \tag{9.3}$$

where ν is the stoichiometric matrix, ξ the extents in mol/s, $C_{cell} = W c_{cat} + C_{ves}/5$, T_{prev} the previous cell's temperature or, for a bed's first cell, the bed inlet $T_{in,b}$, and $\dot{C}_p = \sum_i F_i c_{p,i}$ the heat-capacity flow of the charge, held along the train. Holding $\dot{C}_p$ leaves out the small change of heat capacity by reaction and makes every bed's drop times $\dot{C}_p$ its reaction heat exactly, 11.30, 7.25, 1.90 and 1.16 MW at design with $\dot{C}_p = 160.5\,\mathrm{kW/K}$. The cell composition is found by Newton's method on the extents, four iterations from zero extent with each step damped so no flow falls below a tenth of its value, which makes the right-hand side a pure function of the state. The thermal time constants $C_{bed}/\dot{C}_p$ are 1.55, 1.91, 2.63 and 4.43 min.

9.3.3 Heaters, Tube Metal and the Exchanger

Each heater is first order on its absorbed duty and adds that duty to the gas, and the charge heater's inlet is set by the exchanger, as shown in Equation (9.4). The tube metal is Equation (1.3) without

the coke term, and the exchanger returns heat from the last bed to the first heater as in Section 1.4.5.

$$\begin{aligned} \tau_H \frac{\mathrm{d}Q_b}{\mathrm{d}t} &= \min(Q_{cmd,b}, Q_{cap,b}) - Q_b \\ T_{in,b} &= T_{h,b} + \frac{Q_b}{\dot{C}_p}, \qquad T_{m,b} = T_{in,b} + \theta_m \frac{Q_b}{Q_{max,b}} \\ T_{h,1} &= T_{cold} + \varepsilon\,(T_{out,4} - T_{cold}) \end{aligned} \tag{9.4}$$

where $T_{h,b}$ is the heater inlet, equal to $T_{out,b-1}$ for $b \geq 2$, $Q_{cap,b}$ the rated duty $Q_{max,b}$ times the share of burners in service (zero after a trip), and $\theta_m = 110\,\mathrm{K}$ the rise of the hottest tube metal over the outlet at rated duty. At design the exchanger returns the feed at 445.7 °C and the tube metals sit at 580, 583, 580 and 542 °C. A heater at its rated duty runs 110 K above its outlet, so with an outlet above 530 °C it passes the 640 °C trip. Fuel is the sum of the duties over an efficiency of 0.88, 33.1 MW at design.

9.3.4 Recycle Hydrogen and the Feed

The compressor speed sets the recycle flow, and its purity relaxes toward that of the net gas the reactions make, as shown in Equation (9.5).

$$\begin{aligned} F_{rec} &= \frac{N}{100} F_{rec,max}, \qquad \mathrm{H_2/HC} = \frac{y_H F_{rec}}{F_P + F_N + F_A} \\ \tau_s \frac{\mathrm{d}y_H}{\mathrm{d}t} &= \frac{h}{h+g} - y_H \\ \tau_N \frac{\mathrm{d}N}{\mathrm{d}t} &= \min(N_{cmd}, N_{cap}) - N \end{aligned} \tag{9.5}$$

where h is the net hydrogen made, g the C1 to C3 and half the C4 of the light ends ($0.7F_L$), and the feed flows of P, N and A come from the feed rate and the assay through molar volumes of 150, 128 and 107 $\mathrm{cm^3/mol}$. At design the recycle is 1440 mol/s at 84.7 % speed and 88.9 mol% hydrogen. The naphthenes and aromatics of a new tank reach the reactors through a first-order lag of 6 min.

9.3.5 Coke, Activity and the Regenerator

One coke state per bed treats each bed's catalyst as well mixed. Coke is laid down in proportion to the aromatics and to the inverse cube of the hydrogen pressure, and moves down the train with the catalyst, as shown in Equation (9.6).

$$\begin{aligned} \frac{\mathrm{d}C_b}{\mathrm{d}t} &= r_{c,b} - \frac{W_c}{M_b}(C_b - C_{b-1}), \qquad a_b = e^{-\gamma C_b} \\ r_c &= k_c\, a\, e^{-\frac{E_c}{R}\left(\frac{1}{T} - \frac{1}{T_r}\right)} \frac{p_A}{p_{A,r}} \left(\frac{p_{H,r}}{p_H}\right)^3 \\ B &= \min\left[W_c(C_4 - C_r),\ B_{max}\right] \\ \tau_{reg} \frac{\mathrm{d}C_0}{\mathrm{d}t} &= \max\left(C_r,\ C_4 - B_{max}/W_c\right) - C_0 \end{aligned} \tag{9.6}$$

where $r_{c,b}$ is the mean of the cell rates in bed b, M_b the bed's catalyst, $p_{A,r} = 0.4\,\mathrm{bar}$, $p_{H,r} = 3\,\mathrm{bar}$, $C_r = 0.1\,\mathrm{wt\%}$ the coke a full burn leaves, and the units are reconciled in the code (weight percent, tonnes per hour and kilograms per hour). Below its capacity $B_{max} = 400\,\mathrm{kg/h}$ the regenerator burns the catalyst back to C_r, and above it the catalyst returns to bed 1 carrying coke. At design the beds carry 0.21, 0.49, 1.35 and 2.99 wt% and the regenerator burns 239 kg/h, 60 % of its capacity. The pass through the reactors takes $M_{cat}/W_c = 8.0\,\mathrm{h}$, a deliberate compression (the application's texts put a commercial unit's cycle at 3 to 6 days, about twelve times longer) so that coke moves within a case, and the laydown, the burn and the capacity are compressed by the same factor. The coke in weight percent is therefore of a realistic size while the kilograms per hour are not. The capacity equals the coke on spent catalyst at 4.95 wt% burned back to C_r at the design circulation.

9.3.6 Reformate, Octane and Yield

The reformate is the C5+ liquid of the bed 4 effluent, the P, N and A lumps plus the pentane fifth of the light ends, and its octane is an illustrative correlation on its aromatics and paraffins, as shown in Equation (9.7).

$$\begin{aligned} \mathrm{RON} &= R_0 + b_A A_{vol} - b_P P_{vol} + b_T(\mathrm{WAIT} - T_R) \\ \mathrm{WAIT} &= \textstyle\sum_b f_b T_{in,b} \\ Y_{C5+} &= 100\,\dfrac{\dot{m}_P + \dot{m}_N + \dot{m}_A + \dot{m}_{C5}}{\dot{m}_f} \end{aligned} \tag{9.7}$$

where A_{vol} and P_{vol} are volume percent of the liquid, f_b the catalyst shares 0.10, 0.15, 0.25 and 0.50, and $R_0 = 86.205$ anchors the design point at 95.0. The linear dependence on aromatics follows pilot measurements [17], but the coefficients are teaching values and the correlation holds only for this reformate between WAIT 480 and 535 °C, where it gives 91.8 to 101.2 on the design naphtha and 87.2 to 96.6 on the paraffinic one. The octane and yield of the effluent reach the stabilizer bottoms through two first-order lags of 10 min, the separator and the stabilizer, so the reformate there is on average the reactor effluent of 20 min before.

9.3.7 What the Instruments Report

Every controller reads one measurement packet, built every 10 s. The bed inlet and outlet thermocouples lag 12 s with 0.3 K of noise, the tube-metal thermocouples carry 1 K of noise, the duties, feed and recycle meters 0.5, 0.3 and 0.5 %, and the recycle hydrogen analyzer lags 1 min with 0.002 of noise in mole fraction. The distributed control system computes H_2/HC from the purity, the recycle meter and the feed meter over the assay's molar mass. The NIR analyzer on the stabilizer bottoms draws a sample every 2 min through a 1 min sample line, reports one minute later with 0.1 RON of noise and its own calibration bias b_{NIR}, and is flagged stale after four missed cycles. A knock engine tests a sample every 8 h, reports an hour later with a repeatability of 0.1 RON, and is the reference. Chapter 15 holds the octane of the finished gasoline with the same pair of instruments. The regenerator's burn is read from its air with 2 % of noise and reads the capacity itself at the oxygen limit, the catalyst flow meter carries 1 %, and the assay of the tank in service gives N, A and N+2A. Every execution draws the same 23 normal deviates whatever the settings, so runs that differ in a single setting see identical noise. In the exported file every instrument column is matched by a column of the true value it stands for, whose name begins with `truth`.

9.3.8 Parameters, Starting State and Integration

Table 9.2 lists every parameter with its default. Each case starts lined out at its own operating point. A steady-state solve finds the cell temperatures for the bed inlets the heater loops are asked for, iterating the recycle purity and the coke to a fixed point (the coke bed by bed by bisection, with the burn-limited regenerated coke solved when the capacity binds) and capping any heater short of duty. The state advances in fixed fourth-order Runge–Kutta steps of 5 s. Halving the step moves case 5's octane, yield, WAIT and bed 1 drop at 120 min by 2.6×10^{-8} of their values and leaves cases 1, 6 and 8 unchanged to four decimals. Ten and twenty cells per bed in place of five enlarge bed 1's drop by 2.5 and 3.9 K and raise the octane by 0.15 and 0.23, with every conclusion unchanged. Over the challenge day with the suggested settings, mass, carbon and hydrogen balance at the reactor boundary to 9.2×10^{-16} of the throughput, and the integrated heat balance of the beds closes to 0.04 MJ against 2.5 TJ of heater duty. Over case 6 the 10 509 kg of coke laid down equals the 9425 kg burned plus the 1085 kg stored on the beds and in the regenerator. A simulated minute costs about 2 ms in a JavaScript engine, so the 24 h challenge runs in about 3 s.

9.3.9 Assumptions

Four C7 lumps stand for a naphtha of many carbon numbers, so the ring expansion of cyclopentanes, benzene precursors, heavier aromatics and isomerization are folded into four reactions, and the model

Table 9.2 Parameters of the Simulated Reformer by Access (A Adjustable, H Hidden, F Written Into the Source)

Symbol	Meaning	Value	Unit		Symbol	Meaning	Value	Unit	
P	reactor pressure	4.5	bar	A	θ_m	tube metal over outlet at rated duty	110	K	H
M_{cat}	catalyst, all four beds	66 000	kg	H	$T_{m,max}$	tube-metal limit (teaching value)	620	°C	A
f_b	catalyst shares	0.10, 0.15, 0.25, 0.50	—	F	T_{trip}	heater trip	640	°C	A
c_{cat}	catalyst heat capacity	1050	J/(kg K)	H	η_H	heater efficiency	0.88	—	H
C_{ves}	vessel heat capacity per reactor	8	MJ/K	H	ε	exchanger effectiveness	0.88	—	H
$k_{1,r}$	dehydrogenation at 500 °C	0.40	mol/(s kg bar)	A	T_{cold}	feed and recycle into the exchanger	100	°C	H
E_1	its activation energy	160	kJ/mol	H	k_c	coke laydown at 500 °C, fresh	0.76	wt %/h	A
$k_{2,r}$	ring opening and closing	6.4×10^{-4}	mol/(s kg bar^2)	A	E_c	its activation energy	120	kJ/mol	H
E_2	its activation energy	250	kJ/mol	H	β	laydown order in hydrogen	3	—	H
$k_{3,r}$	paraffin hydrocracking	2.64×10^{-3}	mol/(s kg bar)	A	m	activity exponent of ring closure	2	—	H
$k_{4,r}$	naphthene hydrocracking	6.6×10^{-4}	mol/(s kg bar)	H	γ	activity lost per wt % coke	0.12	1/wt %	A
$E_3 = E_4$	hydrocracking activation	250	kJ/mol	H	C_r	coke left by a full burn	0.1	wt %	H
ΔH_1	dehydrogenation (endothermic)	+210	kJ/mol	H	B_{max}	regenerator burn capacity	400	kg/h	A
ΔH_2, ΔH_3	ring closing, cracking	−40, −60	kJ/mol	H	τ_{reg}	regenerator catalyst residence	30	min	H
A_1, A_2	ln K intercepts	46.19, −7.14	—	H	τ_{lift}	catalyst lift response	5	min	H
$c_{p,i}$	P, N, A, H, L, G vapor	290, 260, 210, 29.7, 70, 65	J/(mol K)	F	$F_{rec,max}$	recycle at full speed	1700	mol/s	H
M_G	recycle light gas	20	g/mol	F	τ_N, τ_s	compressor speed, purity lag	0.3, 3	min	H
$Q_{max,b}$	rated duties	12, 15, 10, 5	MW	A	τ_{feed}	feed line to the reactors	6	min	A
τ_H	heater firing response	2	min	A	τ_{prod}	separator, stabilizer (each)	10	min	A
R_0	octane constant	86.205	—	H	b_A, b_P	octane per vol % A, P	0.38, 0.15	1/vol %	H
b_T, T_R	octane per K of WAIT, reference	0.07, 500	1/K, °C	H	$p_{A,r}$, $p_{H,r}$	coke reference pressures	0.4, 3	bar	F

gives the order of the beds' drops and the slope of the severity trade rather than a particular catalyst's yields. Each bed is five stirred cells with no gas holdup and no pressure drop, which underestimates bed 1's drop by a few kelvin against plug flow. The heaters are first order on their duty with a tube metal proportional to it, leaving out the firebox, the draft and the fuel header. One coke state per bed ignores the coke profile along a bed and the difference between coke on the metal and on the support, and the regenerator is one well-mixed vessel with a burn limit rather than burn, chlorination and reduction zones. The catalyst loop is compressed about twelve times, which keeps the coke in weight percent realistic but makes the kilograms per hour and the capacity twelve times a real unit's. The separator pressure and the feed flow are taken as held by fast loops the model does not include, and the separator and stabilizer are lags on the octane and yield. The octane correlation is illustrative. Because the plan's bed drops at $H_2/HC = 3$ limit the endotherm the model can spend, its 95-octane reformate carries 44.2 vol% of aromatics and its C5+ yield sits about 7 volume points above a published correlation (Section 9.4).

9.4 What the Plant Does Before Any Controller

9.4.1 The First Bed Takes the Heat

Bed 1 meets the naphthenes first (Figure 9.2). Across it the naphthenes fall from 22.0 to 9.4 mole percent of the hydrocarbon and the aromatics rise to 38.4, and the dehydrogenation cools the bed until its own rate slows, 70.4 K at design. Bed 2 finishes the naphthenes, down to 1.4 mole percent,

with a 45.1 K drop, and beds 3 and 4 are left with the slow conversion of paraffins, which takes heat, and hydrocracking, which gives some back, for 11.8 and 7.2 K. The consequence for control is the one the introduction stated. A hotter bed 1 reacts faster and cools itself more, so it passes on only a part of any heat given to it. In case 1, in Manual, heater 1 receives 1.6 MW more at 10 min. By 120 min its outlet is 15.7 K higher, 6.4 K of it from the exchanger, whose outlet has risen as the warmer bed 4 effluent comes around. Of that 15.7 K, bed 1 absorbs 9.3 K in a drop that grows to 79.7 K and passes 6.3 K to bed 2, the WAIT rises 10.5 K, the octane of the reformate goes from 95.0 to 97.3, and heater 1's tube metal from 580 to 610 °C. In Auto the heater loops break the loop through the exchanger.

Pause. Try Exercise 9.1 before reading on.

9.4.2 Severity Buys Octane With Yield

The WAIT moves octane and yield in opposite directions, as shown in Figure 9.3. Lined out with the coke at its new steady state, 10 K of WAIT at design buys 2.01 RON and costs 1.51 points of C5+ yield by mass. With the coke frozen at design the same step gives 2.30 RON and 1.67 points, the difference being the extra coke a hotter catalyst carries. In case 2, stepped from 500 to 510 °C at 10 min, the octane at the reactors is 97.3 by 20 min, the inferred octane passes 96 at 11 min and the analyzer's reading at 27.5 min, and by 4 h the reformate is at 97.1 with the yield down from 94.7 to 93.1 wt%, the aromatics up 2.4 vol%, the hydrogen up from 2.20 to 2.31 t/h and the fuel from 33.1 to 34.7 MW. The coke on the spent catalyst follows over hours, from 3.0 to 3.7 wt% by 10 h. One octane number of giveaway is about three quarters of a point of yield here.

> **TRY IT IN THE APP** **2 · Severity against yield: WAIT +10 °C**
>
> Load the case and press Run. At 10 min raise the WAIT setpoint from 500 to 510 °C. Watch the octane and the C5+ yield. The octane at the reactors is 97.3 by 20 min, and the analyzer's reading passes 96 only at 27.5 min. By 4 h the reformate is at 97.1 and the yield has fallen from 94.7 to 93.1 wt%.

The regenerator puts a ceiling on that curve. Lined out at $H_2/HC = 3$ and the design circulation, the coke made per hour reaches the 400 kg/h capacity at a WAIT of 523.5 °C on the design naphtha, where the octane is 100.1, and at 524.5 °C on the paraffinic naphtha, at 96.1. Above it the regenerated catalyst returns to bed 1 with coke, 3.36 wt% of it at 535 °C on the design naphtha, and five more kelvin buy about 0.5 RON where below the limit they bought 1.1 to 1.2. The laydown and the capacity carry the same compression of the catalyst loop, so the WAIT at which this ceiling appears does not depend on it. The paraffinic naphtha, with an N+2A of 48 against 62, needs 520.2 °C for 95 RON instead of 500, and at WAIT 535 °C it reaches only 96.6.

Pause. Try Exercise 9.3 before reading on.

9.4.3 How the Model's Yield Compares

Gary, Handwerk and Kaiser chart the C5+ volume yield of a reformer against the reformate's octane for naphthas of different N+2A [18]. Read by interpolation for N+2A 62, the chart gives about 84.5 volume percent at 95 RON where the model gives 91.7, and over 93 to 99 RON the model sits 7.0 to 7.6 points above the chart, the range spanning two independent readings of it. The slope agrees within the precision of the reading, 0.93 volume percent per octane number from 93 to 97 RON in the model against 0.83 to 0.87 on the chart, and the leaner naphtha costs the model 4.7 points at 95 RON against the chart's 3.8 to 4.0. The level and the hydrogen agree on the cause. The model makes 1.47 wt% of net hydrogen on feed at 95 RON, where the same source's hydrogen chart gives about 2.4 at its yield, and since each aromatic ring formed releases three molecules of hydrogen or more, the model aromatizes less of its naphtha than the units behind the chart. Its reformate at 95 RON carries 44.2 vol% of aromatics. The plan fixed a drop of about 70 K for the first bed and a few kelvin

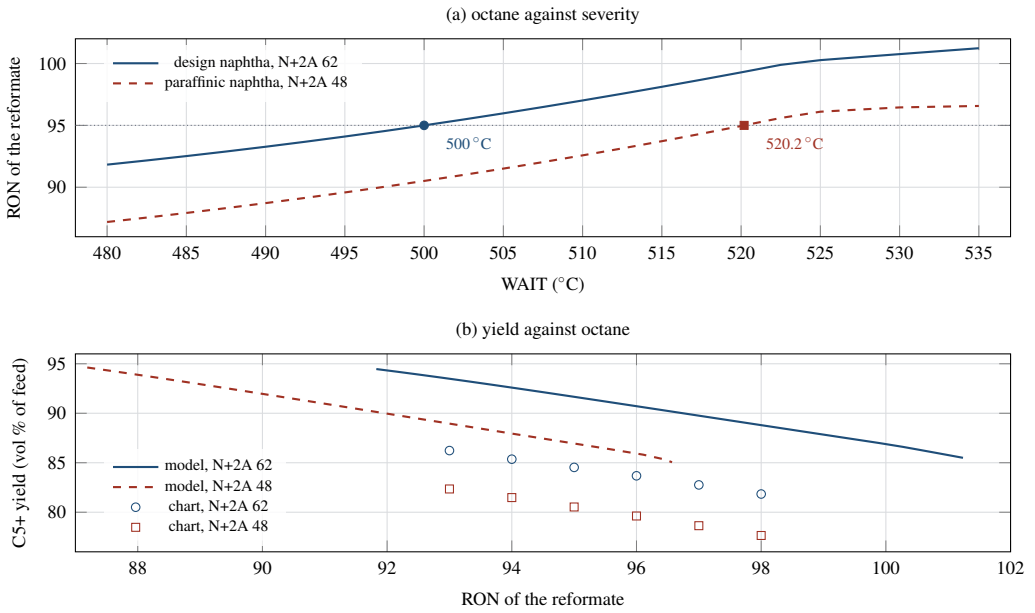


Figure 9.3 The severity trade from the application's own steady state at $H_2/HC = 3$ with the coke lined out at each WAIT from 480 to 535 °C. (a) The octane rises with WAIT on both naphthas (design solid, paraffinic dashed), and the curves bend where the regenerator reaches its burn capacity, from 523.5 °C on the design naphtha and 524.5 °C on the paraffinic one. Dots mark 95 RON. (b) C5+ volume yield against octane, with the yield read from Gary, Handwerk and Kaiser's chart for naphthas of N+2A 62 and 48 (open circles and open squares, interpolated between its curves). The model's slope is close to the chart's, its level about 7 points higher.

for the last at $H_2/HC = 3$, and those drops bound the endotherm, and with it the aromatics, that the model can spend.

9.4.4 Coke Answers in Hours

Coke is the slow variable. Lowering H_2/HC from 3 to 2 at 500 °C raises the lined-out burn demand from 239 to 334 kg/h, still under the capacity, and a cut of the ratio from 3 to 2 at 95 octane leaves the burn demand at 332 kg/h after a day and moves the WAIT the trim needs by only 0.3 K. Only where the regenerator binds does coke reach the octane, and the turndown of case 6 is therefore run at 98 octane. With the lift stopped the coke has no exit at all, and the WAIT for constant octane climbs as it does in a semiregenerative unit (case 8).

9.5 Controllers, Estimator and Optimizer as Coded

The controllers execute every 10 s, two physics steps, and the MPC every 2 min. Each execution builds the truth record, passes it through the instruments, updates the octane estimates and runs the octane layer, the WAIT layer, the heater loops and the H_2/HC loop, in that order, before writing the commands the plant's actuators follow. The header's Manual and Auto switch the four heater duty stations and the compressor station together, and in Manual every loop tracks its station so that the return to Auto is bumpless.

9.5.1 The Shared Velocity-Form PID

All four loop types use the velocity-form PID of Section 1.6.2, with the derivative taken from the measured variable, as shown in Equation (9.8).

$$\begin{aligned} e_k &= \mathrm{SP}_k - \mathrm{PV}_k, \qquad P_k = K_c e_k, \qquad D_k = -K_c \tau_D \dot{\mathrm{PV}}_k \\ s_k &= s_{k-1} + (P_k - P_{k-1}) + (D_k - D_{k-1}) \\ \hat{u}_k &= s_k + I_{k-1}, \qquad u_k = \min(u_{max}, \max(u_{min}, \hat{u}_k)) \\ I_k &= I_{k-1} + \left[\frac{K_c}{\tau_I} e_k + \frac{u_k - \hat{u}_k}{\tau_I} \right] \Delta t \end{aligned} \tag{9.8}$$

where the running sum s_k is never clamped and the back-calculation acts on the integral only, with a tracking time equal to τ_I, as in Equations (1.17) to (1.19). Any of these loops given an integral time of at least 600 min runs proportional-only. No loop here uses derivative action. The error is setpoint minus measurement in every loop, so a low reading raises the output.

9.5.2 Heater Loops and the WAIT Layer

Each heater outlet loop moves its duty between zero and the capacity the control system knows, the rated duty times the burners in service, at $K_c = 0.25\,\mathrm{MW/K}$ and $\tau_I = 2.5\,\mathrm{min}$ for all four. The process gain of an outlet to its duty is $1/\dot{C}_p = 6.2\,\mathrm{K/MW}$ through the 2 min firing lag. The setpoints come from one WAIT controller over a declared distribution, as shown in Equation (9.9).

$$T_{SP,b} = \mathrm{WAIT}_{SP} + s\,\delta_b + \beta, \quad \beta = \mathrm{PI}_W(\mathrm{WAIT}_{SP} - \widehat{\mathrm{WAIT}}) \tag{9.9}$$

where $\widehat{\mathrm{WAIT}}$ is the catalyst-weighted mean of the measured inlets, s the severity shift (0 for equal inlets and 1 for the aged distribution), and $\delta = (-6, -4, 0, +2.4)$ K moves severity from beds 1 and 2 to bed 4 with no change in the catalyst-weighted mean. The bias β is limited to $\pm 15\,\mathrm{K}$, its gain is 0.3 K/K with $\tau_I = 6\,\mathrm{min}$, and its integral is held while every heater sits at the limit the error pushes toward, so it cannot wind up behind them. The heater loops do not watch the tube metal. A saturated heater asked for more simply fires at its cap, and with $\theta_m = 110\,\mathrm{K}$ a WAIT setpoint stepped 30 K at once in Auto trips heater 3 3.9 min later. A tube-metal override on each heater loop would prevent that, but it would also hold the charge heater of case 7 under its limit and remove that case's lesson, so the simulator leaves the tube metals to the operator, the inlet distribution and the MPC.

9.5.3 Octane Trim, Feedforward and the Inferred Octane

The octane trim sets the WAIT setpoint from the octane the controllers read together with a feedforward on the assay of the tank in service, as shown in Equation (9.10).

$$\mathrm{WAIT}_{SP} = \mathrm{PI}_R(\mathrm{RON}_{SP} - \widehat{\mathrm{RON}}) + K_{ff}\,(62 - \overline{\mathrm{N{+}2A}}) \tag{9.10}$$

where $\overline{\mathrm{N{+}2A}}$ is the assay passed through the controllers' 6 min model of the feed line and $K_{ff} = 1.4\,\mathrm{K}$ per unit, which supplies 19.6 K for the paraffinic naphtha against the 20.2 K its lined-out octane needs. The trim's gain is 2 K/RON with $\tau_I = 20\,\mathrm{min}$, and its output limits are the WAIT range, 480 to 540 °C, less the feedforward, so its integral never winds up behind the clamp on their sum. On the analyzer the trim executes once per report, over the time since the last, holds when the analyzer is stale, and holds its integral when every heater is at the limit the error pushes toward.

The octane the controllers read is either the NIR corrected by the laboratory or an inferred octane, as shown in Equation (9.11). The NIR's correction is the laboratory bias update of Equation (1.14).

$$\begin{aligned} \mathrm{RON}_{NIR+lab} &= y_{NIR} + \hat{b}, \quad \hat{b} \leftarrow \hat{b} + \lambda\left(y_{lab} - (\bar{y}_{NIR} + \hat{b})\right) \\ \mathrm{RON}_{inf} &= \mathscr{M}(\widehat{\mathrm{WAIT}}, \overline{\mathrm{N{+}2A}}, \widehat{\mathrm{H_2/HC}}, s) + \hat{b}_s \end{aligned} \tag{9.11}$$

where $\bar{y}_{NIR}$ is the mean of the three NIR reports around the laboratory sample, $\lambda = 0.7$, and $\mathscr{M}$ is the controllers' octane map, a seventeen-term polynomial in $(\text{WAIT} - 500)/10$, $(\text{N+2A} - 62)/7$, $\text{H}_2/\text{HC} - 3$ and s (quadratic with the cube of the first and one cross term) fitted at 663 lined-out points of the commissioned plant with the catalyst as designed, with residuals of 0.020 RON root mean square and 0.103 at worst. The map gives the octane being made at the reactors with no delay. The offset $\hat{b}_s$ carries what the map cannot see, chiefly coke, and follows each corrected NIR report, compared with the map's value passed through the controllers' model of the two 10 min product lags, with a factor $1 - e^{-2/30}$ per report. A corrected report more than 0.4 RON from the inferred octane on two reports running is held as suspect, $\hat{b}_s$ stops following the NIR, each laboratory result then corrects $\hat{b}_s$ directly, the result that finds a suspect NIR re-references it with $\lambda = 1$, and the flag clears when the corrected NIR agrees within 0.2 RON. A calibration shift jumps where coke drifts, and the gate tells them apart. The flag is shown whichever octane the trim reads, but only the inferred source acts on it.

9.5.4 Model Predictive Control

The MPC is the linear MPC of Section 1.6.3 on three decisions, the WAIT setpoint W, the severity shift s and the H_2/HC setpoint H. Its model is the controllers' own lined-out maps of octane, C5+ yield, burn, the four duties and the four tube metals, linearized at every execution by central differences around the measured WAIT and H_2/HC and the present shift, made dynamic by first-order lags of 4, 4 and 1.5 min on the decisions and discretized exactly for a zero-order hold. Each predicted output carries the offset between its measurement and its map, filtered over 10 min for the duties and tube metals and taken from the octane estimator for the octane, while the burn has none and is held on the lined-out map. This output bias is the simplest way to remove offset, where Volume 9, Chapter 4 estimates an unmeasured input with a Kalman filter instead. The program over 30 steps of 2 min in blocks of 1, 1, 2, 4, 8 and 14 is Equation (9.12).

$$\begin{aligned} \min\ & \textstyle\sum_k \left[w_R (\text{RON}_k - \text{RON}_{SP})^2 - w_Y Y_k \right] + c_H \sum_j H_j \\ & + \textstyle\sum_j \Delta u_j^{\mathsf{T}} S \, \Delta u_j + \text{slack penalties} \\ \text{s.t. } & 480 \le W \le 540, \ \ 0 \le s \le 1, \ \ 2.7 \le H \le H_{cap} \\ & |\Delta W| \le 8, \ \ |\Delta s| \le 0.25, \ \ |\Delta H| \le 0.25 \\ & B_k \le B_{max} - 10, \ \ T_{m,b,k} \le T_{m,max} - 8 \\ & Q_{b,k} \le Q_{cap,b} - 0.05 \end{aligned} \tag{9.12}$$

where $w_R = 20$ per RON squared, $w_Y = 1$ per weight percent, $c_H = 0.4$ per unit of H_2/HC, $S = \text{diag}(0.5, 50, 50)$, the output limits are soft with quadratic and linear slack prices of 10^4 and 10^3, H_{cap} is the ratio the compressor can reach, and the lower bound of 2.7 keeps a margin of 0.2 over the floor of 2.5. A heater already at its duty limit has that limit dropped, its tube-metal limit kept, and the WAIT layer moves its share of the severity to the others. A solution that is not finite or not inside its bounds is rejected and the last decisions are held. A dual active-set quadratic programming module, shared by the predictive controllers of the series, solves the program. Over the challenge day the MPC solves 721 times with no fallback.

9.5.5 Trips and the Models Kept Apart

Heater protection sits apart from the loops. It reads the true tube metal, trips a heater that reaches 640 °C and stops the run there. The 620 °C limit, the 640 °C trip, the 400 kg/h capacity and the H_2/HC floor of 2.5 are teaching values. Real ones come from the tube alloy and an API 530 heater design, from the regenerator's air and oxygen system and from the licensor's operating guidelines. The application keeps three models apart. The plant is everything in Section 9.3. The identified model is what a learner fits from a step, such as the 6.2 K/MW heater gain. The controllers' model is the set of maps with the catalyst as designed, which knows neither the present coke nor the analyzer's bias.

Table 9.3 The Nine Cases

#	Case, preset and events	Teaching point
1	Manual, heater 1 duty raised by the learner	the staircase
2	WAIT setpoint raised by the learner, 500 to 510 °C	severity against yield
3	Heater loops 0.08/8 min, WAIT controller off, WAIT SP +10 at 20, heater 2 at 60 % at 120 min	heater loops, WAIT layer
4	Trim on NIR, NIR +0.8 RON at 30 min, laboratory at 150, reported 210 min	a biased analyzer
5	Trim on NIR, paraffinic naphtha at 30 min	feedforward from N+2A
6	Trim on NIR at 98 RON, compressor at 62 % at 30 min	H_2/HC, coke, regenerator
7	Paraffinic, trim on inferred with feedforward, heater 2 at 60 % at 30 min	tube metal, distribution
8	Trim on NIR, catalyst lift stops at 30 min	semiregenerative decline
9	Challenge: paraffinic at 4, heater 2 at 10, NIR +0.8 at 16 h	the whole problem

Table 9.4 Measured Outcomes by Case, Default Seed

#	Setting	Outcome
1	+1.6 MW at 10 min	bed 1 drop 79.7, outlet +6.3, heater 1 outlet 515.7 °C at 120 min
2	WAIT 510 °C	RON 97.1, C5+ 94.7 to 93.1 wt% at 4 h, coke 3.0 to 3.7 wt%
3	as shipped	WAIT dips to 501.6, holds 507.0, bed 2 20 K short
	heaters 0.25/2.5 min	dip to 505.3 °C, deficit unchanged
	+WAIT 0.3/6 min	bias 3.4 K, WAIT 510.0 °C by 180 min
4	trim on NIR	true RON 94.24, under 94.7 from 54 to 248 min
	trim on inferred	true RON never under 94.97
5	no feedforward	lowest 92.2, 112 min out of 95 ± 0.3
	feedforward	lowest 94.9, never out of the band
6	as shipped	H_2/HC 2.08, burn limit at 312 min, WAIT +2.2 K by 24 h
	feed 120 t/h at 8 h	H_2/HC 2.56, C_0 back to 0.10 wt%
7	equal inlets	heater 1 tube metal 619 to 626 °C
	aged at 45 min	under 620 °C at once, near 605, RON 95.1
8	as shipped	WAIT 499.6, 507.4, 513.2 °C at 0, 12, 24 h, bed 4 10.5 wt%

9.6 What the Nine Cases Show

The cases appear in Table 9.3. All start from the defaults of Section 9.5, WAIT only at 500 °C with equal inlets, the trim on the NIR and the design naphtha, and each row lists only what differs from them. The values below come from the engine of Section 1.6.1 advanced without a display at the default seed, with each learner move applied at the time the text gives. A case run first, after another case or from Reset produces the same trajectory. The outcomes are collected in Table 9.4.

9.6.1 Heater Loops Against Their Neighbors

Case 3 ships the heater loops detuned, 0.08 MW/K with an 8 min integral time, and the WAIT controller off. The WAIT setpoint step of 10 K at 20 min reaches the beds, but the inlets of beds 3 and 4 overshoot to about 512 °C as each warmer upstream outlet arrives at the next heater as a disturbance. At 120 min heater 2 loses burners and can give 60 % of its rated duty. The WAIT falls to 501.6 °C and stays 3 K short at 507.0 °C, with bed 2's inlet 20 K under its setpoint. Retuned loops at 0.25 MW/K and 2.5 min cut the dip to 505.3 °C and leave the deficit, because no heater loop can make heater 2 bigger. The WAIT controller at 0.3 K/K and 6 min adds a 3.4 K bias to every setpoint, beds 1, 3 and 4 follow it, and the WAIT is back at 510.0 °C by 180 min.

9.6.2 Octane Measured Late, Wrong and Rarely

In case 4 the NIR starts reading 0.8 RON high at 30 min. The trim on the analyzer does what it is told, lowering the WAIT setpoint to 495.8 °C and holding the corrected reading at 95 while the true reformate falls to 94.24, under 94.7 from 54 min. The knock-engine sample drawn at 150 min and reported at 210 min moves the analyzer's correction to −0.83, and the true octane is back over 94.7 at 248 min. On the inferred octane the validation holds the NIR as suspect from 33.1 min, two reports after the shift reaches the sample, to 211.0 min, and the true octane never falls under 94.97. At seeds

1, 2 and 3 the trim on the NIR takes the true octane down to 94.24, 94.22 and 94.19, with laboratory corrections of −1.00, −0.94 and −0.79, while on the inferred octane it never falls under 94.99.

Case 5 switches to the paraffinic naphtha at 30 min. Bed 1's drop starts to shrink as soon as the leaner feed arrives, from 70.3 to 63.9 K by 60 min even as the WAIT climbs, and the octane at the reactors passes under 94.7 at 30.5 min. The reformate follows at 38 min and the analyzer reports it at 41.5 min. The trim on the analyzer then chases a late signal, the true octane falls to 92.2 and stays outside 95 ± 0.3 for 112 min (139, 120 and 133 min at seeds 1, 2 and 3), and the WAIT settles near 519 °C. With the feedforward the WAIT rises with the feed and the lowest octane is 94.9 (94.82 to 94.92 at the four seeds). The leaner naphtha costs about 4 points of yield at the same octane, 94.7 to 90.6 wt%, and at that severity the charge heater's tube metal exceeds its limit for 11 min, reaching 621 °C.

> **TRY IT IN THE APP** **5 · A paraffinic naphtha: the octane falls before the analyzer says so**
>
> Load the case and press Run. Touch nothing. At 30 min the feed changes to a leaner naphtha. Watch the true octane and the analyzer's reading. The octane at the reactors is under 94.7 at 30.5 min, and the analyzer reports it at 41.5 min. The true octane falls to 92.2. Then turn N+2A feedforward On, press Reset and run again. The lowest octane is 94.9.

9.6.3 Hydrogen, Coke and the Regenerator

Case 6 holds 98 RON at a WAIT of 514.1 °C, since at 95 the regenerator never limits, and limits the recycle compressor to 62 % speed at 30 min. The ratio falls to 2.08, and no setpoint or speed can raise it. Lined out with the compressor at 20 to 62 %, H_2/HC is 0.71 to 2.08 and stays at 2.08 for any speed asked above 62. The coke on the spent catalyst rises from 4.1 to 5.2 wt% by 8 h, the burn demand passes the capacity at 312 min, the regenerated catalyst returns with 1.15 wt% of coke by 24 h, and the trim raises the WAIT by 2.2 K to keep 98 RON while the coke still climbs. Lined out, the WAIT needed for 98 RON against this compressor is 6.0 K higher. The feed rate is the honest remedy. Cut to 120 t/h at 8 h, it raises H_2/HC to 2.56, the burn falls under its limit at 677.5 min, and the regenerated catalyst is clean again by 24 h.

Pause. Try Exercise 9.2 before reading on.

Case 8 stops the catalyst lift at 30 min. The trim on the analyzer raises the WAIT steadily, about 0.6 K an hour, from 499.6 °C to 507.4 by 12 h and 513.2 by 24 h, holding the octane within 0.15 of 95. Bed 4, with half the catalyst and the most aromatics, cokes fastest, to 10.5 wt% by 24 h, its activity falls to 0.28, and its drop shrinks from 6.8 to 0.9 K. The C5+ yield hardly moves, 94.7 to 95.0 wt%, because the deactivated catalyst cracks less while its hotter beds crack more, whereas a platinum catalyst lost yield steadily at constant octane in a published life test [14]. Restarted at 12 h, the lift raises bed 4's activity from 0.42 to 0.62 by 24 h while the regenerator, at its limit, burns the stored coke, and the trim brings the WAIT setpoint back from 507.3 to 501.6 °C.

9.6.4 A Heater Short of Duty Near Its Tube-Metal Limit

Case 7 runs the paraffinic naphtha at 95 RON, WAIT 520 °C, with the trim on the inferred octane and the feedforward on, and the charge heater's tube metal already at 619 °C. At 30 min heater 2 drops to 60 % of its duty, bed 2's inlet falls 15.7 K, the WAIT controller adds 2.7 K to every setpoint, and heater 1's tube metal passes 620 °C within 2 min, reaching 626 °C. Selected at 45 min, the aged distribution takes bed 1 6 K and bed 2 4 K cooler and bed 4 2.4 K hotter. Bed 1 holds a tenth of the catalyst, so its cooler inlet costs little octane and spares the most duty, and heater 1's tube metal falls under the limit at once and settles near 605 °C, heater 3's rises to about 605 and heater 4's peaks at 596 °C, with the octane at 95.1.

Table 9.5 The Challenge Over 24 h at the Default Seed. Share of Judged Time Within 95 ± 0.3, Tube-Metal Peak, Burn-Demand Peak, Lowest H_2/HC and Average C5+

Settings	Band (%)	T_m (°C)	B (kg/h)	H_2/HC	C5+ (wt %)	Verdict
shipped (WAIT only, equal)	13.6	587	239	2.95	94.17	fails
trim on NIR	77.9	625	361	2.95	91.36	fails
trim on NIR, ff, aged	84.7	607	360	2.90	91.22	fails
trim on inferred, ff, equal	100	624	366	2.90	91.19	fails
trim on inferred, aged, no ff	87.7	606	364	2.93	91.27	fails
trim on inferred, ff, aged[a]	100	606	364	2.90	91.16	passes
MPC on NIR	86.4	617	383	2.61	91.23	fails
MPC on inferred	100	617	387	2.61	91.16	passes

[a]The suggested settings. Here ff is the N+2A feedforward.

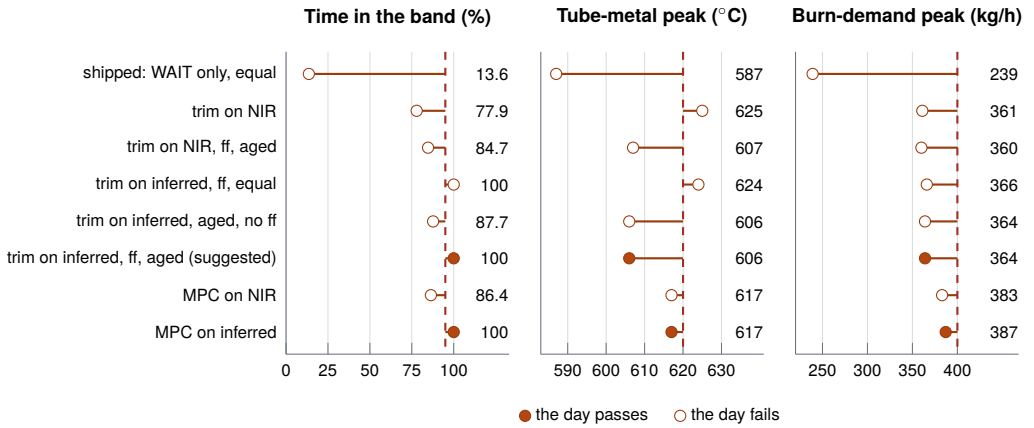


Figure 9.4 Three columns of Table 9.5 against their limits (dashed): the share of judged time within 95 ± 0.3 against 95 %, the tube-metal peak against 620 °C and the burn-demand peak against 400 kg/h. A filled circle marks a day that passes and an open circle a day that fails. Notice that every row on the analyzer stops short of the band target whatever else it has, that equal inlets put two rows over the tube-metal limit, and that the MPC runs closer to both limits at the right than the suggested trim does. The lowest H_2/HC and the C5+ yield, which no row fails, are not drawn.

9.6.5 The Challenge and the Analyzer It Cannot Trust

The challenge runs 24 h with the paraffinic naphtha from 4 h, heater 2 at 60 % from 10 h and the NIR reading 0.8 RON high from 16 h, with knock-engine samples at 2, 10 and 18 h. Its target is the true reformate octane within 95 ± 0.3 for 95 % of the time after the first hour, every tube metal at or under 620 °C, the regenerator's burn demand at or under 400 kg/h, the true H_2/HC at or over 2.5, an average C5+ yield within 0.5 points of the best the unit can make at 95 RON over the day (91.18 wt%, so 90.68 wt% or more), and no heater trip. The variants are listed in Table 9.5. Figure 9.4 draws three columns of the table against their limits.

As shipped the octane falls to 90.4 after the feed switch and never returns, 13.6 % of the time in the band. Each event needs its own layer. The feedforward handles the switch (without it the switch alone costs 169 min out of the band), the aged distribution keeps heater 1 under its limit (with equal inlets it reaches 624 °C after the heater 2 limit), and the inferred octane rides through the analyzer's shift. The suggested trim on the inferred octane with the feedforward and the aged distribution stays in the band all day, at 606 °C, 364 kg/h of burn demand, H_2/HC of at least 2.90 and 91.16 wt% of C5+, and passes at seeds 1, 2 and 3 with 91.17 to 91.20 wt%. The MPC on the inferred octane passes as well, using the room to 617 °C of tube metal and 387 kg/h of burn that the trim leaves unused, with

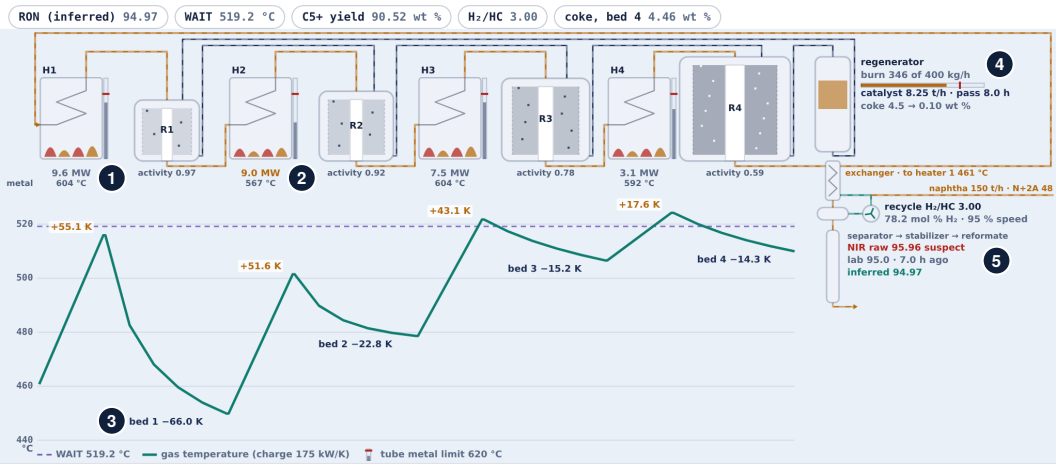


Figure 9.5 The challenge at $t = 18$ h, Unit view, cropped to the plant view with numbered callouts, with the suggested settings (octane trim on the inferred octane, N+2A feedforward, aged distribution). On the paraffinic naphtha with heater 2 held at its 9.0 MW cap (2), the staircase rises 55.1, 51.6, 43.1 and 17.6 K through the heaters and drops 66.0 (3), 22.8, 15.2 and 14.3 K through the beds for a charge of 175 kW/K, the exchanger returns the feed at 461 °C, and the tube metals read 604 (1), 567, 604 and 592 °C. The beds' activities are 0.97, 0.92, 0.78 and 0.59, and the regenerator (4) burns 346 kg/h of its 400 by its air reading, the catalyst passing at 8.25 t/h, a pass every 8.0 h, with coke falling from 4.5 to 0.10 wt%. The NIR's raw report (5), 95.96, is flagged suspect beside a laboratory value of 95.0 reported 7.0 h earlier and an inferred octane of 94.97. The WAIT stands at 519.2 °C.

the same yield.

> ✦ **DID YOU KNOW? A Better Controller Cannot Mend a Wrong Reading**
>
> From the simulator (Table 9.5). On the graded day the MPC that reads the analyzer keeps the octane in its band 86.4 % of the time and fails. A plain octane trim on the inferred octane, with the feedforward and the aged distribution, holds 100 % and passes. The MPC passes too, once it reads the inferred octane.

No strategy that believes the analyzer can pass. The judged time is 1380 min, so the rule allows 69 min outside the band, but the shift at 16 h lasts until the sample drawn at 18 h is reported at 19 h, and the trim on the corrected NIR loses 190.5 min after 16 h with or without the other layers, the MPC on it 188.5 min. The best of them, the MPC on the NIR, keeps the octane in the band 86.4 % of the time and the best trim 84.7 %. This is why the simulator offers the inferred octane as a measurement choice, and why the target is not met with the plan's 8 h laboratory alone. The challenge scores the regenerator's burn demand, which the plant caps at the capacity, so a demand over 400 kg/h is a violation even though the burn never exceeds it. The state at 18 h under the suggested settings is shown in Figure 9.5.

9.7 Technology in Practice

Figure 9.6 draws the loops of Section 9.5 with the hardware a refinery installs.

What Measures It

The bed inlets and outlets and the heater tubes are read by thermocouples, which are less accurate than resistance thermometers but more rugged and wider in span [19]. Their 12 s lag in Section 9.3 is

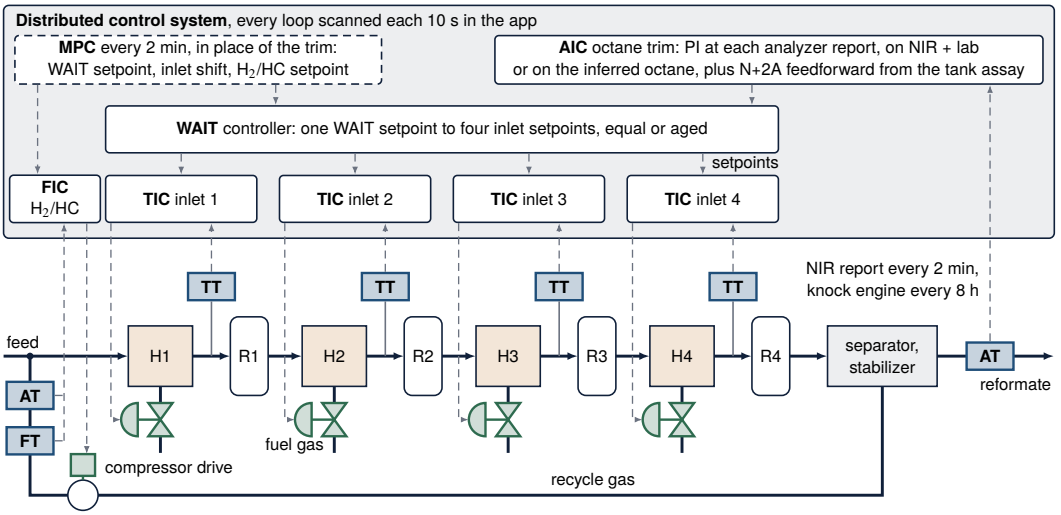


Figure 9.6 The loops of the reformer with their hardware. Lettered boxes are sensing devices, thermocouples with their transmitters (TT), the recycle flow meter (FT) and the analyzers (AT). Valve symbols and the compressor drive are final elements, and dashed arrows are signals. Notice that four fuel gas valves answer to one octane number through three controllers in cascade, and that the only measurement of that octane stands at the far end of the product path. Not to scale, and not a piping or safety diagram.

short beside the 2 min firing lag of a heater. The measurement that sets the pace is the octane, which the analyzer reads on a product that left the reactors about 20 min earlier.

> **TECHNOLOGY** **Sensor: Near-Infrared Octane Analyzer**
>
> Octane is defined by an engine test, and the research method is ASTM D2699 [20]. An online analyzer runs no engine. It passes near-infrared light through the sample, where the overtone and combination bands of the carbon and hydrogen bonds carry information on the types of hydrocarbon, and a multivariate model turns the spectrum into a property. A measurement takes a minute or less. The method is a secondary one: its model is built from samples rated by the reference method, may need hundreds or even thousands of them, and has to be updated frequently as the sample matrix changes [21]. A reading can therefore repeat well and still be wrong, by the error of its model. The bias b_{NIR} of Section 9.3 stands for that error, and the knock engine every 8 h for the reference method.

What Moves It

Four fuel gas valves carry the severity, the drive of the recycle compressor carries the hydrogen ratio, and the catalyst lift carries the circulation. A review of commercial practice describes the pattern the model has [13]. Heaters stand before and between the reactors, the temperature falls fastest in the first reactor, and the first reactor is the smallest while the third and fourth are very large. Its table of typical conditions for three reactors gives temperature drops of 69, 31 and 6 K, beside the 70.4, 45.1, 11.8 and 7.2 K of the model's four beds.

TECHNOLOGY Final Element: Fuel Gas Valve of a Fired Heater

A heater's duty is moved through a control valve on the fuel gas to its burners. At one refinery reactor furnace the inlet temperature controller set the furnace temperature controller, which in turn set the fuel gas pressure [22]. A pneumatic valve that air opens fails closed, and the failure mode is to be chosen from the risks of the process and not from habit. Beside the control valve stands a burner management system. It watches for low fuel pressure, high fuel pressure and loss of flame, and stops the fuel with shutoff valves of its own. Where a master controller can call for a setpoint that is unsafe, a limit function belongs between the two controllers [19]. The model's heater is a 2 min lag on its duty, capped by the burners in service, and its loops carry no tube-metal limit.

Where the Controller Runs

A refinery runs these loops in a distributed control system. Its loops are sometimes nested, the set point of one being set by another, and their cycle times range from milliseconds to minutes [23]. The app scans every loop each 10 s and solves its MPC every 2 min, inside that range. The four heater controllers are PID blocks, the WAIT controller sets their setpoints and the octane trim sets the WAIT setpoint, a cascade three deep. Industrial predictive controllers meet the two awkward cases of this chapter with standard features. When an analyzer delivers no new value, the typical answer is to skip the bias update for a number of intervals. When a lower-level controller is saturated, a temporary hard constraint keeps the move from going further the wrong way [6]. The chapter's trim holds on a stale analyzer, and its MPC drops the duty limit of a heater already at its cap.

IN PRACTICE A Regenerator That Burns on a Leash

The model's regenerator is one vessel with a burn limit. A real one works in steps, which a federal rule lists as depressurization, purge, coke burn-off, catalyst rejuvenation with a chloride compound and a final purge [24]. The burn is held back on purpose. Oxygen and air flow are regulated so that the coke burns under control, because higher temperatures cost the catalyst surface area. The screens that hold the moving catalyst plug with fines, which leaves coke unburned [7]. In the United States, 40 CFR Part 63, Subpart UUU limits the hydrogen chloride vented during coke burn-off and rejuvenation, as a percent reduction or a concentration, and requires continuous monitoring and a written operation, maintenance and monitoring plan [24].

What Is Modernizing It

The first development is the spectroscopic analyzer, which replaced a slow engine test by a spectrum read in a minute. Fiber optics let one instrument watch several points of a plant up to 100 m away [21]. The second is the soft sensor, of which the inferred octane is one. A review of such models names a use beyond prediction, the detection of a failed sensor and a back-up for it. The same review warns that after launch the prediction usually deteriorates as the process drifts, catalyst deactivation among the causes, and that telling a process drift from a sensor drift takes expert knowledge. A sensor drift calls for recalibration and for no corrective action on the process [25]. The trim of case 4 moved the process instead. The third is predictive control, of which a 2003 survey counted more than 4600 applications, with refining holding 67 % of those classified [6].

9.8 Limitations

Every size, rate constant, limit and the octane correlation are generic teaching values for a 150 t/h low-pressure continuous reformer, chosen so that the design point gives the plan's bed drops, and none was fitted to a unit. The kinetics depart from Smith's in two declared ways, a dehydrogenation constant 8.5 times his and hydrocracking first order in hydrogen. The model's 95-octane reformate carries less aromatics and its C5+ yield is about 7 volume points higher than a published correlation gives, so the slope of the trade is right and its level is not. One coke state per bed, a well-mixed regenerator and a catalyst loop compressed about twelve times make the coke lessons visible within a

day at the price of coke rates twelve times a real unit's. An aging catalyst here loses almost no yield. Pressure drop, the heater fireboxes, the fuel header, the separator pressure loop and the stabilizer's composition profile are outside the model. No refinery switches its naphtha as a step or shifts an analyzer by exactly 0.8 RON, but clean events let each effect in the cases be attributed to one cause.

The inferred octane uses maps of the commissioned plant, so it learns coke only through its offset, and its validation is a jump test against the analyzer. The MPC linearizes those maps at each execution and holds the burn on the lined-out map, its limits are soft, and it carries no guarantee beyond its hour. The heater loops do not protect the tube metals, which only the MPC, the distribution and the operator do, and the trip reads the truth and stops the run rather than model a restart. The challenge scores the burn demand, not the capped burn, and sets no price on hydrogen or fuel.

> ⚠ **OPEN CHALLENGE** **Octane With Less Benzene**
>
> A reformer makes octane by making aromatics, and the lightest aromatic is regulated. A review of 2003 called the limits on benzene and vapor pressure that followed the Clean Air Act Amendments of 1990 the biggest challenge facing reforming, one that forced a complete rethink of the process. It expected modified catalysts, new processes after the reformer and reluctance to reform the full range of naphtha [13]. The federal fuel rule now holds a manufacturer's gasoline to an average of 0.62 volume percent benzene [26]. The four lumps of this chapter fold benzene precursors and heavier aromatics into one aromatic, so its octane target cannot see the limit. Progress would be a severity target that weighs octane, yield, hydrogen and benzene together.

9.9 Conclusion

A continuous catalytic reformer is a train of endothermic beds whose first bed does most of the work, run on an octane that is never measured where it is made, and this simulator computes the cell temperatures, the coke on each bed, the analyzer, the laboratory and an inferred octane together from a forty-three-state model. At design the beds cool the gas by 70.4, 45.1, 11.8 and 7.2 K, and heat added to the charge heater is mostly absorbed by bed 1, 9.3 K of a 15.7 K rise. Ten kelvin of WAIT buy 2.01 RON for 1.51 points of yield, until the regenerator's capacity, reached near 100 RON on the design naphtha, cuts what each further kelvin buys. A trim that trusts a shifted analyzer holds its reading while the true octane falls to 94.24, a feedforward from N+2A keeps a leaner naphtha's octane above 94.8 where feedback alone loses 112 min, and the aged distribution keeps the charge heater under its tube-metal limit. The graded day passes with the inferred octane and fails on any strategy that believes the analyzer, because the shift outlasts what the target allows. The octane correlation is illustrative and the yield level high. Further work would add plug-flow beds with a coke profile, a regenerator with burn and chlorination zones, a pressure and stabilizer model, and an octane correlation fitted to measured reformates. The simulator shows that a reformer is controlled through its temperature profile and its measurements as much as through its chemistry.

KEY POINTS

- A reformer cools its own feed, so four fired heaters must put the heat back between the beds.
- The first bed absorbs most of a heater step, 9.3 K of a 15.7 K rise.
- Ten kelvin of WAIT buy 2.01 RON for 1.51 points of yield, until the regenerator reaches its burn capacity.
- A trim that trusts a shifted analyzer holds its reading while the true octane falls to 94.24.
- The graded day passes with the inferred octane and fails on any strategy that believes the analyzer.

Where else this applies. A soft sensor corrected by a delayed laboratory returns in Volume 5, Chapter 10, the same pair of octane instruments in Chapter 15, beds in series that amplify where these damp in Chapter 7, and an optimum on a moving constraint in Volume 9, Chapter 8.

References

[1] P. Doolin, D. Zalewski, and S. Oyekan, "Catalyst regeneration and continuous reforming issues," in *Catalytic Naphtha Reforming*, 2nd ed., ser. Chemical Industries, G. J. Antos and A. M. Aitani, Eds. New York: Marcel Dekker, 2004, vol. 100, pp. 414–435.

[2] L. Turpin, "Control systems for commercial reformers," in *Catalytic Naphtha Reforming*, 2nd ed., ser. Chemical Industries, G. J. Antos and A. M. Aitani, Eds. New York: Marcel Dekker, 2004, vol. 100, pp. 470–518.

[3] U. Taskar and J. B. Riggs, "Modeling and optimization of a semiregenerative catalytic naphtha reformer," *AIChE Journal*, vol. 43, no. 3, pp. 740–753, 1997.

[4] D. Li, S. L. Shah, T. Chen, and K. Z. Qi, "Application of dual-rate modeling to CCR octane quality inferential control," *IEEE Transactions on Control Systems Technology*, vol. 11, no. 1, pp. 43–51, 2003.

[5] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[6] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[7] A. Gupta and S. K. Gupta, "Catalyst regeneration techniques in naphtha reforming: Short review," *Chemical and Process Engineering*, vol. 43, no. 2, pp. 101–108, 2022.

[8] K.-m. Liang, H.-y. Guo, and S.-w. Pan, "A study on naphtha catalytic reforming reactor simulation and analysis," *Journal of Zhejiang University-SCIENCE B*, vol. 6, no. 6, pp. 590–596, 2005.

[9] A. J. Lazar and J. H. Klatt, "Modern control of a catalytic reformer in distributed control," in *Proc. 1983 American Control Conference*, Jun. 1983, pp. 7–13.

[10] V. Haensel and M. J. Sterba, "Comparison of Platforming and thermal reforming," in *Progress in Petroleum Technology*, ser. Advances in Chemistry. American Chemical Society, 1951, pp. 60–75.

[11] R. Larraz and R. Arvelo, "Modeling catalytic naphtha reforming," in *Catalytic Naphtha Reforming*, 2nd ed., ser. Chemical Industries, G. J. Antos and A. M. Aitani, Eds. New York: Marcel Dekker, 2004, vol. 100, pp. 519–561.

[12] V. Haensel, "Process of reforming a gasoline with an alumina-platinum-halogen catalyst," U.S. Patent 2 479 110, Aug. 16, 1949.

[13] U. T. Turaga and R. Ramanathan, "Catalytic naphtha reforming: Revisiting its importance in the modern refinery," *Journal of Scientific & Industrial Research*, vol. 62, no. 10, pp. 963–978, 2003.

[14] H. E. Kluksdahl, "Reforming a sulfur-free naphtha with a platinum-rhenium catalyst," U.S. Patent 3 415 737, Dec. 10, 1968.

[15] E. L. Pollitzer, "Continuous Platforming," *Platinum Metals Review*, vol. 20, no. 1, pp. 2–6, 1976.

[16] J. J. Kelly, C. H. Barlow, T. M. Jinguji, and J. B. Callis, "Prediction of gasoline octane numbers from near-infrared spectral features in the range 660–1215 nm," *Analytical Chemistry*, vol. 61, no. 4, pp. 313–320, 1989.

[17] K. Moljord, H. G. Hellenes, A. Hoff, I. Tanem, K. Grande, and A. Holmen, "Effect of reaction pressure on octane number and reformate and hydrogen yields in catalytic reforming," *Industrial & Engineering Chemistry Research*, vol. 35, no. 1, pp. 99–105, 1996.

[18] J. H. Gary, G. E. Handwerk, and M. J. Kaiser, *Petroleum Refining: Technology and Economics*, 5th ed. Boca Raton, FL: CRC Press, 2007.

[19] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[20] U.S. Federal Trade Commission, "Automotive fuel ratings, certification and posting," Code of Federal Regulations, Title 16, Part 306, section 306.0, edition of January 1, 2026, 2026, govinfo.gov.

[21] C. Pasquini, "Near infrared spectroscopy: fundamentals, practical aspects and analytical applications," *Journal of the Brazilian Chemical Society*, vol. 14, no. 2, pp. 198–219, 2003.

[22] U.S. Environmental Protection Agency, "EPA chemical accident investigation report: Tosco Avon refinery, Martinez, California," U.S. Environmental Protection Agency, Chemical Emergency Preparedness and Prevention Office, Washington, DC, Tech. Rep. EPA 550-R-98-009, 1998.

[23] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[24] U.S. Environmental Protection Agency, "National emission standards for hazardous air pollutants for petroleum refineries: Catalytic cracking units, catalytic reforming units, and sulfur recovery units," Code of Federal Regulations, Title 40, Part 63, Subpart UUU, sections 63.1566, 63.1567 and 63.1579, edition of July 1, 2025, 2025, govinfo.gov.

[25] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[26] U.S. Environmental Protection Agency, "Regulation of fuels, fuel additives, and regulated blendstocks," Code of Federal Regulations, Title 40, Part 1090, section 1090.210, edition of July 1, 2024, 2024, govinfo.gov.

9.10 Exercises

9.1 Concept check · 2 min

In case 1 the reformer is in Manual and heater 1 receives about 1.6 MW more, enough for about 10 K at its outlet. Two hours later heater 1's outlet has risen 15.7 K, yet bed 1's outlet has risen only 6.3 K. Why does so little reach bed 2?

(a) The hotter bed reacts faster, and its endothermic reactions keep most of the heat in a larger drop

(b) Heater 1 is held at its rated duty, so most of the extra firing never reaches the gas at all

(c) The feed–effluent exchanger takes heat back out of heater 1's inlet as the outlet warms

(d) The recycle hydrogen rises with the temperature and spreads the added heat over more gas

9.2 Concept check · 2 min

In case 6 the recycle compressor is limited to 62 % speed while the unit takes 150 t/h of naphtha, and H_2/HC falls to 2.08, under its floor of 2.5. Which move brings the ratio back over the floor?

(a) Raise the H_2/HC setpoint, so that the loop asks the compressor for more speed

(b) Cut the naphtha feed rate, which lowers the hydrocarbon moles in the ratio

(c) Raise the WAIT, which makes more hydrogen and so raises the recycle purity

(d) Raise the catalyst circulation, so that the regenerator burns more coke

9.3 Concept check · 1 min

Lined out on the design naphtha at $H_2/HC = 3$, five kelvin of WAIT buy 1.1 to 1.2 RON below a WAIT of 523.5 °C and about 0.5 RON above it. What happens at that WAIT?

(a) The octane correlation stops being linear in the aromatics above 100 RON

(b) Heater 4 reaches its rated duty and can no longer lift bed 4's inlet

(c) The coke made reaches the regenerator's burn capacity, and coke returns to bed 1

(d) Dehydrogenation reaches equilibrium in bed 1, so no more aromatics can form

9.4 Theory and calculation · 7 min

At the design point every bed inlet is at 500 °C, the bed drops are 70.4, 45.1, 11.8 and 7.2 K, the charge carries $\dot{C}_p = 160.5\,\mathrm{kW/K}$, and the heaters fire 8.72, 11.30, 7.25 and 1.90 MW of their rated 12, 15, 10 and 5 MW, with $\theta_m = 110\,\mathrm{K}$.

(a) Compute each bed's reaction heat and the total endotherm. *Checkpoint:* 11.30, 7.24, 1.89 and 1.16 MW; 21.6 MW.

(b) Compute the four tube metals from Equation (9.4). *Checkpoint:* 579.9, 582.9, 579.8 and 541.8 °C.

(c) In case 7 all four inlets are at 520.2 °C. What share of its rated duty can a heater fire before its tube metal reaches 620 °C, and what duty is that for heater 1? At $t = 0$ heater 1 fires 10.62 MW. What is its tube metal, and at what outlet would a heater at rated duty reach the 640 °C trip? *Checkpoint:* 0.907 and 10.89 MW; 617.5 °C; 530 °C.

9.5 Theory and calculation · 6 min

The four beds hold 66 t of catalyst, circulated at $W_c = 8.25\,\mathrm{t/h}$. A full burn leaves $C_r = 0.1$ wt% and the regenerator can burn $B_{max} = 400\,\mathrm{kg/h}$.

(a) How long does the catalyst take to pass once through the beds? *Checkpoint:* 8.0 h.

(b) At design the spent catalyst leaving bed 4 carries 2.99 wt%. Use Equation (9.6) to compute the burn and its share of the capacity. *Checkpoint:* About 238 kg/h, 60 %.

(c) At what coke on the spent catalyst does the burn reach its capacity? *Checkpoint:* 4.95 wt%.

(d) In case 6 the spent catalyst carries 6.02 wt% at 24 h. What coke does the regenerated catalyst return to bed 1 with, once the regenerator has lined out? *Checkpoint:* 1.17 wt%.

9.6 Theory and calculation · 6 min

The design naphtha carries 21 vol% of naphthenes and 20.5 vol% of aromatics, the paraffinic one 16 vol% of each. The trim's feedforward is Equation (9.10) with $K_{ff} = 1.4\,\mathrm{K}$ per unit of N+2A.

(a) Compute N+2A for both naphthas and the feedforward after the switch. Lined out, the paraffinic naphtha needs a WAIT of 520.2 °C for 95 RON. How much must the trim's own output supply? *Checkpoint:* 62 and 48; 19.6 K; 0.6 K.

(b) The trim's output limits are the WAIT range, 480 to 540 °C, less the feedforward. Give the trim's own range on the paraffinic naphtha, and say why the limits are written this way. *Checkpoint:* 460.4 to 520.4 °C.

(c) The challenge judges 1380 min and requires the octane in 95 ± 0.3 for 95 % of it. The NIR shifts at 16 h, and the next knock-engine sample is drawn at 18 h and reported an hour later. How many minutes may the run spend outside the band, and for at least how long does a trim that believes the NIR act on the shifted reading? *Checkpoint:* 69 min; 180 min.

9.7 App challenge · 8 min

Load case 1, "The staircase: four drops, four rises (Manual)", at the default seed. Before pressing Run, raise Heater 1 duty (TIC · four heater outlets card) from 8.72 to 10.30 MW. Run to 120 min at speed 20× (about 6 s) and export the CSV.

(a) From the columns `bed1_in_C` and `truth_drop1_K`, read heater 1's outlet and bed 1's drop at 0 and 120 min. How much of the outlet's rise does bed 1 keep in its drop, and how much reaches bed 1's outlet? *Checkpoint:* 500.0 to 515.7 °C (±0.8 K); drop 70.39 to 79.65 K. About 9.3 K is kept and 6.3 K passed on.

(b) The extra 1.58 MW over $\dot{C}_p = 160.5\,\mathrm{kW/K}$ accounts for about 9.8 K. Where does the rest of the outlet's rise come from, and why does the outlet keep climbing for an hour after the step? *Checkpoint:* The feed–effluent exchanger, about 6 K. The heat goes round the train and back.

9.8 App challenge · 9 min

Load case 4, "The octane trim on a late analyzer and an 8 h laboratory", at the default seed and speed 20×. It loads with Severity control on Octane trim and Octane measurement on NIR + lab, and the NIR starts reading 0.8 RON high at 30 min.

(a) Run as shipped to 300 min and export the CSV. In `truth_ron_product`, find the lowest true octane, the first row under 94.7 and the first row back over 94.7. Read `nir_bias_est` after the laboratory result at 210 min. *Checkpoint:* 94.24 (±0.02); 53.5 min; 248 min; −0.83.

(b) Select Inferred + lab in Octane measurement, press Reset and run to 300 min again. What is the lowest true octane now, and how does the WAIT setpoint (`wait_sp_C`) compare with run (a)? *Checkpoint:* 94.97. It stays near 500 °C where run (a) lowered it to 495.8 °C.

(c) Why is the change in (b) followed by Reset, rather than made at 60 min into run (a)? *Checkpoint:* By then the trim has already lowered the WAIT and the octane.

9.9 App challenge · 8 min

Load case 9, the challenge, at the default seed. Before pressing Run, set Severity control to Octane trim, N+2A feedforward to On and Inlet distribution to Aged, and leave Octane measurement on NIR + lab. Set the speed to 300× (the day takes about 5 s) and run to the stop at 24 h.

(a) Read the verdict. Which targets are met and which are missed? *Checkpoint:* Missed: 84.7 % in the band. Met: tube metal 607 °C, burn demand 360 kg/h, H_2/HC 2.89, C5+ 91.22 wt%.

(b) Select Inferred + lab, press Reset and run the day again. Read the verdict, and say which of the day's three events the change dealt with. *Checkpoint:* Passes: 100.0 %, 606 °C, 364 kg/h, 2.90, 91.16 wt%. It deals with the NIR shift at 16 h.

9.10 Application area · 4 min

The heater loops of this simulator do not watch the tube metal, and a WAIT setpoint stepped 30 K at once in Auto trips heater 3 3.9 min later. On a real reformer, name two layers that keep a tube metal under its limit, say where the limit itself comes from, and explain why the simulator leaves a tube-metal override out of its heater loops.

9.11 Application area · 4 min

Near design one octane number of giveaway costs about three quarters of a point of C5+ yield.

(a) On 150 t/h of naphtha, what does a trim that holds the reformate 0.5 RON above target cost in reformate per hour? *Checkpoint:* About 0.56 t/h.

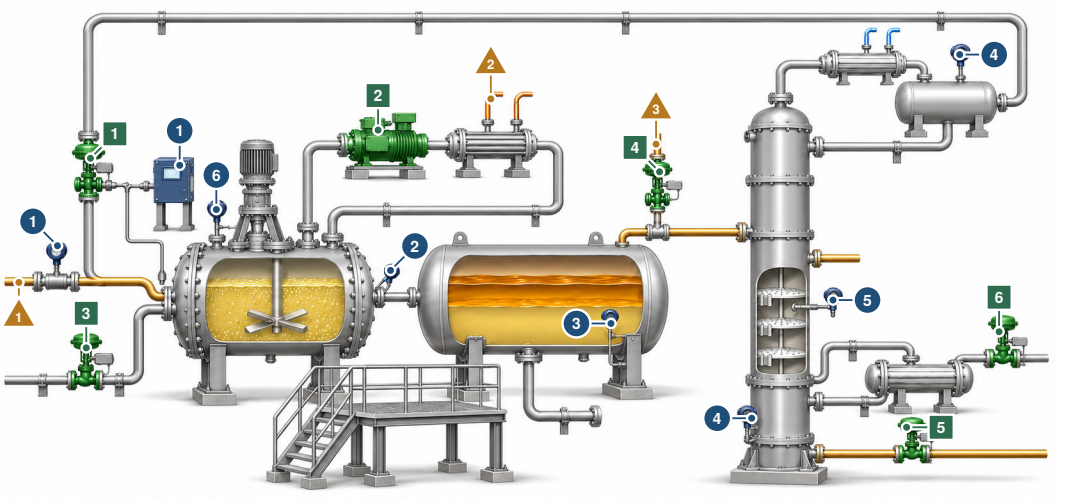
(b) The model's yield sits about 7 volume points above a published chart while its slope agrees. Does the answer to (a) survive that, and why? *Checkpoint:* Yes, because it depends on the slope.

9.12 Review the agent · 5 min

An AI assistant reviews case 4 and writes: "The true octane falls because the NIR reports about 20 min late. Raise the octane trim gain from 2 to 4.0 K per RON and cut its integral time from 20 to 10 min so the trim corrects sooner, and the octane will stay within 0.3 of 95." Find the flaw, and say how you would check the claim in the app.

CHAPTER 10

Three Clocks Behind One Octane in a Sulfuric Acid Alkylation Unit



The controller moves 1 valve on the isobutane recycle 2 refrigeration compressor 3 fresh acid valve 4 isobutane makeup valve 5 alkylate valve at the column bottom 6 steam valve of the reboiler

The controller reads 1 olefin flow, isobutane in the recycle 2 contactor temperature 3 acid density 4 receiver and column-bottom levels 5 column tray temperature 6 compressor suction pressure

What disturbs it ▲1 more olefin, propylene or water in the feed ▲2 warm cooling water ▲3 a lost isobutane makeup

Notice the three stocks side by side: the cold made by the compressor loop above the stirred contactor, the isobutane that the column's receiver sends back from the right, and the acid lying under the hydrocarbon in the settler. Illustration, not to scale.

IN PLAIN TERMS Three Clocks Behind One Octane

The main idea. An alkylation unit joins two light liquids into a clean gasoline that resists engine knock, a quality called octane. A strong acid makes them join, and the unit must stay cold, rich in one of the two liquids and strong in its acid.

The challenge. The cold changes in minutes, the mix of liquids in hours and the acid in more than a day. The acid is tested by a lab only every four hours, and a weak acid spoils itself faster and faster.

Why it matters. The acid is the largest running cost, and the same waste that weakens it also lowers the octane.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the History boxes in Sections 10.1 and 10.2, the Did you know? boxes in Sections 10.1 and 10.6, the Try it boxes in Sections 10.3 and 10.4, and Key points in Section 10.9. No equations.
Study	65 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	25 min	This page, Section 10.6 with Figure 10.5, Section 10.7 and Section 10.8 with its Open challenge box.

The control question The octane of a sulfuric acid alkylation unit rests on three stocks that answer on three clocks, the contactor's cold in minutes, the recycle's isobutane in hours and the acid's strength in more than a day, while the acid is titrated only every four hours. Which loops, sensors and supervisors hold the octane and the acid bill through a day of upsets, and where does each of them lose?

Learning objectives

- Compute the fresh acid that holds a strength from the acid balance, and the acid-soluble oil the activity law adds as the strength falls.
- Explain why a ratio loop on the flows reads on setpoint while the true isobutane-to-olefin ratio drifts with the recycle's purity.
- Compare a strength loop on a four-hourly titration with one on a density soft sensor, and state what the soft sensor cannot see.
- Run the acid to its runaway floor and the graded day, and read each outcome from the app's verdict and exported log.
- Judge whether a predictive controller's lead comes from its model or its plan, and test it on a plant whose acid law has changed.

Prerequisites Sections 1.4.6, 1.4.7, 1.6.2 and 1.6.3, with Section 1.4.3 for the column. Chapter 8 for a late analyzer in this volume.

The app https://apmonitor.com/apps/alkylation-unit/

See also Volume 5, Chapter 10 for a soft sensor corrected by a delayed laboratory and Volume 5, Chapter 3 for a loop on a slow analyzer. Volume 9, Chapter 4 for an offset-free predictive controller with a bias in its estimator and Volume 9, Chapter 8 for a supervisor whose optimum sits on a moving constraint.

10.1 Introduction

Nine moles of isobutane pass the contactor of an alkylation unit for every mole of olefin that reacts there. In the generic unit simulated here, 530 kmol/h of butenes meet that excess at 7 °C over 80 t of acid at 92 wt% and become 60.0 t/h of alkylate with a research octane number (RON) of 95.00, while 251 kg/h of the olefin turns into acid-soluble oil (ASO) that dilutes the acid. A fresh acid flow of 3831 kg/h, 63.8 kg per tonne of alkylate, keeps the strength, and a compressor that boils 2321 kmol/h of light ends off the contactor, at 85 % of its capacity, keeps the cold. Each of the three stocks moves at its own speed. A cut of 0.05 bar in the compressor's suction pressure takes the contactor 63 % of the way to its new temperature in 2.1 min, 15 % more olefin drifts the recycle's isobutane purity 63 % of the way in 7 h 00 min, and 20 % less fresh acid needs 36.3 h to move the strength as far.

Alkylate is valued because it carries octane without aromatics or olefins, and Chapter 15 blends it with reformate and cracked gasoline toward a tank-average octane. Sulfuric acid is one of the two acids on which the commercial process is run [1–3]. The acid is also its largest running cost. A process patent puts sulfuric acid consumption at approximately 30 % of the overall process costs and calls an improvement of half an octane number significant [4], and feedstock consumption, acid usage and alkylate quality have been compared across the processes in use [5]. Every operating decision therefore trades octane against acid, and the acid is consumed by the same side reaction that lowers the octane.

+ DID YOU KNOW? Seven Percent of the Crude Capacity

On January 1, 2003 the refineries of the United States could alkylate 1,170,019 barrels per calendar day, 7.0 % of their crude distillation capacity [6]. The process shrinks what it is fed. Two light feedstocks become a denser product, and the volume falls by about 30 % at the same weight [7]. The unit of this chapter makes 60.0 t/h of alkylate.

Two beliefs are corrected by this simulator. The first is that a ratio loop reading on setpoint holds the ratio. In case 2 the loop counts the recycle as 89.3 % isobutane and reads an isobutane-to-olefin

ratio (I/O) of 9.0 throughout, while the true ratio is 8.88 at 4 h and 8.75 at 12 h, because the recycle's composition belongs to a distillation column and its n-butane can leave only through a side draw. The second is that the acid strength is a fast variable to be held by a fast loop. It is an inventory renewed at 4.79 % an hour, it is measured four-hourly, and below about 89 wt% it accelerates its own decline. With the fresh acid withheld from 90 wt%, the strength falls 0.35 wt% in the first hour and 0.52 in the last, and reaches the declared floor of 87 wt% after 7 h 30 min.

The class of problem is general. A quality that depends on a recycle through a separation column inherits the column's inventories, the setting in which snowball effects in reactor and separator processes with recycle were analyzed [8]. A quality measured too rarely to close a loop on is inferred from secondary measurements, which inferential control [9] and the soft sensors of the process industry [10] have addressed for decades. Section 1.4.6 sets out the analyzer delay, the soft sensor and the laboratory bias update that this unit combines. Setpoints chosen by a steady-state economic optimization above the loops are the real-time optimization of current practice [11]. A catalytic reactor whose catalyst is a slowly renewed inventory, assayed in a laboratory, with a recycle through a column and refrigeration that runs out on a hot day, has all three at once. Between the loops and the optimizer, industry places model predictive control [12], and minimum-time constrained control of the acid strength has been published for a sulfuric acid alkylation unit [13].

> ⧖ HISTORY **1997: A Controller Takes Over the Fresh Acid**
>
> In 1996 a refinery in France decided to install a controller to limit the acid consumption of its alkylation unit [13]. Extra acid lowers the risk and raises the cost, so the operators tried to hold the unit just above a minimum strength, below which undesired reactions cause serious malfunctions. The unit was too slow to steer by hand, since a change of the fresh acid flow was fully felt only after about a week. The controller estimated the strength in the second settler from an analysis made downstream and late, and it solved a minimum-time problem as successive linear programs every 15 min. It entered full service in January 1997, ran more than 98 % of the time and saved about 5 % of the cost of the acid. It surprised the operators at first by moving the fresh acid before the measured strength had changed.

However, the alkylation literature treats the chemistry, the kinetics and the process apart [14–16], and the control literature treats ratio loops, soft sensors and optimizers as separate structures [9–11]. This chapter puts numbers on how ratio, temperature and strength loops, a laboratory-corrected soft sensor, a moisture feedforward and an economic supervisor hold the octane and the acid bill through a day of olefin, water and heat upsets, using a browser simulator of a sulfuric acid alkylation unit that logs the true ratio, strength and octane beside every instrument. The selectivity law, the activity law and the octane correlation are declared teaching relations of published direction, and the controllers are standard ones. One finding is that neither the soft sensor nor the moisture feedforward can be made necessary by a target on a single day, because the soft sensor's blindness to water trims back the feedforward's acid. A model predictive controller added as a fifth structure gives a second finding, that its lead on a four-hourly titration belongs to its acid balance reading the feed more than to its plan, and that the lead is lost once the plant's acid law departs from that balance.

Section 10.2 traces the process and the evidence behind its octane and acid laws. Section 10.3 builds the unit, a 108-state model whose hydrocarbon and acid balances close to 7.4×10^{-8} over a day, with a 24-tray deisobutanizer on the column module the apps share and a boundary that keeps every controller to the instruments. Section 10.4 measures its three clocks, its floor and its trip. Section 10.5 adds the loops, the soft sensor, the supervisor and the predictive controller. Section 10.6 runs the nine cases and the graded day, each of which credits an action only when the unit responds to it. Section 10.7 sets the hardware of a real unit beside the model, Section 10.8 states the limitations, Section 10.9 concludes, and the exercises of Section 10.10 work through the cases.

10.2 Isoparaffin Alkylation, Acid Strength and Octane

The process dates from 1939, and its acid strength was an economic question from the start, as the History box below tells. A thesis of 1969 reproduces Putney's pilot-plant curves of F-2 (motor) clear octane against acid strength, which reach a maximum between 95 and 96 wt% titratable acidity [17]. Read from those four curves, the octane falls by 0.33 to 0.39 numbers per point of strength between 94 and 90 wt%. Its review also records that one gallon of butadiene depletes 58 lb of fresh acid from 98 to 90 wt%, that water in the feed reduces the acidity, and that a reactor temperature between 40 and 50 °F was then recommended, because warmer raised the acid consumption and colder thickened the acid [17].

> ⧖ HISTORY **1939 to 1969: A Wartime Fuel and the Cost of Its Acid**
>
> The addition of olefins to isoparaffins to make high-octane isoparaffinic fuels was described in 1939 [18]. A federal manual of refining dates the process to the early 1940s, when it produced more high-octane aviation gasoline [19]. By 1949 the alkylation of isobutane with propylene over sulfuric acid was reported as a commercial production [20]. Acid strength was an economic question from the start. A study of lower acidity sulfuric acid alkylation appeared in 1951 [21]. A thesis of 1969 reports from it that the alkylate yield could be held at lower acidity by raising the reactor temperature, and concludes that some optimum strength based on minimum cost must exist [17].

Kinetics followed. A 17-reaction mechanism for isobutane and propylene, measured in a continuous stirred tank over 95 wt% acid, held from 81 to 135 °F and departed from it at 65 °F, with a change of selectivity reported when the acid was lowered to 90 wt% [15]. The chemistry, catalysts and processes were reviewed in 1993 [1], the physicochemical sequence of events in the sulfuric acid process in 2003 [14], and a kinetic model of olefin and isobutane alkylation over sulfuric acid was published in 2022 [16]. A process patent of 1992 moved the contactor colder, to −6 to 0 °C over 92 to 95 wt% acid. Its examples report about half an octane number more at 28 °F than at 49 °F, acid consumption of 0.197 against 0.323 lb per gallon of alkylate, and maximum octane at 93 to 94 wt% with 0.5 to 1 wt% water [4].

The textbooks describe the unit this simulator is drawn from: a refrigerated contactor, an acid settler, a compressor on the flashed light ends and a deisobutanizer that returns the isobutane [3]. No commercial unit's control configuration was found in the literature consulted here, so the loops below follow the structure the process implies and are not offered as any plant's practice.

10.3 The Unit as Modeled

The unit is generic, sized without any licensor's data, as shown in Figure 10.1. A cracked C_4 cut of 963.6 kmol/h (55 % butenes, 30 % isobutane, 15 % n-butane) and the isobutane recycle enter a stirred contactor over acid. The cut comes from a catalytic cracker such as that of Chapter 6. The hydrocarbon boils at the pressure a compressor leaves it, the vapor is condensed against cooling water and returns at once. The settler sends the net hydrocarbon to a deisobutanizer (DIB), whose receiver supplies both the reflux and the recycle, whose vapor side draw takes the n-butane and whose bottoms is the alkylate. Fresh acid at 98.5 wt% enters the contactor and spent acid leaves over the settler's weir. The engine's clock is the minute, hydrocarbon flows are in kmol/h, acid flows in kg/h and heat in megawatts.

The variables are classified in Table 10.1. The integrated state holds the contactor's isobutane and n-butane fractions, its temperature, the compressor's vapor flow, the acid's three masses, the product octane, 78 column states (a holdup and two fractions on each of the receiver, 24 trays and the sump) and eight flow-loop lags, 94 physical states, and 14 ledgers that integrate what enters, leaves and reacts so that each balance can be checked.

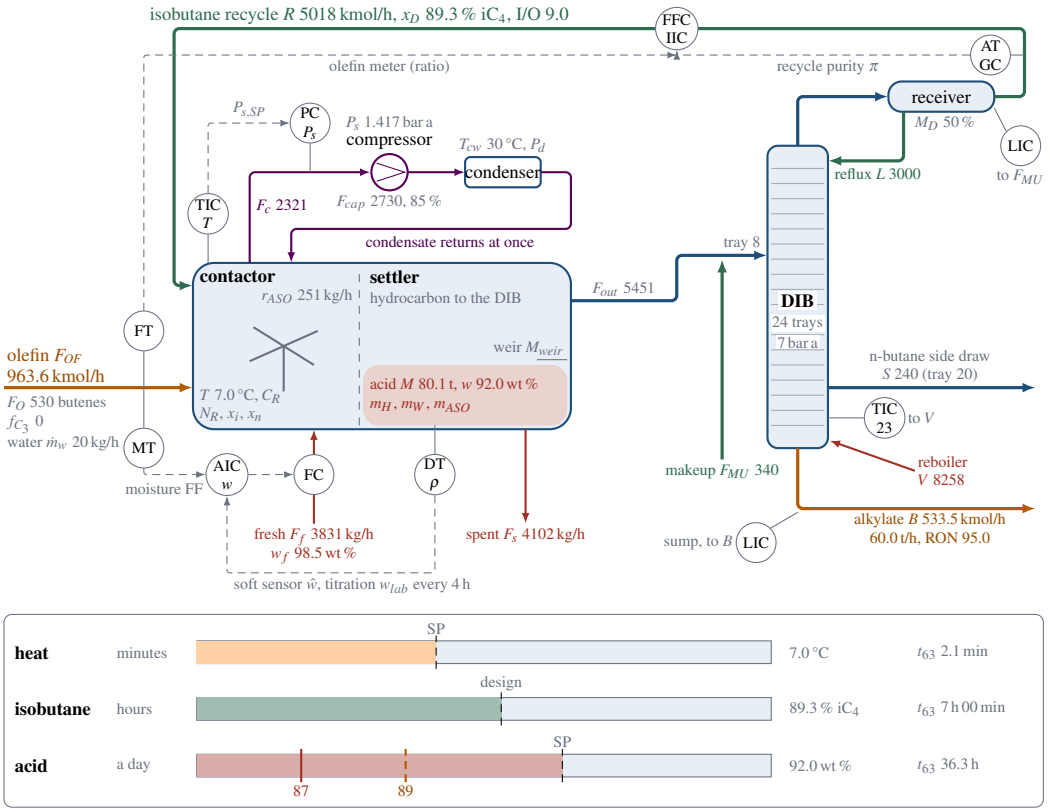


Figure 10.1 The sulfuric acid alkylation unit as modeled, case 1 at $t = 0$, after the app's Unit view, at the design values. The olefin feed F_{OF} (olefin F_O, propylene share f_{C_3}, water $\dot{m}_w$) and the recycle R of purity x_D meet in the contactor (N_R, C_R, T, x_i, x_n), the compressor holds the suction pressure P_s that the TIC asks for up to its capacity F_{cap}, which falls as the condenser's water T_{cw} warms, and the settler's acid (M, m_H, m_W, m_{ASO}, strength w) is renewed by F_f at w_f and leaves over the weir as F_s. The net hydrocarbon F_{out} and the makeup F_{MU} feed tray 8 of the DIB, the side draw S leaves from tray 20, the receiver M_D supplies the reflux L and R, and the reboiler's boilup V is set on tray 23. The IIC sets R in ratio to the olefin meter with the recycle chromatograph, the AIC sets F_f on the density soft sensor or the titration with the moisture feedforward, and two LICs hold the receiver and the sump. The strip below shows the three inventories, each against its setpoint or design mark, with their 63 % times. The drawing shows the model, not the piping or a safety layout, and is not to scale.

10.3.1 How the Olefin Splits

The olefin reacts within seconds and is not stored. It splits between alkylation, isobutane plus olefin to alkylate, and conjunct polymerization to ASO, in the ratio of two declared rates, as shown in Equation (10.1).

$$\rho = \frac{r_{ASO}}{r_{alk}} = \kappa_0 \, \mathrm{e}^{\theta(T-7)} \left(\frac{9}{I}\right)^n \frac{1 + c_3 f_{C_3}}{a(w)}$$

$$r_{alk} = \frac{F_O}{1+\rho}, \qquad r_{ASO} = F_O - r_{alk} \tag{10.1}$$

$$I = \frac{R\,x_{D,i} + z_I F_{OF}}{F_O}$$

Table 10.1 Variables, Their Roles and Design Values (Case 1 at $t = 0$)

Variable	Role	Design	Range
I	CV, isobutane to olefin	9.00	SP 6–14
R	MV, recycle	5017.5 kmol/h	0–9000
T	CV, contactor	7.00 °C	SP 2–14
$P_{s,SP}$	MV, suction pressure	1.417 bar a	0.6–2.5
w	CV, acid strength	92.00 wt%	SP 88–96
F_f	MV, fresh acid	3831 kg/h	0–12 000
M_D, M_B	CV, receiver, sump	50 %, 50 %	SP 20–80
F_{MU}, B	MV, makeup, bottoms	339.9, 533.5 kmol/h	0–1200, 0–1500
T_{23}	CV, DIB tray 23	87.26 °C	SP 70–110
V	MV, boilup	8257.5 kmol/h	0–10 500
F_{OF}	DV, olefin feed	100 %	40–130
$\dot{m}_w$	DV, feed water	20 kg/h	0–1000
T_{cw}	DV, cooling water	30 °C	20–42
f_{C_3}	DV, propylene in olefin	0 %	0–60
—	DV, makeup supply	100 %	0–100
x_D	state, recycle iC_4	89.31 %	—
F_c, F_{cap}	state, output: compressor	2321, 2730 kmol/h	—
M	output, acid held	80.05 t	—
RON	state, product	95.00	—

where $F_O = z_O F_{OF}$ is the olefin, $z_O = 0.55$ and $z_I = 0.30$ are the feed's olefin and isobutane shares, $x_{D,i}$ is the recycle's isobutane fraction, $\kappa_0 = 0.0085$, $\theta = 0.06\,\mathrm{K}^{-1}$, $n = 1.5$, $c_3 = 1.5$ and I is floored at 0.5 inside the law. Alkylation runs in proportion to the acid's activity, which is the declared law of Equation (10.2).

$$a(w) = \frac{\mathrm{e}^{-\beta(92-w)}}{1 + A\,\mathrm{e}^{(w_r - w)/s}} \tag{10.2}$$

with w in wt%, $\beta = 0.06$ per point, $A = 0.5$, $w_r = 87$ and $s = 0.6$. Near design the ASO rises 6.25 % per point of strength lost, and $1/a$ is 1.063, 1.131, 1.219, 1.391 and 2.025 at 91, 90, 89, 88 and 87 wt%. At design $a = 0.99988$, $\rho = 0.008501$, 525.53 kmol/h is alkylated and 4.468 kmol/h of olefin becomes 250.7 kg/h of ASO. The alkylate leaves at $r_{alk}(M_i + M_O)$, 60.032 t/h, with $M_i = 58.12$ and M_O the olefin's molar mass, 56.11 kg/kmol for butenes and 42.08 for propylene. The collapse replaces a linear activity law, which made the design point unstable at a fixed fresh acid.

10.3.2 Contactor, Settler and Autorefrigeration

The contactor and settler are one stirred hydrocarbon holdup N_R at fixed moles, so the olefin consumed leaves the outflow, as shown in Equation (10.3).

$$\begin{aligned} N_R \frac{\mathrm{d}x_i}{\mathrm{d}t} &= R\,x_{D,i} + z_I F_{OF} - r_{alk} - F_{out}\,x_i \\ N_R \frac{\mathrm{d}x_n}{\mathrm{d}t} &= R\,x_{D,n} + z_N F_{OF} - F_{out}\,x_n \\ F_{out} &= R + F_{OF} - F_O \end{aligned} \tag{10.3}$$

where $z_N = 0.15$, the alkylate fraction is $1 - x_i - x_n$, flows are divided by 60 in the balances because the clock is the minute, and $N_R = 1500\,\mathrm{kmol}$. The temperature follows an energy balance in which

the vapor the compressor takes carries the heat away, as shown in Equation (10.4).

$$\begin{aligned} C_R \frac{\mathrm{d}T}{\mathrm{d}t} &= Q_{rx} + Q_{mix} + (R + F_{OF}) c_p (T_f - T) \\ &\quad - \lambda F_c \\ Q_{rx} &= r_{alk}(-\Delta H_a) + r_{ASO}(-\Delta H_s) \\ P_s &= P_{bub}(T, x) - F_c / k_v \\ \tau_c \frac{\mathrm{d}F_c}{\mathrm{d}t} &= \min\left(k_v (P_{bub} - P_{s,SP}),\ F_{cap}\right) - F_c \end{aligned} \tag{10.4}$$

where the contactor's heat capacity is $C_R = 300\,\mathrm{MJ/K}$, the heats of alkylation and polymerization are $-\Delta H_a = 80$ and $-\Delta H_s = 90\,\mathrm{MJ/kmol}$, the mixers add $Q_{mix} = 1.5\,\mathrm{MW}$, the hydrocarbon's molar heat capacity is $c_p = 0.13\,\mathrm{MJ/(kmol\,K)}$ and the feed enters at $T_f = 10\,^\circ\mathrm{C}$. The coefficient $k_v = 28\,000\,\mathrm{kmol/(h\,bar)}$ is the vapor leaving the boiling liquid per bar over the suction pressure, and the compressor's speed governor, equipment rather than a loop, follows with $\tau_c = 0.5\,\mathrm{min}$. The latent heat is $\lambda = R_g B = 21.618\,\mathrm{MJ/kmol}$ for every component. At design $Q_{rx} = 11.790\,\mathrm{MW}$, the mixers and the feed add 1.5 and 0.648 MW, and the compressor carries 2321.1 kmol/h. The bubble pressure uses the deisobutanizer's own constant-volatility law, as shown in Equation (10.5).

$$P_{bub} = \sum_j x_j P_0 \,\mathrm{e}^{B(1/T_{b,j} - 1/T)}, \qquad \frac{1}{T_{b,j}} = \frac{1}{T_{b,n}} + \frac{\ln \alpha_j}{B} \tag{10.5}$$

with $P_0 = 1.013\,25\,\mathrm{bar}$, $B = 2600\,\mathrm{K}$, $T_{b,n} = 272.65\,\mathrm{K}$ and $\alpha = (1.3, 1, 0.1)$ for isobutane, n-butane and alkylate, so isobutane boils at −7.80 °C and the alkylate at 86.3 °C. At 7 °C the contactor's liquid has a bubble pressure of 1.5000 bar, the suction is 1.4171 bar a, and the contactor runs 5.39 K above isobutane's own saturation temperature at that pressure because its liquid also holds n-butane and alkylate.

10.3.3 The Compressor and Its Condenser

A fixed-volume machine moves fewer moles at a lower suction pressure, and fewer again as a warmer condenser raises its pressure ratio, as shown in Equation (10.6).

$$\begin{aligned} F_{cap} &= F_{c,0} \frac{P_s}{P_{s,0}} \left(1 - \kappa_c (r - r_0)\right) \\ r &= \frac{P_{sat,i}(T_{cw} + \Delta T_{cd})}{P_s} \\ W &= \frac{F_c}{3600} \frac{\gamma}{\gamma - 1} \frac{R_g T}{\eta_c} \left(r^{(\gamma-1)/\gamma} - 1\right) \end{aligned} \tag{10.6}$$

where $F_{c,0} = 2730\,\mathrm{kmol/h}$ at the design suction $P_{s,0}$ and ratio $r_0 = 3.276$, $\kappa_c = 0.15$, the condenser approach $\Delta T_{cd} = 11\,\mathrm{K}$ puts its pressure at 4.643 bar with 30 °C water, $\gamma = 1.1$, $\eta_c = 0.75$, and $W = 2.509\,\mathrm{MW}$ at design. With 38 °C water the same suction pressure leaves a capacity of 2424 kmol/h. The capacity is proportional to the suction pressure because, without that term, power alone never bound and the supervisor ran the contactor to its coldest box at every price.

> **TRY IT IN THE APP 3 · Temperature by suction pressure, and a hot afternoon**
>
> Load the case and press Run. Before 2 h lower the Contactor temperature setpoint from 7 to 6.0 °C. The suction pressure falls from 1.409 to 1.362 bar a and the temperature follows within 11 min. From 2 h the cooling water warms. The compressor reaches its capacity at 2 h 17 min, and the contactor peaks at 7.10 °C at 5 h 26 min.

10.3.4 The Acid as an Inventory

The acid is three masses, sulfuric acid m_H, water m_W and ASO m_{ASO}, renewed by fresh acid and spilled over a weir, as shown in Equation (10.7).

$$\begin{aligned}
\frac{\mathrm{d}m_H}{\mathrm{d}t} &= w_f F_f - \frac{m_H}{M} F_s \\
\frac{\mathrm{d}m_{ASO}}{\mathrm{d}t} &= M_O r_{ASO} - \frac{m_{ASO}}{M} F_s \\
\frac{\mathrm{d}m_W}{\mathrm{d}t} &= (1 - w_f) F_f + \dot{m}_w - \frac{m_W}{M} F_s \\
F_s &= \max\left(0,\ \frac{M - M_{weir}}{\tau_w}\right), \qquad w = 100\,\frac{m_H}{M}
\end{aligned} \tag{10.7}$$

where $M = m_H + m_W + m_{ASO}$, $w_f = 0.985$, $M_{weir} = 78\,\mathrm{t}$ and $\tau_w = 30\,\mathrm{min}$. Both diluents lower the strength. At steady state the fresh acid must balance the dilution $D = M_O r_{ASO} + \dot{m}_w$, as shown in Equation (10.8).

$$F_f^* = \frac{w D}{w_f - w}, \qquad F_s^* = F_f^* + D \tag{10.8}$$

with w as a fraction. At design $D = 270.7\,\mathrm{kg/h}$, $F_f^* = 3831.1\,\mathrm{kg/h}$ and $F_s^* = 4101.8\,\mathrm{kg/h}$, the acid held is 80 051 kg, sulfuric acid enters and leaves at 3773.6 kg/h, and each kilogram of water costs 14.15 kg of fresh acid at 92 wt%. The fresh acid renews 4.79 % of the inventory an hour. At the design point an extra 100 kg/h of water lowers the strength 0.115 wt% an hour, a contactor 2 K warmer 0.036 and an I/O cut from 9 to 7.65 0.079, so every diluent pulls the same way.

10.3.5 Octane, the Deisobutanizer and the Instruments

The octane of the alkylate formed is a declared teaching correlation, mixed through the alkylate holdup, as shown in Equation (10.9).

$$\begin{aligned}
\mathrm{RON} &= R_0 + 3a \tanh\frac{I - 9}{3} - b(T - 7) \\
&\quad + 3c \tanh\frac{w - 92}{3} - 100\,d\,f_{C_3} \\
\tau_R \frac{\mathrm{dRON}_p}{\mathrm{d}t} &= \mathrm{RON} - \mathrm{RON}_p
\end{aligned} \tag{10.9}$$

with $R_0 = 95.0$, $a = 0.45$, $b = 0.12\,\mathrm{K^{-1}}$, $c = 0.35$, $d = 0.06$ per percent of propylene and $\tau_R = 30\,\mathrm{min}$, stated for I/O 6 to 14, 2 to 16 °C, 86 to 96 wt% and up to 60 % propylene. The directions follow the literature. The strength term, 0.35 per point near 92 wt%, lies on the 0.33 to 0.39 read from Putney's curves between 94 and 90 wt% [17], although those curves peak near 95 to 96 wt% where the tanh only saturates. The temperature term is 0.12 per kelvin against the patent's half octane number over the 11.7 K from 49 to 28 °F, 0.043 $\mathrm{K^{-1}}$ [4], about three times as steep, and the sources read give the I/O term only its sign. The coefficients were raised from 0.30, 0.10 and 0.25 because a smaller octane response let the shipped challenge pass.

The deisobutanizer is the column module the apps share, a total condenser and receiver, 24 equilibrium trays and a partial reboiler and sump at 7 bar a, with constant relative volatility, constant molar overflow, 30 kmol on each tray and a linearized weir of time constant 0.1 min. Vapor and liquid flows follow $V_i = V_{i+1} - W_i^V$ and $L_i = \bar{L}_i + (M_i - \bar{M}_i)/\tau_L$, the vapor is $y_{ij} = \alpha_j x_{ij} / \sum_k \alpha_k x_{ik}$, and each stage's temperature is its closed-form bubble point $T = B / \ln\left((P_0/P)\sum_j x_j \mathrm{e}^{B/T_{b,j}}\right)$. The net hydrocarbon and the makeup, 95 % isobutane, feed tray 8, the n-butane leaves as a 240 kmol/h vapor side draw from tray 20, the reflux is 3000 kmol/h, and the receiver holds 2000 kmol at full. At design the recycle is 5017.5 kmol/h at 89.305 % isobutane, the makeup 339.9 kmol/h, the boilup

8257.5 kmol/h (89.3 t/h of steam at 2000 MJ/t) and the bottoms 533.5 kmol/h with 1.79 % butanes. Isobutane closes over the unit as 289.09 fed plus 322.90 of makeup isobutane against 525.53 alkylated, 85.46 in the side draw and 1.00 in the bottoms. A pump draws a reduced share once its vessel holds less than 2 % of its capacity, and every flow station follows its setpoint with a lag of 0.2 min, the boilup 0.4 min.

The instruments are flow meters with noise of 0.5 to 12 kmol/h or 8 kg/h, a contactor thermometer with a 0.2 min lag, the suction pressure, the compressor's flow and its governor's limit signal, a feed moisture analyzer with a 3 min lag, a vibrating-tube acid density meter with a 1 min lag, the two DIB levels and the tray 23 temperature (0.3 min). A recycle chromatograph samples every 10 min and reports 5 min later with $\sigma = 0.001$, and the laboratory titrates a sample every 4 h and reports it 30 min later with $\sigma = 0.1$ wt%, both on the framework's analyzer module, whose delay and sampling combine as in Equation (1.12). The density, the soft sensor's input, is a teaching law, as shown in Equation (10.10).

$$\rho = \left(\frac{1 - x_{ASO}}{1814 + 4(w_{HW} - 90)} + \frac{x_{ASO}}{900} \right)^{-1} \tag{10.10}$$

with ρ in kg/m^3, x_{ASO} the ASO mass fraction and w_{HW} the strength on an ASO-free basis. At design $\rho = 1734.54\,\mathrm{kg/m^3}$ with 1.889 wt% water and 6.111 wt% ASO. Per kilogram ASO lowers the density 6.6 times as much as water does.

10.3.6 Lining Out, Integration and Balances

Table 10.2 gives each parameter and whether a learner may change it. Every case lines out at its own setpoints before $t = 0$. A pass runs from a recycle composition through the contactor to the column's Newton steady state and back to the receiver, and Newton's method on the receiver's two fractions closes the loop, because the plain fixed point contracts only by 0.87 a pass while the side draw purges the n-butane. An independent solver of the contactor and the 24-tray column reproduces the recycle, makeup, compressor flow, suction and fresh acid to four decimals. The largest relative derivative at the design point is 2.9×10^{-15}. A fixed 2 s step of fourth-order Runge–Kutta integrates the state, the controllers execute each 6 s, and the historian keeps one row per 30 s, on the engine of Section 1.6.1. Over the challenge's day the isobutane, n-butane, alkylate and olefin ledgers (kmol) and the sulfuric acid, water and ASO ledgers (kg) close to 7.4×10^{-8}, and halving the step moves the strength, temperature, I/O and octane after 8 h by 1.2×10^{-8}.

10.3.7 Assumptions

The olefin reacts completely within seconds, and only its split between two products is modeled, by a declared ratio of rates. A real contactor runs an emulsion whose interfacial area, set by the mixers and the acid-to-hydrocarbon ratio, governs both rates, and its byproducts other than ASO, light ends, heavy alkylate and esters, are left out. The contactor and settler are one hydrocarbon holdup and one acid holdup, where real units run several contactors with their own temperatures and strengths. The refrigerant returns condensed at once, with no economizer, receiver or depropanizer, and the vapor pressures follow the column's constant-volatility law, so isobutane boils at −7.80 °C at one atmosphere instead of −11.7 °C. The column is 24 theoretical trays at a fixed pressure, its condenser and reboiler duties are not modeled beyond the boilup and its steam limit, and its tray count is a design choice that sets how far the recycle drifts. The strength is the sulfuric acid mass fraction, where a titration counts total acidity. The selectivity, activity, octane and density laws, the floor at 87 wt%, the STRENGTH LOW alarm at 89 wt% and the low-I/O trip at an I/O of 5 for 2 min are teaching values. A real minimum strength comes from the licensor's operating guidelines and the refiner's procedures, and a low-isobutane interlock and its settings from the unit's safety review.

Table 10.2 Parameters. A: Adjustable in the Model Dialog. H: Hidden and Fixed

Symbol	Meaning	Value	
$F_{O,0}$	design olefin	530 kmol/h	A
z_O, z_I	olefin, isobutane shares	0.55, 0.30	H
μ_I, $F_{MU,max}$	makeup purity, capacity	0.95, 700 kmol/h	H, A
N_R, C_R	contactor holdup, heat capacity	1500 kmol, 300 MJ/K	A
$-\Delta H_a$, $-\Delta H_s$	alkylation, polymerization heat	80, 90 MJ/kmol	A, H
Q_{mix}, T_f, c_p	mixers, feed, heat capacity	1.5 MW, 10 °C, 0.13	H
k_v	boiling coefficient	28 000 kmol/(h bar)	A
κ_0, θ, n, c_3	selectivity law	0.0085, 0.06, 1.5, 1.5	A
β, A, w_r, s	activity law	0.06, 0.5, 87, 0.6	A
R_0, a, b, c, d	octane law	95.0, 0.45, 0.12, 0.35, 0.06	A
τ_R	alkylate holdup	30 min	H
τ_c, $F_{c,0}$, κ_c	governor, capacity, loss	0.5 min, 2730, 0.15	A
ΔT_{cd}, η_c	condenser approach, efficiency	11 K, 0.75	H
M_{weir}, τ_w	acid to the weir, weir lag	78 t, 30 min	A, H
w_f	fresh acid	98.5 wt%	A
w_{floor}, w_{low}	runaway floor, STRENGTH LOW	87, 89 wt%	A
$\dot{m}_{w,0}$	design feed water	20 kg/h	H
P_{DIB}, α_I, α_A	pressure, volatilities	7 bar, 1.3, 0.1	H, A, H
B, $T_{b,n}$	latent heat over R_g, n-C_4	2600 K, 272.65 K	H
M_t, τ_L	tray holdup, hydraulic lag	30 kmol, 0.1 min	A, H
$M_{D,cap}$, $M_{B,cap}$	receiver, sump capacity	2000, 400 kmol	A, H
V_{max}	reboiler (steam) limit	10 500 kmol/h	A
$\bar{L}$, $\bar{S}$, B_l	design reflux, side draw, butanes	3000, 240, 8 kmol/h	H
τ_Q	flow loop lag	0.2 min	H

10.4 Three Clocks, a Floor and a Trip

10.4.1 Minutes, Hours and More Than a Day

The three clocks were measured in the app, loops and instruments included, as shown in Figure 10.2. With the temperature loop in Manual, a cut of 0.05 bar in the suction pressure takes the contactor from 6.984 to 5.994 °C, 63 % of the way in 2.07 min. A linearization of Equation (10.4) with $\mathrm{d}P_{bub}/\mathrm{d}T = 0.0497\,\mathrm{bar/K}$ gives 0.86 min, and the measured response is slower because the step asks the compressor for more than its capacity during the first minute. In case 2, with the ratio held on the meters and the design purity, 15 % more olefin drifts the recycle's purity from 89.305 to 86.094 % a day later, 63 % of the way in 7 h 00 min. The extra n-butane leaves the loop only through the fixed side draw, and the column, receiver and sump hold about 1900 kmol beside the contactor's 1500 kmol, which turn over in about 35 min at 5800 kmol/h. The purity is the column's to set, not the loop's.

> **TRY IT IN THE APP** **2 · A ratio on flows and the recycle's purity**
>
> Load the case and press Run. Touch nothing. At 30 min the olefin rises 15 %. Watch the I/O trend. The loop's value stays at 9.0 while the analyzer's value, the dashed line, sinks below 8.9 at 2 h 56 min. Set Recycle purity from to Analyzer. The loop raises the recycle from 5770 to 5842 kmol/h and the analyzer's I/O returns to within 0.1 of 9.

On Fixed flows, 20 % less fresh acid takes the strength from 92.000 to 89.608 wt% over four days, 63 % of the way in 36.3 h. A linearization of Equation (10.7) at fixed fresh acid gives 29.4 h, with the ASO rising 14.96 kg/h per point lost. The measured response is slower because the activity's slope steepens as the strength falls, so each point lost adds more ASO and weakens the restoring rate. A loop on the acid therefore acts on a plant whose time constant is longer than a day, through a measurement that arrives every four hours.

Pause. Try Exercise 10.3 before reading on.

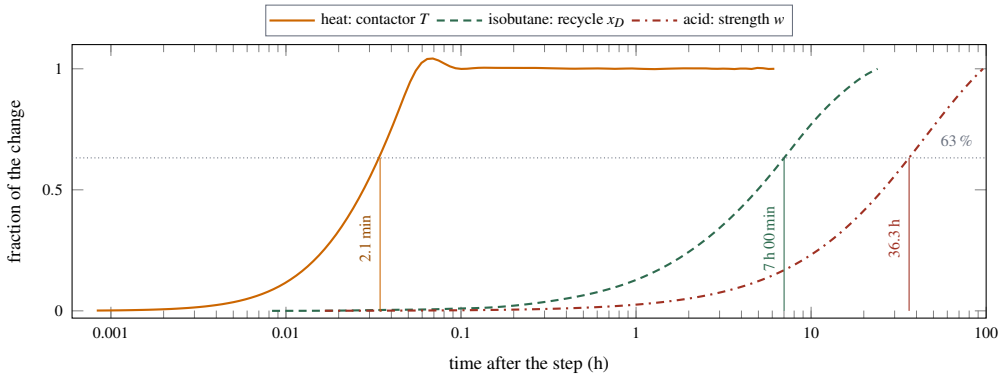


Figure 10.2 The three clocks, from the app's engine, each response divided by its own change. Heat: the contactor after a 0.05 bar suction cut with the temperature loop in Manual (it overshoots by 4 %). Isobutane: case 2 as shipped, the recycle's purity after 15 % more olefin, against its value a day later. Acid: Fixed flows, the strength after 20 % less fresh acid, over 96 h.

10.4.2 An Activity That Collapses

The floor is chemical. Below about 89 wt% the activity of Equation (10.2) collapses, each mole of olefin makes more ASO, and the ASO dilutes the acid further. In case 5 the unit lines out at 90 wt% on 3211 kg/h of fresh acid, and the acid is withheld from 5 min. The strength passes 89 wt% at 2 h 53 min and 88 wt% at 5 h 28 min, falling 0.35 points in the first hour and 0.52 in the last, the ASO grows from 283 to 504 kg/h, and the ACID RUNAWAY stop comes at 7 h 35 min. The stop latches on the first step below 87 wt% and never above it. Restoring the design point's 3831 kg/h at 87.6 wt% turns the strength back without a stop, while the case's own 3211 kg/h recovers from 88.0 wt% but not from 87.6.

With the strength loop in Auto the floor is out of reach of any single disturbance slider for a day. Feed water at its 1000 kg/h maximum takes the strength to 88.52 wt% at 5 h 14 min with no stop. The floor comes only when the fresh acid falls short of what dilutes the acid. On Fixed flows, with the design acid held, 130 % olefin kept from 30 min runs the acid away at 14 h 03 min, feed water raised from 20 to 420 kg/h at 13 h 37 min, and a recycle held at 3000 kmol/h, an I/O of 5.6, at 18 h 14 min. On Loops with feedforward the acid runs away at 5 h 48 min only with the olefin, water, propylene and cooling water sliders all at their limits and the contactor at 14 °C, the fresh acid pinned at its 12 t/h maximum. A contactor at 14 °C, cooling water at 42 °C and 60 % propylene never reach the floor even with the acid held, ending the day at 89.34, 91.89 and 88.80 wt%.

10.4.3 A Trip on the Isobutane

A unit short of isobutane polymerizes its olefin. The low-I/O interlock stops the run once the true I/O has been below 5 for 2 min, a teaching value. It was added because a lost makeup otherwise ran the I/O down into a second, isobutane-driven path to the acid floor, which the app's design review asked to exclude. In case 7 the makeup supply is lost at 30 min, the receiver loop asks for 689 kmol/h by 2 h, the receiver falls below 30 % at 1 h 42 min and below 10 % at 2 h 51 min, the recycle pump loses suction, and the trip comes at 4 h 00 min with the I/O at 4.22. Cutting the olefin to 60 % at 1 h 42 min delays it to 4 h 23 min, and to 40 % only to 4 h 39 min. The feed's isobutane, 30 % of the stream, is less than its olefin, 55 %, consumes, and the side draw keeps taking isobutane, so no olefin rate or I/O setpoint in the app's sweep holds the ratio, every run tripping between 4.00 and 4.93 h.

Table 10.3 Loops, Pairings and Default Tunings

Loop	PV → MV	K_c	τ_I	Output
IIC	$I \to R$ (over FF)	300 per unit	30 min	0–9000 kmol/h
TIC	$T \to P_{s,SP}$	0.03 bar/K	3 min	0.6–2.5 bar a
AIC	$w \to F_f$ (over FF)	3000 per point	120 min	0–12 000 kg/h
LIC	$M_D \to F_{MU}$	8 kmol/h per %	60 min	0–1200 kmol/h
TIC	$T_{23} \to V$ (over L+R+S)	100 kmol/h K	20 min	0–10 500 kmol/h
LIC	$M_B \to B$, reverse	20 kmol/h per %	—	0–1500 kmol/h

10.5 Loops, a Soft Sensor and a Supervisor

Six loops act on the unit, with the header's Manual and Auto belonging to the I/O loop alone. The temperature, strength and three column loops carry their own Auto and Manual switches, because case 1 needs the strength loop live while the recycle is in the learner's hands. A control structure option sets how the quality loops act: Fixed flows holds the recycle and the fresh acid at their stations with only the temperature loop acting, Loops puts a PI controller on each measurement, Loops + FF adds ratio feedforwards from the olefin meter, Supervisor takes the I/O, temperature and strength setpoints from the economic optimization, and MPC replaces the three quality loops with the predictive controller of Section 10.5.5. Under the loops every manipulated flow is a feedforward plus a PI trim, and in Manual the trim tracks the station minus the feedforward, so every switch is bumpless. A single function builds the measurement packet the controllers read, and an application test checks that it carries measurements only.

10.5.1 The Shared PI Law

The loops share the framework's velocity-form controller of Section 1.6.2, Equations (1.16) to (1.19), with no derivative action in any loop here, a control period $\Delta t = 0.1$ min and the tracking time T_t equal to τ_I. Only the emitted output is limited, and back-calculation drags the integral back toward what the clamp let through. At the largest value its slider allows, τ_I removes integral action. Table 10.3 gives each pairing with its default tuning. The bottoms loop is proportional and reverse acting. The temperature loop's output is the suction pressure setpoint, and when the governor reports that the compressor is at its capacity the loop's lower limit closes onto the measured suction pressure less 0.005 bar, so its integral does not wind down against a pressure that is not coming.

10.5.2 Ratio Feedforwards and the Recycle Analyzer

The I/O the loop sees is computed from the meters and a purity π, and the feedforwards supply the recycle and the fresh acid that a ratio needs, as shown in Equation (10.11).

$$
\begin{aligned}
I_m &= \frac{R_m\,\pi + z_I F_{OF,m}}{z_O F_{OF,m}} \\
R_{ff} &= \frac{(I_{SP} z_O - z_I) F_{OF,m}}{\pi} \\
F_{f,ff} &= \frac{w_{SP}(M_{C_4} r_0 + \dot{m}_{w,0})}{w_f - w_{SP}} + k_w\,\Delta\dot{m}_w \\
r_0 &= \frac{\rho_0}{1+\rho_0}\, z_O F_{OF,m}
\end{aligned}
\tag{10.11}
$$

where π is the chromatograph's last report or the design 89.3 %, ρ_0 is the design ASO ratio, so r_0 is the ASO the metered olefin makes at design, $M_{C_4} = 56.11$ kg/kmol, $k_w = w_{SP}/(w_f - w_{SP})$ is the acid per kilogram of water, $\Delta\dot{m}_w = \dot{m}_{w,m} - \dot{m}_{w,0}$ is the measured water above design, and the last term, the moisture feedforward, acts on the feed analyzer when it is switched on. The ratio feedforward and the strength term follow the olefin meter, and the PI controllers trim them.

10.5.3 The Soft Sensor and the Laboratory

The density soft sensor solves Equation (10.10) for the strength by bisection, assuming the diluent is water and ASO in their design shares, 23.6 % water by mass, and adds a bias, as shown in Equation (10.12).

$$\hat{w} = w(\rho_m;\ \varphi_0) + b, \qquad b \leftarrow b + \lambda\left(w_{lab} - \hat{w}_{raw}(t_s) - b\right) \tag{10.12}$$

where $\lambda = 0.5$, the residual is taken against the raw value at the instant t_s the sample was drawn, and the bias is bounded at 3 wt%. This is the laboratory bias update of Equation (1.14), and Volume 5, Chapter 10 corrects a psychrometric soft sensor from a sampled, delayed laboratory in the same way. Because ASO lowers the density 6.6 times as much as water per kilogram, the soft sensor reads a water dilution at 19 % of its size and an ASO dilution at 125 %. The bias starts where a lined-out unit's titrations would have left it, at the case's own strength setpoint less the raw reading, a choice made after a zero bias left case 5 adding acid to a unit its briefing calls lined out. With the soft sensor Off the strength loop acts once per laboratory report, over the time since the previous one, and holds between reports. The STRENGTH LOW alarm reads the soft sensor either way, because an alarm on the titration would be four hours late.

Pause. Try Exercise 10.2 before reading on.

10.5.4 The Economic Supervisor

Every 30 min the supervisor chooses the I/O, temperature and strength targets on the app's own steady-state correlations, fed with the measured olefin, water, propylene, cooling water and recycle purity, as shown in Equation (10.13).

$$\begin{aligned} &\max_{I,T,w}\ p_R\,\mathrm{RON}\,\dot{m}_{alk} - p_a F_f^{*} - p_e W - p_s \dot{m}_{st} \\ &\text{s.t. } F_c \le 0.97 F_{cap}, \qquad V \le 0.97 V_{max} \\ &\qquad 7 \le I \le 12, \quad 4 \le T \le 12\,^{\circ}\mathrm{C}, \quad 89.5 \le w \le 95\,\mathrm{wt\%} \end{aligned} \tag{10.13}$$

where $p_R = \$4$ per RON and tonne, $p_a = \$150$ per tonne of acid, $p_e = \$70$ per MWh, $p_s = \$20$ per tonne of steam, $V = L + R + S$ and $\dot{m}_{st} = \lambda V / 2000\,\mathrm{MJ/t}$. Colder always pays in this economics, so the temperature is the coldest the compressor holds, an optimum on a limit as Section 1.4.7 explains, found by bisection on its capacity row. The I/O and the strength are found by successive linear programming, with the margin and the reboiler row linearized by central differences, the framework's bounded simplex solving each linear program inside a trust region, and the longest of 1, 1/2, 1/4 or 1/8 of the step that improves the margin taken. A start beyond the reboiler row, where a feed increase leaves the previous target, is brought back inside it by the same row linearized there with an elastic slack priced a thousand times above the margin's own rate, a step being taken when it lowers the excess. The problem is reduced to two variables because a three-variable version stalled on the active compressor row. The targets move at most 1 in I/O, 1 K and 0.5 wt% a period. Against a grid of 0.05 over the boxes the reduced problem is \$0.43, \$1.52 and \$0.66 an hour better at 80, 150 and \$250 per tonne, its temperature sits on the compressor row at 97.00 % of capacity, and 0.2 K warmer costs \$8.5 to \$12.0 an hour. It is a real-time optimization in the sense of [11], a steady-state optimization that runs, not a dynamic one.

10.5.5 Model Predictive Control

The fifth structure gives the three quality handles to one model predictive controller (MPC) of the kind described in Section 1.6.3, and leaves the column loops, the supervisor and every protection acting as they do under Supervisor. Every $\Delta t_M = 2$ min, twenty scans of the loops, it solves a quadratic program over $N_p = 120$ steps, a horizon of 4 h, for the I/O request u_1, the suction pressure setpoint u_2 and the fresh acid u_3, applies the first move and holds it to the next decision. The recycle is not itself

a handle. At every scan the recycle station sets $R = (u_1 z_O - z_I) F_{OF,m}/\pi$, the ratio of Equation (10.11), so an olefin step reaches the recycle within 6 s instead of at the next decision. The setpoints I_{SP}, T_{SP} and w_{SP} it tracks are the targets of Equation (10.13), which the supervisor keeps choosing beneath it every 30 min, and the price of the fresh acid enters only there.

The controller's model has five states, as shown in Equation (10.14).

$$\begin{aligned} \tau_i \frac{\mathrm{d}\hat{z}_i}{\mathrm{d}t} &= K_i v_i - \hat{z}_i, \qquad i = 1,2,3 \\ \hat{T} &= T_{ref} + \hat{z}_1 + \hat{z}_2 + \hat{z}_3 + b_T \\ \widehat{I} &= u_1 + b_I, \qquad \tilde{w} = \hat{w} + b_w \\ \frac{\mathrm{d}\hat{w}}{\mathrm{d}t} &= \frac{(w_f - \tilde{w}) F_f - \tilde{w}\,(\dot{m}_{w,m} + M_O \hat{r}_{ASO})}{60\,\hat{M}} \\ \hat{M} &= M_{weir} + F_{s,m}\tau_w/60 \end{aligned} \tag{10.14}$$

where a hat marks a quantity of the controller's model, $v_1 = u_2$, $v_2 = R$, v_3 is the metered olefin feed in percent of design, T_{ref} places $\hat{T}$ on the measured temperature at the start, $\tilde{w}$ is the strength the MPC controls, the strengths and w_f are in wt%, $\dot{m}_{w,m}$ and $F_{s,m}$ are the moisture analyzer and the spent acid meter, and F_f is the fresh acid meter in the estimate and u_3 in the prediction. The ASO rate $\hat{r}_{ASO}$ is that of Equation (10.1) at the declared parameters, evaluated at the metered olefin, the measured I/O and contactor temperature, $\tilde{w}$ and the propylene share f_{C_3}, which a feed chromatograph reports as an ideal instrument, without lag, noise or sampling.

The three temperature paths are first-order-plus-dead-time fits to steps up and down at the lined-out design point, each dead time added to its lag. The suction pressure gives $K_1 = 19.63\,\mathrm{K/bar}$ with $\tau_1 = 1.1\,\mathrm{min}$, the recycle $K_2 = -0.000\,552\,\mathrm{K}$ per kmol/h with $\tau_2 = 18.2\,\mathrm{min}$, the contactor's composition time N_R/F_{out} of 16.5 min plus its dead time, and the olefin, with the I/O held by the ratio, $K_3 = 0.0155\,\mathrm{K}$ per percent with $\tau_3 = 0.75\,\mathrm{min}$. An independent refit with the noise off gives 19.58 and 19.80 K/bar upward and downward, −0.000528 and −0.000584, and 0.01477 and 0.01499, with residuals of 0.3 to 3.2 % of the step. Through the station one unit of I/O cools the contactor by 0.33 K at design.

The strength row is not identified. It is the plant's own acid balance, Equation (10.7) written for the strength with the weir law inverted for $\hat{M}$, and its declared ASO law, with the declared parameters frozen when the application loads and fed by instruments only, so it is not an independent model. At the shipped parameters, integrated for a day after a step in the fresh acid, the I/O, the olefin, the feed water or the propylene, with the strength loop open, it stays within 0.0056 wt% of the true strength, a difference set by the instruments. With the plant's κ_0 raised 29 % the same steps leave it 0.99 to 1.81 wt% away. At each decision the row is linearized by central differences about $\tilde{w}$ and the applied commands, the temperature entering it through its predicted paths, and the five states are discretized with an exact zero-order hold over Δt_M.

The model runs at every 6 s scan under whichever structure acts, and the MPC therefore takes over from an estimate that is already current. The biases b_T and b_I are the measured less the modeled temperature and I/O, filtered over 0.5 and 1 min. The strength bias takes half of each titration's residual, $b_w \leftarrow b_w + 0.5\big(w_{lab} - \hat{w}(t_s) - b_w\big)$, bounded at 3 wt% and taken against the model at the instant t_s the sample was drawn, as Equation (10.12) does for the soft sensor. With the soft sensor On, b_w follows the corrected soft sensor through a 30 min filter instead. Such biases are the integrating disturbances through which predictive control absorbs plant-model mismatch [22], and Volume 9, Chapter 4 builds an offset-free predictive controller whose estimator carries a fuel bias of this kind. Here b_w corrects the output alone while $\hat{w}$ keeps integrating its own rate between titrations, so a rate the model gets wrong returns at every report as a fresh offset, and the MPC is not offset-free when the plant's acid law departs from the declared one.

Each decision minimizes the objective of Equation (10.15), a case of Equation (1.21) with a funnel on the floor band, over the three handles held in twelve move blocks, which lengthen from a single

step at the front of the horizon to 75 steps at its end, 36 decisions and one slack ε.

$$\begin{aligned}\min_{u,\varepsilon}\ &\sum_{k=1}^{N_p}\Big[100\big(\widehat{I}_k - I_{SP}\big)^2 + 100\big(\hat{T}_k - r_T\big)^2 \\ &\qquad + 400\big(\tilde{w}_k - r_w\big)^2\Big] + 10^5\varepsilon^2 + 10^5\varepsilon \\ &+ \Lambda\sum_{j=1}^{12}\Big[100\,\Delta u_{1,j}^2 + \Big(\frac{\Delta u_{2,j}}{0.01}\Big)^2 + 10\Big(\frac{\Delta u_{3,j}}{1000}\Big)^2\Big] \\ \text{s.t. } & u_{i,min} \le u_{i,j} \le u_{i,max}, \qquad \varepsilon \ge 0 \\ & \tilde{w}_k \ge w_b - \max(0,\, w_b - \tilde{w}_0)\,\mathrm{e}^{-k\Delta t_M/\tau_f} - \varepsilon\end{aligned} \tag{10.15}$$

where the tracking weights are the inverse squares of 0.1 in I/O, 0.1 K and 0.05 wt%, $\Delta u_{i,j}$ is the change of handle i at block j, the first of them taken from the command in force, u_2 is in bar and u_3 in kg/h, and $\Lambda = 10$ is the move suppression, chosen once on case 9 and never retuned. The floor band w_b is the run's runaway floor, read as the protection reads it, plus 2 wt%, 89 wt% at the default, $\tilde{w}_0$ is the estimate now, and $\tau_f = 60$ min is a funnel by which an estimate already below the band is brought back. The band is soft and priced far above tracking, with one slack for the whole horizon. The strength target is $r_w = \max(w_{SP}, w_b)$, so the target never pulls the plan under the band. While the governor reports the compressor at its capacity the temperature target is $r_T = \max(T_{SP}, \hat{T}_0)$, the temperature loop's own rule, and $r_T = T_{SP}$ otherwise. Without that rule the plan reached for a temperature the compressor could not give through the recycle's cooling, which in the challenge raised the I/O request by 0.461 in all while it stood above its target and the butanes in the alkylate to 32.3 %, against 0.030 and 16.3 % with it.

The bounds on the handles are hard and are those the shipped loops clamp to, the I/O request between the ratios the station gives at 0 and 9000 kmol/h of recycle, the suction pressure between 0.6 and 2.5 bar a with its lower bound closing onto the held pressure less 0.005 bar while the compressor is at its capacity, and the fresh acid between 0 and 12 000 kg/h. A handle whose own loop is in Manual is pinned at its station, so the plan works around a recycle held by hand. The framework's dense quadratic program solver returns each plan, and a plan is kept only if the solver's status is optimal, the first move, the plan, the predictions, the cost and the slack are finite, the slack is nonnegative and every move lies inside its bounds. A plan that fails is rejected, the historian records MPC FALLBACK, and the Supervisor structure, whose PI controllers track the applied commands at every scan without preload, holds the unit until the next valid plan takes back. With the noise off, a rejection forced on a quiet scan of case 9 moves no handle, and one on the scan a titration arrives moves the fresh acid by 867.3 kg/h, the same move the shipped structure makes when its strength loop goes from Manual to Automatic on that scan, its proportional response to the new titration. No plan was rejected in any run reported here (718 to 720 a day). The runaway stop, the low-I/O trip and the alarms act above the MPC on the run's own limits. A plan takes 0.78 to 0.95 ms in Node and 0.75 to 0.83 ms in Chrome, 0.04 % of one core at 1 min/s and 11 to 12 % at the top speed of 300 min/s, which the application still holds at 59 to 60 frames a second.

10.5.6 Three Models Kept Apart

Section 10.3 is the plant the engine integrates. A learner who steps the recycle and fits its gain on the I/O builds an identified model of it. What the loops and the supervisor assume is narrower still: PI gains, the feedforwards' design ASO ratio and design purity, the soft sensor's design diluent split, and the supervisor's copy of the correlations with the scenario's parameters, which sees an upset only through the instruments. The MPC's model of Equation (10.14) is identified in its temperature row, while its strength row is the plant's own acid law at the declared parameters read through the instruments, exact on the shipped plant but for those instruments and wrong in its rate on a changed one. The status strip, the verdict and the protection layer read the truth and act on no controller.

Table 10.4 The Nine Cases and the Challenge

#	Case and events	Teaching point
1	Manual recycle, olefin 115 % at 10 min	I/O, octane, acid use
2	Ratio on the design purity, olefin 115 % at 30 min	the recycle's purity
3	105 % olefin, learner sets 6 °C, water 38 °C from 2 h	refrigeration limit
4	Manual recycle, olefin 115 % at 1 h, titration alone	sampled strength
5	Lined out at 90 wt %, learner withholds acid	the runaway floor
6	30 % propylene at 30 min	a poorer olefin
7	Makeup lost at 30 min	an infeasible ratio
8	Supervisor at $80 then $250/t acid	the economic optimum
9	MPC against Supervisor, titration alone, 30 % propylene at 1 h	a model between titrations
10	Challenge, 24 h: olefin 115 % at 2 h, slug at 8 h, 38 °C 12–18 h	the whole unit

Table 10.5 Case Outcomes Measured on the Engine

Case and condition	Result
1, as shipped	I/O 9.0 → 7.9; ASO 251 → 353 kg/h; acid use 73.4 (2 h), 82.0 kg/t (6 h); product RON 94.69 (40 min), 94.45 (6 h)
1, recycle 5775 at 20 min	RON 94.94, acid use 65.0 kg/t at 6 h
2, design purity	purity 88.1, 87.6, 86.7, 86.1 % (4, 6, 12, 24 h); loop I/O 9.0, true 8.88 (4 h), 8.75 (12 h)
2, analyzer at 2 h 56	recycle 5770 → 5842 kmol/h; within 0.1 of 9 0.6 min later
3, SP 6 °C	suction 1.409 → 1.362 bar a; within 0.15 K in 11 min
3, afternoon	at capacity 2 h 17; +0.3 K at 2 h 37; peak 7.10 °C at 5 h 26; back 6 h 39
4, titration alone	strength 91.45–92.31 wt %; fresh acid 3205–6472 kg/h
4, soft sensor	strength 91.92–92.03 wt %
5, acid withheld	89 at 2 h 53, 88 at 5 h 28; ACID RUNAWAY 7 h 35; ASO 283 → 504 kg/h
6, as shipped	inferred RON 95.00 → 93.18; ASO 250 → 335 kg/h; acid use 64.1 → 89.1 kg/t (4 h)
6, I/O 11	RON 94.00 (4 h); ASO 249 kg/h; acid use 66.7 (4 h), 64.4 kg/t (12 h); steam 89.4 → 101.9 t/h
7, as shipped	receiver < 30 % 1 h 42, < 10 % 2 h 51; LOW I/O TRIP 4 h 00 (I/O 4.22)
7, olefin 60 / 40 %	trip 4 h 23 / 4 h 39
8, $80/t	targets I/O 10.90, 4.93 °C, 94.25 wt % at 4 h
8, $250/t	strength target < 93 at 5 h 06; 12.00, 4.87, 92.79 at 12 h; true 92.70 at 24 h
9, MPC / Supervisor	strength from 4 h 93.12–93.23 / 92.68–93.78 wt %; IAE 1.20 / 6.18 wt % h; fresh acid 79.8 / 89.0 t

10.6 Nine Cases and a Day

The cases are listed in Table 10.4 and their outcomes in Table 10.5, and Section 1.6.4 describes how scenarios, completion rules and the challenge work. Each value comes from the engine stepped headless on the default seed, as the browser steps it, with the case as published and the learner's moves made at the instants its briefing names. A measuring script re-run on the app reproduces all 122 values the app's texts quote, and runs at seeds 1 and 2 and with a halved step keep every lesson.

10.6.1 A Ratio on Flows and a Column's Purity

Case 1 raises the olefin 15 % at 10 min against a recycle the learner holds. The I/O falls from 9.0 to 7.9 at once, and with less isobutane around it a larger share of the olefin polymerizes, so the ASO rises from 251 to 353 kg/h, 41 % for 15 % more olefin. The strength loop answers with acid, and the acid use climbs from about 64 to 73.4 kg/t two hours later and 82.0 kg/t at 6 h, while the product octane slips from 95.00 to 94.69 within 40 min. A recycle raised to about 5775 kmol/h at 20 min holds the I/O near 9 and leaves 94.94 and 65.0 kg/t at 6 h. In case 2 the feedforward moves the recycle from 5018 to 5774 kmol/h at the step, and the loop reads 9.0 for the rest of the day while the purity drifts. The analyzer's I/O falls below 8.9 at 2 h 56 min, the true I/O at 3 h 13 min, and switched to the chromatograph then, the loop raises the recycle from 5770 to 5842 kmol/h in 30 min and the analyzer's I/O is back within 0.1 of 9 0.6 min later.

🕮 **Pause.** Try Exercise 10.1 before reading on.

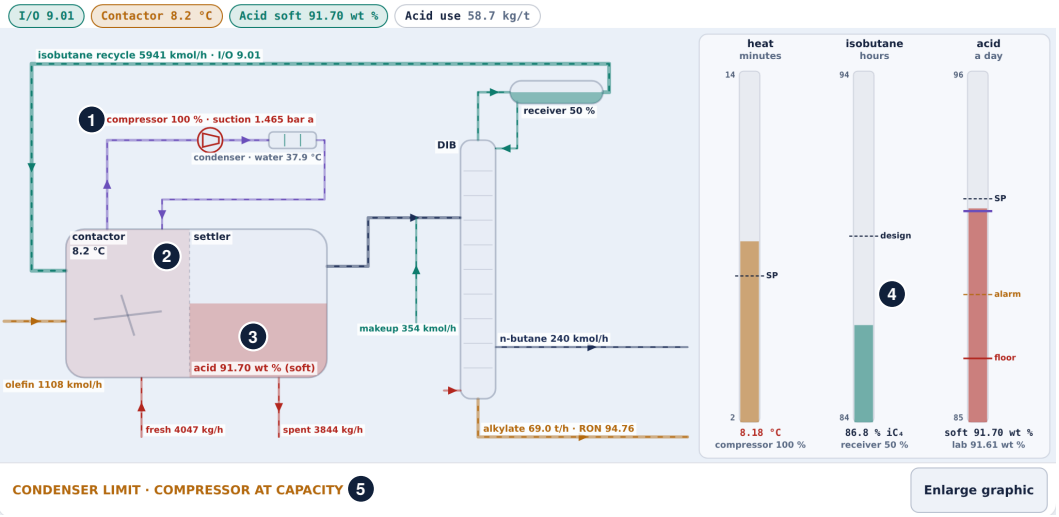


Figure 10.3 The challenge on its suggested settings at $t = 13$ h, Unit view, light theme, an hour into the hot afternoon, cropped to the plant view with numbered callouts. The badges read an I/O of 9.01, the contactor at 8.2 °C (2), the acid at 91.70 wt% on the soft sensor (3) and an acid use of 58.7 kg/t. The compressor is at 100 % of its capacity holding 1.465 bar a with the condenser water at 37.9 °C (1), the olefin is 1108 kmol/h, the recycle 5941 kmol/h on its meter, the makeup 354, the fresh and spent acid 4047 and 3844 kg/h, and the alkylate 69.0 t/h at an inferred RON of 94.76. The three gauges (4) show the heat at 8.18 °C above its setpoint, the recycle at 86.8 % isobutane below its 89.3 % design mark and the acid at 91.70 against its last titration of 91.61, all three inventories off their marks at once, under CONDENSER LIMIT and COMPRESSOR AT CAPACITY (5). The true strength is 91.46 wt%.

10.6.2 Refrigeration Until It Runs Out

In case 3, at 105 % olefin, a setpoint lowered from 7 to 6 °C takes the suction pressure from 1.409 to 1.362 bar a and the temperature within 0.15 K of it 11 min later, with the compressor carrying 2452 of its 2570 kmol/h. As the cooling water warms to 38 °C from 2 h, the compressor reaches its capacity at 2 h 17 min, the contactor is 0.3 K above its setpoint at 2 h 37 min, and it peaks at 7.10 °C at 5 h 26 min while the machine holds the suction pressure it can, 1.416 bar a at 4 h. When the water cools the temperature is back within 0.15 K at 6 h 39 min.

10.6.3 An Acid on a Four-Hour Laboratory

Case 4 raises the olefin at 1 h with the recycle fixed, so the ASO rises from 251 to 354 kg/h. On the titration alone, with the default gain of 3000 kg/h per point and 120 min of integral time, the strength swings between 91.45 and 92.31 wt% over the day and the fresh acid between 3205 and 6472 kg/h, because each report moves the loop by a whole period of integral action on an error that is already hours old. On the soft sensor, with the same tuning, the strength stays between 91.92 and 92.03 wt%. The soft sensor's value is in the variance of the strength, which a minimum on a single day does not see.

10.6.4 A Poorer Olefin, a Lost Makeup and Two Prices

In case 6 30 % propylene takes the inferred octane from 95.00 to 93.18 within the hour and the acid use from 64.1 to 89.1 kg/t by 4 h. Raising the I/O to 11 buys about half of the octane back, 94.00 at 4 h, because the tanh term flattens above 9 while the propylene penalty does not depend on the I/O, and it costs the column 12.5 t/h more steam. Case 7 is honestly infeasible, as Section 10.4 measured.

Table 10.6 Shipped Structures Against the MPC, Default Seed, 24 h. IAE Is That of the True Strength Against the Target in Force in Each Run, Travel the Sum of the Fresh Acid's Moves, and Margin the Supervisor's Objective Evaluated on the Simulator's Truth

Case	Structure	IAE (wt % h)	Largest error (wt %)	Fresh acid (t)	Acid and power ($)	Margin ($)	Acid travel (kg/h)	Plans, rejected
9, titration alone, propylene at 1 h	Supervisor (the comparison)	6.182	0.551	88.96	17 961	459 372	14 198	—
	MPC as loaded	1.202	0.329	79.78	16 583	460 939	18 821	720, 0
	Supervisor, hindsight K_c 6000, τ_I 480 min	4.256	0.351	79.24	16 503	460 875	14 706	—
	Supervisor on the soft sensor	1.483	0.298	79.14	16 488	461 011	108 069	—
	the same, hindsight 12 000, 240 min	1.007	0.305	79.60	16 556	460 955	419 501	—
	Supervisor's PI on the MPC's estimate, no plan	2.026	0.296	78.88	16 449	461 000	12 983	—
	the same, hindsight 12 000, 30 min	1.259	0.327	79.84	16 592	460 905	29 326	—
	Supervisor with a propylene feedforward	13.407	1.378	89.22	18 001	458 987	27 133	—
	MPC without the propylene	6.835	0.458	71.71	15 373	461 672	27 205	720, 0
	MPC, propylene sampled every 10 min	1.184	0.329	79.76	16 581	460 939	18 876	720, 0
	Loops + FF on 11.4, 4.9 °C, 93.5 wt %	2.399	0.223	89.75	18 071	460 484	12 125	—
	MPC on the same fixed setpoints	0.771	0.078	92.34	18 461	460 098	5 626	720, 0
6, propylene at 30 min	Loops + FF, soft sensor (shipped)	1.195	0.113	118.58	22 003	451 321	107 537	—
	the same, hindsight 12 000, 240 min	0.672	0.065	118.99	22 065	451 298	429 600	—
	MPC on the shipped setpoints	0.754	0.057	119.28	22 108	451 291	5 484	720, 0
	Supervisor	5.458	1.084	99.94	19 605	456 819	110 007	—
	MPC	3.359	0.823	99.62	19 556	456 794	27 166	720, 0
	Loops + FF on the model's optimum	6.245	1.200	99.51	19 542	456 861	105 583	—
4, recycle by hand, olefin at 1 h	Loops + FF, titration (shipped)	3.853	0.549	122.31	23 188	555 277	12 031	—
	the same, hindsight 12 000, 240 min	2.340	0.347	132.79	24 760	553 589	12 421	—
	MPC on the shipped setpoints	0.751	0.078	128.14	24 061	554 403	5 980	720, 0
	Supervisor, titration	20.577	1.786	149.52	27 285	554 215	31 086	—
	MPC	2.168	1.022	165.86	29 736	551 800	21 641	720, 0
8, acid $80, $250/t	Supervisor (the case's path)	12.336	1.455	76.17	16 037	490 925	93 556	—
	MPC on the same path	6.720	1.175	75.77	15 975	490 985	35 551	718, 0

Case 8 switches the supervisor on at $80 a tonne of acid. Within four hours it settles on an I/O of 10.90, 4.93 °C and 94.25 wt%, cheap acid buying octane. At $250 its strength target is below 93 wt% at 5 h 06 min, and by 12 h it reads an I/O of 12.00, its upper box, 4.87 °C and 92.79 wt%, with the compressor at 97 % of its capacity. Dear acid is used weaker, and more isobutane makes less ASO for it to absorb. The strength itself follows over a day, to 92.70 wt% at 24 h. With the acid above its new target the strength loop stops the fresh acid altogether from 6 h 46 min to 12 h 57 min, using the strength already paid for in the settler, and the strength's lowest after the change is 92.44 wt% at 17 h 10 min, more than three points above the alarm. The stop is physical and intended, and the case's explanation says so.

10.6.5 A Predictive Controller Between Titrations

Case 9 starts lined out where the supervisor places the unit, an I/O of 11.4, 4.9 °C and 93.5 wt%, with the strength on the titration alone, and at 1 h turns 30 % of the olefin to propylene. The ASO rises from 141 to about 182 kg/h, and the supervisor moves its strength target to 93.17 to 93.23 wt%. The case loads under the MPC and is compared with the same day under the Supervisor structure after Reset, whose targets differ from the MPC run's by at most 0.0001 wt%. Table 10.6 measures cases 9, 6, 4 and 8 under the shipped structures and under the MPC with the settings case 9 ships, on the default seed over 24 h with the same instruments, noise and events in every row. Beside them stand the stronger classical settings, the best of 20 strength tunings (K_c of 1000 to 12 000 per point, τ_I of 30 to 480 min) chosen in hindsight, the classical loop given the signal the MPC's model reads, the MPC given the loops' fixed setpoints, and the MPC without the propylene.

Under the MPC the true strength stays within 93.12 to 93.23 wt% from 4 h on, its largest error in each later four-hour period is 0.042 to 0.097 wt%, and the six titrations differ from the model at their

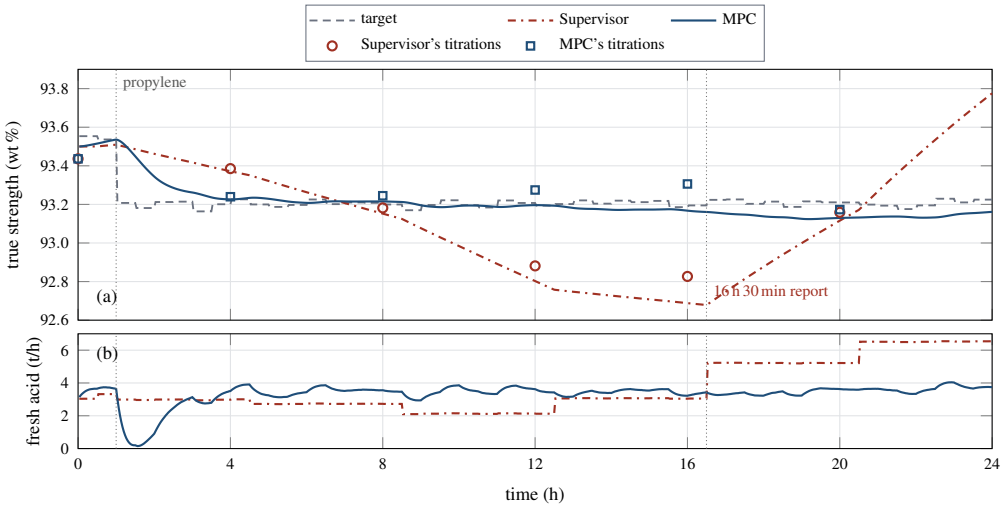


Figure 10.4 Case 9 on the default seed with the strength on the titration alone, the MPC run as the case loads (solid) and the Supervisor run after Reset (dash-dotted). (a) The true strength against the supervisor's target in force (gray, dashed), with each run's titrations at their sample times (squares for the MPC, circles for the Supervisor), each reaching its controller 30 min later. The floor band at 89 wt% lies below the panel. (b) The fresh acid applied, in the same line styles. The dotted lines mark the propylene step at 1 h and the Supervisor's report at 16 h 30 min.

samples by −0.064 to 0.099 wt%, inside the titration's own scatter. The supervisor's loop acts once per report on an error already hours old. Its strength swings between 92.68 and 93.78 wt%, and its largest error grows from 0.151 in the second period to 0.551 wt% in the last, as shown in Figure 10.4. At the 16 h sample its true strength is 92.69 wt%, and the titration, reported at 16 h 30 min, reads 92.83 against a target of 93.22, so its fresh acid jumps from 3050 to 5233 kg/h, while the MPC's strength at the same sample is 93.17 wt%. Over the day the MPC integrates 1.20 wt% h against 6.18 on 79.8 t of fresh acid against 89.0 t, and spends $1,378 less on acid and power. Over seeds 1 to 3 it scores 0.76 to 1.31 against 5.93 to 8.15, and halving the step leaves both rows unchanged to the third decimal.

The gain belongs mostly to the model between titrations and less to the plan. With the propylene left out of its model the MPC scores 6.84, no better than the titration alone. The supervisor's own PI controller acting every scan on the MPC's titration-corrected estimate, with no plan, scores 2.03, and 1.26 with the best of the 20 tunings, against the MPC's 1.20, so the quadratic program contributes the last part. A propylene feedforward of the design ASO law makes the titration loop worse (13.41), and so does one computed from the measured I/O, temperature and propylene (9.95), because a static feedforward does not carry the strength between reports. The feed chromatograph is ideal, but one sampled every 10 min and reported 5 min later changes the MPC's day only to 1.18. On the case's fixed setpoints, with the supervisor off, the MPC scores 0.77 against 2.40 for the loops with feedforward on the titration.

The density soft sensor closes most of the gap without any model. The supervisor on it scores 1.48 on 79.1 t, and with its best tuning in hindsight 1.01, better than the MPC's 1.20 but with 22 times its fresh-acid travel, while none of the 20 tunings moves the acid as little as the MPC. On margin the MPC does not lead. It earns $460,939 against $459,372 for the supervisor on the titration, but the supervisor on the soft sensor earns $461,011, and with its acid price raised to $250 a tonne $462,520, by running the acid down to 91.97 wt%.

✦ DID YOU KNOW? A Feedforward That Makes the Loop Worse

From the simulator (Table 10.6). In case 9 the supervisor's strength loop on the four-hourly titration integrates an error of 6.182 wt% h over the day. A propylene feedforward added to it scores 13.407, because a static feedforward does not carry the strength between reports. The MPC, whose model does, scores 1.202.

The MPC loses when its acid law is wrong. With the plant's κ_0 at 0.011, 29 % more ASO than the model expects, each titration corrects the estimate's offset but not its rate, the estimate stands 0.50 wt% above the truth before the last titration, and the MPC scores 7.71 against 5.58 for the supervisor on the titration alone. With fresh acid of 97.5 wt% instead of 98.5 it scores 5.92 against 5.62. With κ_0 at 0.0065 or c_3 at 2.5 it still leads (5.66 against 6.84 and 5.53 against 5.97), and a weir raised to 90 t costs it little (1.28 against 6.25). A drift estimate updated from each residual over the time since the previous sample cuts the κ_0 case to 3.43 but raises the nominal one from 1.20 to 1.97, because it also integrates the titration's noise, and it is not part of the controller.

Case 6 ships the loops with feedforward and the soft sensor on fixed setpoints of 9, 7 °C and 92 wt%. There the MPC, on the supervisor's targets, scores 3.36 against the Supervisor's 5.46 with a quarter of its fresh-acid travel and an equal margin ($456,794 against $456,819). A comparison with the shipped loops must hold their setpoints, and on them the MPC scores 0.75 against 1.20, while the loops tuned in hindsight reach 0.67, better than the MPC, with 78 times its travel. The $5,473 a day the MPC earns over the shipped loops belongs to the operating point, since the loops on the model's optimum as fixed setpoints earn $456,861.

Case 4 is a loss. With the recycle held by hand, the MPC follows the supervisor's targets of 93.21 to 93.31 wt% at 2.17 against the shipped loop's 3.85 on its 92 wt%, but it spends 165.9 t of fresh acid against 122.3 t and earns $3,477 less margin. The loss belongs to the supervisor's targets, which assume an I/O the fixed recycle cannot give. On the loop's own setpoints the MPC scores 0.75 on 128.1 t, against 2.34 for the loop tuned in hindsight, and the Supervisor on the titration reaches 20.58 on 149.5 t. In case 8, switched at 5 min with the acid raised from $80 to $250 a tonne at 4 h 05 min, the MPC follows the moving target at 6.72 against 12.34, with 35 551 against 93 556 kg/h of fresh-acid travel at an equal margin.

10.6.6 A Day of Olefins, Water and Heat

The challenge runs 24 h. The olefin rises 15 % at 2 h, a coalescer upset sends 400 kg/h of free water with the feed for 90 min from 8 h, and the cooling water warms to 38 °C over the hour from 12 h and cools again from 18 h. It is met when the product octane averages at least 94.5 and never falls below 94.0, the true strength never falls below 91 wt%, and the day's fresh acid at $150 a tonne plus compressor power at $70 a megawatt hour cost at most $24,500, with no stop. The verdict reads the simulator's truth every step and its bill uses those fixed prices, whatever the learner's price boxes say. The bill was moved to fixed prices after a strength setpoint of 94 wt% with the acid box at $20 passed while it cost $31,003. It ships on Fixed flows with the soft sensor and the moisture feedforward off. The configurations are compared in Table 10.7, and Figure 10.5 draws three of its columns against their limits.

As shipped the unit misses the octane and the strength. The I/O falls to 7.9 at 2 h 05 min and drifts to 7.6 by the end of the day as the recycle's purity falls, the octane's lowest is 93.52 at 18 h 08 min, and the acid weakens under the extra ASO and the slug to 89.08 wt%. The soft sensor, reading the heavier ASO a quarter high (8.76 wt% of the acid against 6.11 at design), puts it at 88.87 and sounds STRENGTH LOW from 22 h 44 min. The suggested settings close the ratio and strength loops with feedforward, the chromatograph, the soft sensor and the moisture feedforward, and pass with 94.91, 94.65, 91.45 wt% and $21,995, and they pass at seeds 1 to 3 with 94.90 to 94.92 and $22,040 to $22,277. The supervisor passes at 95.89 for $23,839, spending more acid on a stronger acid because the octane is worth it, and at seeds 1 to 3 for $23,933 to $24,218. The MPC with the same sensors passes at 95.91 for $23,185, and at seeds 1 to 3 for $23,186 to $23,449, but the verdict does not

Table 10.7 Configurations Tried on the Challenge Day, Default Seed

Configuration	Avg RON	Low RON	w_{min} (wt %)	Cost ($)	
Fixed flows (shipped)	94.02	93.52	89.08	18 738	fail
Loops + FF, GC, soft, FF (sugg.)	94.91	94.65	91.45	21 995	pass
Supervisor, same sensors	95.89	95.00	92.00	23 839	pass
MPC, same sensors	95.91	95.00	92.00	23 185	pass
suggested, strength loop Man	94.72	94.44	90.76	18 719	fail
suggested, no moisture FF	94.91	94.72	91.38	21 889	pass
suggested, no soft sensor	94.91	94.66	91.43	21 914	pass
suggested, neither	94.95	94.66	91.37	24 250	pass
neither, design purity	94.86	94.53	91.33	24 916	fail
suggested, I/O 10	95.31	95.00	91.41	19 899	pass
suggested, strength 94, box $20	95.43	95.00	92.00	31 003	fail

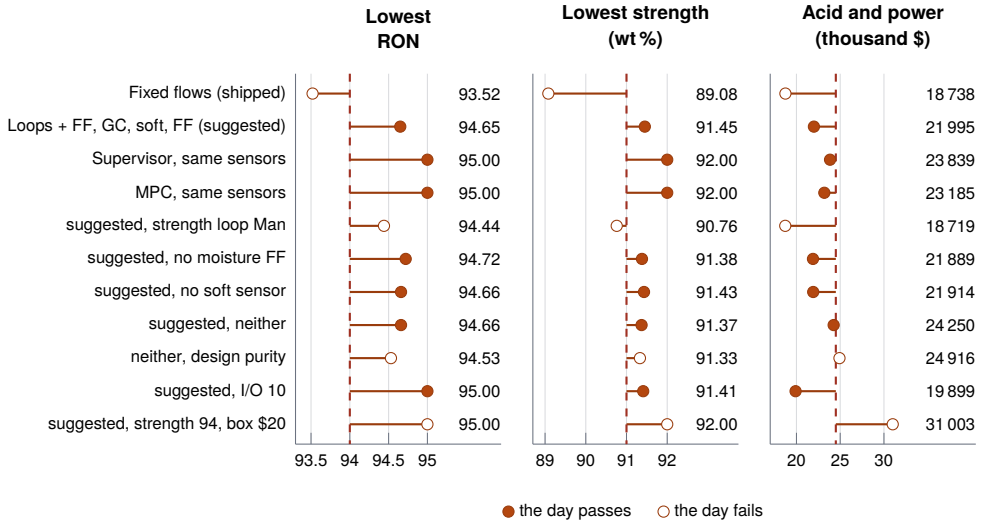


Figure 10.5 Three columns of Table 10.7 against their limits (dashed): the lowest product octane against 94.0, the lowest true strength against 91 wt% and the day's acid and power against $24,500. A stem runs from the limit to each value, and a filled circle marks a day that passes. Notice that the two cheapest days fail because they let the acid weaken, and that two other days fail on cost alone. The average octane is not drawn, because the only day that misses its 94.5 is the shipped one, which misses the lowest octane as well.

score the column. At its first period after the olefin step, 2 h 30 min, the supervisor lowers its I/O target from 11.44 to 10.75 to bring the boilup back inside its own reboiler row, and the targets stay on that row for the rest of the day. The deisobutanizer's reboiler therefore sits at its steam limit only within the hour after the step, for 32 min under the Supervisor and 52 min under the MPC. Until 2 h 30 min the MPC holds the I/O within 0.04 of its target with 105 to 165 kmol/h more recycle than the Supervisor's loop, whose I/O runs 0.16 to 0.23 below its target, and then brings its request down over about 20 min, so the alkylate carries up to 16.3 % butanes for 38 min, against 4.9 % under the Supervisor. On the loops' fixed setpoints the MPC passes at $21,786 with no minute at the steam limit.

The strength target was first set at 89 wt% and then raised to 91. At 89 the suggested settings with the strength loop left in Manual passed for $18,719, less than the suggested $21,995, by running the acid inventory down to 90.76 wt% to keep the bill low, so the day needed only the I/O loop. With the strength loop in Manual the day ends at 90.70 to 90.76 wt% at seeds 0 to 3, with or without the

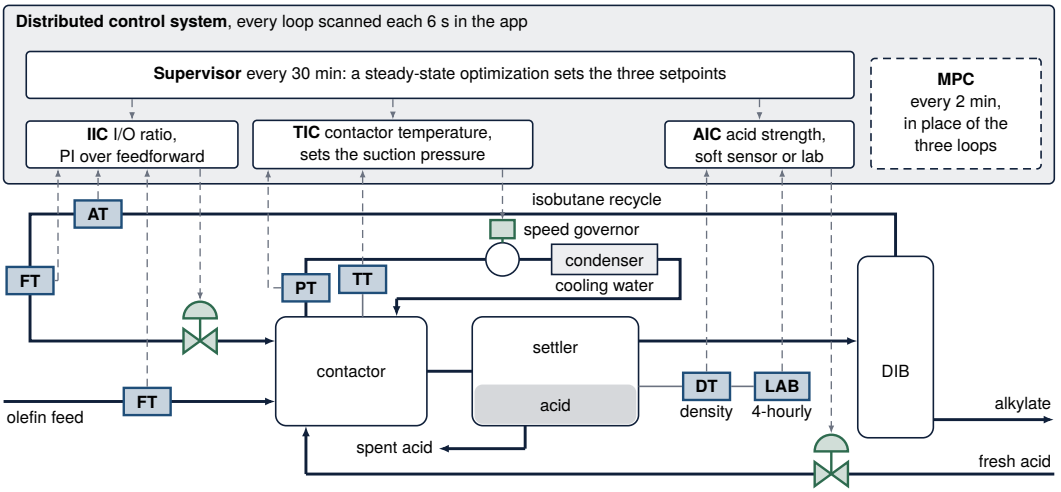


Figure 10.6 The three quality loops with their hardware. Lettered boxes are sensing devices: flow meters (FT), the recycle chromatograph (AT), the contactor thermometer (TT), the suction pressure (PT), the acid density meter (DT) and the laboratory sample (LAB). The two valve symbols and the compressor's governor are final elements, and dashed arrows are signals. Notice that the strength loop is the only one whose reference measurement is carried to a laboratory by hand, and that the temperature loop ends in a machine and not in a valve. Not to scale, and not a piping or safety diagram.

feedforwards, and the weakest closed-loop structure keeps 91.26 wt%, so 91 separates them while the alarm keeps its place at 89.

Neither the soft sensor nor the moisture feedforward is necessary for the verdict. Without either the day still passes, and with both off the cost reaches \$24,250 on the chromatograph, inside the cap at the default seed and outside it at seeds 1 to 3 (\$25,221 to \$25,423), and \$24,916 on the design purity, a miss. The two interact. The slug takes 0.65 wt% off the strength with the soft sensor and no feedforward, and 0.13 wt% with the feedforward on the titration alone, but 0.50 wt% with both on. The soft sensor reads only 0.11 of the 0.65 drop, a water dilution being nearly invisible to it, and with both on it reads 92.01 then 92.00 while the feedforward's acid lifts the true sulfuric acid share, so the PI controller trims the added fresh acid over 8 to 12 h from 6397 kg, on the titration, to 2389 kg. The 600 kg of water slugged would need 8490 kg of acid. The largest error of the corrected soft sensor over the suggested day, 0.61 wt% high, comes at 9 h 35 min, just after the slug. The bill leaves out the column's steam, so a higher I/O is free here, and an I/O setpoint of 10 on the suggested settings averages 95.31 for \$19,899.

The suggested day at 13 h is shown in Figure 10.3. In this day the compressor reaches its capacity at 12 h 00 min and stays there for 99 % of the steps from 13 to 18 h, and the contactor peaks at 8.36 °C at 17 h 20 min, 1.36 K over its setpoint, before it is back under 7.15 °C at 18 h 50 min.

10.7 Technology in Practice

Figure 10.6 draws the three quality loops of Section 10.5 with the hardware a refinery installs.

What Measures It

The ratio loop reads two flow meters and a chromatograph on the recycle. A chromatograph separates the species of a sample in time, so cycle times of minutes are not uncommon and some dead time is inevitable [23]. The model carries it as a sample every 10 min reported 5 min later. The strength

has no such instrument in the model. A laboratory titrates it every 4 h, and a soft sensor fills the gaps from a density meter. A review of soft sensors names this as their original and still dominant use, the prediction of variables that can be determined only at low sampling rates or by off-line analysis [10].

> **TECHNOLOGY** **Sensor: Acid Strength by Titration and by Density**
>
> The reference for acid strength is a titration. In the laboratory runs of a process patent the acid sample was centrifuged and a weighed portion titrated with standardized sodium hydroxide to an indicator's end point [4]. A federal case study describes a refinery sample station that gave two samples, one for an immediate density analysis by the operator and one for the laboratory to determine the weight percent of acid [24]. A vibrating-tube meter makes the density continuous. In a Coriolis flow meter, which reads density the same way, the tubes are shaken at resonance, the resonant frequency changes with the density of the fluid, and a resistance thermometer corrects for the change of the metal's elasticity with temperature [23]. Density cannot tell one diluent from another. In Section 10.5 the soft sensor reads a water dilution at 19 % of its size.

What Moves It

Two control valves and one machine carry the three loops. The recycle is the captive flow of a ratio whose wild, uncontrolled flow is the olefin [23]. The fresh acid valve moves an inventory of 80 t that it renews at 4.79 % an hour, so the valve's speed never limits its loop. The compressor does limit its loop. It gives the suction pressure that the temperature controller asks for only up to a capacity that falls on a hot afternoon.

> **TECHNOLOGY** **Final Element: Refrigeration Compressor**
>
> The contactor's temperature is moved through the compressor that draws vapor off the boiling hydrocarbon. Varying a compressor's speed sets a gas flow with less energy than running at full speed and throttling the discharge or recycling gas to the suction. Such a machine often cannot change a flow as rapidly as a control valve and offers limited options for a fail-safe state. A centrifugal compressor surges, violently, when the pressure across it is too high and the flow through it too low, and a recycle valve from discharge to suction prevents it [23]. A federal manual warns that loss of the cooling water needed to maintain process temperatures could upset an alkylation unit [19]. The model's machine has a 0.5 min governor and a capacity that falls as its condenser water warms, and it carries no surge limit.

Where the Controller Runs

A refinery runs these loops in a distributed control system, in which a supervisory level oversees local controllers, loops are sometimes nested and cycle times range from milliseconds to minutes [25]. The app scans its loops every 6 s, solves its MPC every 2 min and its supervisor every 30 min. The acid controller of the History box in Section 10.1 executed every 15 min [13], a pace that suits a strength whose clock is longer than a day. The ratio and strength loops are PI blocks over a feedforward, under a supervisor and an MPC that write setpoints. Industrial predictive controllers meet a measurement that is not available at every execution, as from an analyzer, by skipping the bias update for a number of intervals [12]. On the titration alone the chapter's MPC corrects its strength bias only when a titration reports.

> **IN PRACTICE** **An Acid Run Close to Its Limit and Sampled by Hand**
>
> The runaway of case 5 is a known event in service. A federal case study of two acid releases at one refinery in 2014 states that runaways can occur when the acid concentration falls below the operating limit, that they cost alkylation capacity and can damage equipment, and that running close to the limit is more economical, so there is a financial incentive to run weak. Operators there doubted that their controls could hold the unit at that edge, while nearby refineries ran economically with more advanced control systems. The study also shows what a sample costs. Strength control requires frequent sampling and testing, and a tubing connector beside a sample station came apart and sprayed two operators with acid [24].

What Is Modernizing It

The first development is the inferred measurement. Soft sensors also back up a hardware instrument and detect its faults, and their predictions deteriorate after launch as the process changes [10], which is why this chapter's sensor keeps a laboratory bias. The second is predictive control, of which a 2003 survey counted more than 4600 applications, with refining holding 67 % of those classified [12]. The third stands outside the control room. One refinery designed closed-loop sample stations, with a spring-loaded isolation valve that must be held open, to protect operators from the acid. Its stations plugged, foamed and were taken out of service, while another refinery made the change to more robust stations with success [24].

10.8 Limitations

The unit is generic, and its sizes, laws and limits serve teaching. The selectivity and activity laws have published directions but no fitted constants, so the floor near 87 wt%, the collapse below 89 and the clocks of 36.3 h and 7 h belong to this model, its 80 t of acid and its 24 trays. The octane correlation's strength term lies on Putney's curves, its temperature term is about three times the patent's, and its I/O term has only its sign, while the curves' maximum near 95 to 96 wt% is absent because the term only saturates. The contactor is one holdup without an emulsion, the refrigerant returns at once, and the column is ideal at a fixed pressure.

The feedforwards assume the design ASO per olefin and the soft sensor the design diluent split, so both are wrong by design when the feed or the water changes, which is the lesson of the slug. The supervisor optimizes the same correlations the plant uses and meets model error only through the instruments, and its bill and the challenge's leave out the column's steam. The loop tunings stay fixed as the feed changes, and an acid dump and refill, a contactor trip and a start-up are outside every case.

The MPC's model leaves out the deisobutanizer and its steam limit, which the column reaches for 52 min after the challenge's olefin step, and holds the compressor's capacity only as a bound on the suction pressure. Its strength row is the plant's own acid law at the declared parameters, so on the shipped plant the comparison credits a model that is exact but for its instruments, which no real unit offers, and its titration bias corrects an offset and not the drift a changed acid law produces. The feed chromatograph that gives it the propylene share is ideal, although a sampled one moves case 9 by only 0.02 wt% h.

> **OPEN CHALLENGE** **Alkylation Without a Liquid Acid**
>
> Every clock of this chapter belongs to a liquid acid that is diluted, titrated, renewed and regenerated. A review of 2003 counted four widely used processes, two on sulfuric acid and two on liquid hydrofluoric acid, and none yet on a solid catalyst [14]. In 2009 the solid catalysts proposed all showed rapid deactivation [2]. A kinetic study of 2022 reports that the liquid acids in use raise problems of safety and environmental protection, and that solid acids and ionic liquids had not reached large-scale production, the first because they deactivate easily and the second because they are expensive [16]. Progress would be a catalyst that keeps its activity at a cost near that of sulfuric acid.

10.9 Conclusion

A sulfuric acid alkylation unit holds its octane with three inventories, and this chapter measured their clocks on one engine whose balances close to 7.4×10^{-8}: 2.1 min for the contactor's temperature, 7 h 00 min for the recycle's purity behind a 24-tray column and 36.3 h for the acid's strength. A ratio on the meters reads 9.0 while the true I/O falls to 8.75, a loop on a four-hourly titration swings 0.86 points where a density soft sensor holds 0.11, and an acid withheld from 90 wt% reaches the 87 wt% floor in seven and a half hours, faster at the end than at the start. On a day of olefin, water and heat, loops with feedforward, the chromatograph, the soft sensor and the moisture feedforward pass at $21,995, the supervisor at $23,839 with the contactor on its compressor's row and the boilup on its reboiler's, while the soft sensor's blindness to water trims back the feedforward's acid so that

neither aid is necessary on its own. A model predictive controller added to the unit holds the strength between four-hourly titrations at 1.20 against the supervisor's 6.18 wt% h because its acid balance reads the feed, and it trails at 7.71 against 5.58 once the plant's acid law departs from that balance. A fuller model would fit the selectivity and octane laws to measured data, model the emulsion and several contactors, and add the column's steam to the bill. The simulator shows that a quality loop is only as fast as the slowest inventory behind it.

KEY POINTS

- An alkylation unit holds its octane with three stocks, the contactor's cold, the recycle's isobutane and the acid's strength, and each moves at its own speed.
- The clocks are 2.1 min for the temperature, 7 h 00 min for the recycle's purity and 36.3 h for the strength.
- A ratio on the meters reads 9.0 while the true I/O falls to 8.75.
- A loop on a four-hourly titration swings 0.86 points where a density soft sensor holds 0.11, and the soft sensor reads a water dilution at 19 % of its size.
- A quality loop is only as fast as the slowest inventory behind it.

Where else this applies. A soft sensor corrected by a delayed laboratory returns in Volume 5, Chapter 10, a loop on a slow analyzer in Volume 5, Chapter 3, a predictive controller with a bias in its estimator in Volume 9, Chapter 4, and a supervisor whose optimum sits on a moving constraint in Volume 9, Chapter 8.

References

[1] A. Corma and A. Martínez, "Chemistry, catalysts, and processes for isoparaffin–olefin alkylation: Actual situation and future trends," *Catalysis Reviews*, vol. 35, no. 4, pp. 483–570, 1993.

[2] L. F. Albright, "Present and future alkylation processes in refineries," *Industrial & Engineering Chemistry Research*, vol. 48, no. 3, pp. 1409–1413, 2009.

[3] J. H. Gary, J. H. Handwerk, M. J. Kaiser, and D. Geddes, *Petroleum Refining*. CRC Press, 2007.

[4] K. Kranz, "Cold temperature alkylation process," U.S. Patent 5 095 168, Mar. 10, 1992.

[5] L. F. Albright, "Alkylation of isobutane with C_3–C_5 olefins: Feedstock consumption, acid usage, and alkylate quality for different processes," *Industrial & Engineering Chemistry Research*, vol. 41, no. 23, pp. 5627–5631, 2002.

[6] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[7] U.S. Energy Information Administration, "Alkylation is an important source for octane in gasoline," Today in Energy, February 13, 2013, eia.gov/todayinenergy/detail.php?id=9971, read November 2026.

[8] W. L. Luyben, "Snowball effects in reactor/separator processes with recycle," *Industrial & Engineering Chemistry Research*, vol. 33, no. 2, pp. 299–305, 1994.

[9] B. Joseph and C. B. Brosilow, "Inferential control of processes: Part I. steady state analysis and design," *AIChE Journal*, vol. 24, no. 3, pp. 485–492, 1978.

[10] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[11] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson, "RTO: An overview and assessment of current practice," *Journal of Process Control*, vol. 21, no. 6, pp. 874–884, 2011.

[12] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[13] N. Petit, Y. Creff, L. Lemaire, and P. Rouchon, "Minimum time constrained control of acid strength on a sulfuric acid alkylation unit," *Chemical Engineering Science*, vol. 56, no. 8, pp. 2767–2774, 2001.

[14] L. F. Albright, "Alkylation of isobutane with C_3–C_5 olefins to produce high-quality gasolines: Physicochemical sequence of events," *Industrial & Engineering Chemistry Research*, vol. 42, no. 19, pp. 4283–4289, 2003.

[15] J. R. Langley and R. W. Pike, "The kinetics of alkylation of isobutane with propylene," *AIChE Journal*, vol. 18, no. 4, pp. 698–705, 1972.

[16] Z. Xin, H. Jiang, Z. Zhang, Y. Chen, and J. Wang, "Kinetic model of olefins/isobutane alkylation using sulfuric acid as catalyst," *ACS Omega*, vol. 7, no. 11, pp. 9513–9526, 2022.

[17] J. D. Freel, "A study of an alkylation reaction," M.Eng. thesis, McMaster University, Hamilton, Ontario, Sep. 1969.

[18] S. F. Birch, A. E. Dunstan, F. A. Fidler, F. B. Pim, and T. Tait, "High-octane isoparaffinic fuels production by the addition of olefins to isoparaffins," *Industrial & Engineering Chemistry*, vol. 31, no. 7, pp. 884–891, 1939.

[19] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[20] E. C. Oden and W. J. Burch, "Alkylation of isobutane with propylene: Commercial production using sulfuric acid catalyst," *Industrial & Engineering Chemistry*, vol. 41, no. 11, pp. 2524–2530, 1949.

[21] E. C. Hughes, D. G. Stevens, and F. Veatch, "Lower acidity sulfuric acid alkylation," *Industrial & Engineering Chemistry*, vol. 43, no. 6, pp. 1447–1451, 1951.

[22] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free model-predictive control," *AIChE Journal*, vol. 49, no. 2, pp. 426–437, 2003.

[23] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[24] U.S. Chemical Safety and Hazard Investigation Board, "Tesoro Martinez refinery process safety culture case study, Martinez, California, February 12, 2014 and March 10, 2014 sulfuric acid release incidents," U.S. Chemical Safety and Hazard Investigation Board, Washington, DC, Case Study 2014-02-I-CA, 2016.

[25] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

10.10 Exercises

10.1 Concept check · 2 min

In case 2 the I/O loop holds the ratio on the meters with the recycle's design purity of 89.3 % isobutane. The olefin rises 15 % at 30 min, and the loop reads an I/O of 9.0 for the rest of the day while the true I/O is 8.75 at 12 h. Why does the loop not see the shortfall?

(a) The olefin meter reads low after the step, so the loop counts less olefin than the contactor receives

(b) The integral time of 30 min is too long for the loop to follow a step of 15 % in the olefin

(c) The contactor's 1500 kmol holdup delays the new ratio by half a day, longer than the loop looks

(d) The recycle carries more n-butane as the column fills with it, while the loop still counts the design purity

10.2 Concept check · 2 min

The density soft sensor assumes the acid's diluent is water and ASO in their design shares. In the challenge a coalescer upset sends 400 kg/h of free water with the feed for 90 min. How does the soft sensor read the dilution this slug causes?

(a) It sees only about a fifth of it, because a kilogram of water lowers the density far less than a kilogram of ASO does

(b) It sees all of it within its 1 min meter lag, because density falls in proportion to the diluent's mass

(c) It sees a quarter more than the truth, because the design split counts every kilogram of diluent as ASO

(d) It sees none of it until the next titration, because its reading moves only when the laboratory bias updates

10.3 Concept check · 2 min

On Fixed flows, 20 % less fresh acid moves the strength 63 % of the way to its new value in 36.3 h, while a linearization of the acid balance at the design point gives 29.4 h. Why is the measured response slower?

(a) The titration reports four-hourly and 30 min late, which adds its delay to the measured time constant

(b) The activity's slope steepens as the strength falls, so each point lost adds more ASO and weakens the restoring rate

(c) The weir holds spent acid in the settler for 30 min, which slows the renewal of the whole inventory

(d) The acid inventory grows as ASO accumulates, so the same fresh acid renews a smaller share of it each hour

10.4 Theory and calculation · 7 min

Use the acid balance of Equations (10.7) and (10.8) at the design point: dilution $D = 270.7\,\mathrm{kg/h}$ (250.7 kg/h of ASO and 20 kg/h of water), strength 92.00 wt%, fresh acid 98.5 wt% and 80 051 kg of acid held.

(a) Compute the steady fresh acid F_f^* and spent acid F_s^*. *Checkpoint:* 3831 kg/h and 4102 kg/h

(b) How much fresh acid does each kilogram of water cost at 92 wt%, and how much would the 600 kg of the challenge's slug need? *Checkpoint:* 14.15 kg, about 8490 kg

(c) Hold F_f at its design value and linearize the strength balance, with M constant and the ASO rising 14.96 kg/h per point of strength lost. Find the time constant, and the value it would have if the ASO did not respond to the strength. *Checkpoint:* 29.4 h and 19.5 h

10.5 Theory and calculation · 7 min

Case 5 lines the unit out at 90 wt% with the design I/O of 9, the contactor at 7 °C, no propylene and 20 kg/h of feed water, then withholds the fresh acid. Use Equation (10.1) with $\kappa_0 = 0.0085$, $F_O = 530\,\mathrm{kmol/h}$, $M_O = 56.11\,\mathrm{kg/kmol}$ and the chapter's values $1/a = 1.131$ at 90 wt% and 2.025 at 87 wt%.

(a) Compute the ASO made, in kg/h, at 90 wt% and at 87 wt%. *Checkpoint:* 283 kg/h and 503 kg/h

(b) What fresh acid holds the unit at 90 wt%? *Checkpoint:* 3211 kg/h

(c) With the fresh acid at zero the acid held settles near 78.2 t, at the weir. Estimate the rate at which the strength falls at 90 wt% and at 87 wt%, and compare with the 0.35 points lost in the first hour and 0.52 in the last. *Checkpoint:* 0.35 and 0.58 wt%/h

10.6 Theory and calculation · 7 min

The ratio feedforward of Equation (10.11) sets the recycle $R = (I_{SP}\,z_O - z_I)F_{OF}/\pi$ with $z_O = 0.55$, $z_I = 0.30$ and the design olefin feed $F_{OF} = 963.6\,\mathrm{kmol/h}$.

(a) Compute the recycle for an I/O of 9 at the design purity $\pi = 0.893$. *Checkpoint:* 5018 kmol/h

(b) Case 2 raises the olefin feed by 15 %. What recycle does the feedforward now ask for? *Checkpoint:* 5770 kmol/h

(c) At 12 h of case 2 the recycle's true purity is 86.7 % while the loop still uses 0.893. With the recycle of part (b), compute the true I/O from Equation (10.1), and the recycle that would restore an I/O of 9. *Checkpoint:* 8.75 and 5943 kmol/h

10.7 App challenge · 6 min

Load case *4 · Acid strength on a four-hour laboratory*, default seed, which opens with the recycle in Manual and the Soft sensor Off. Run the full 24 h at a speed of 300 min/s and export the CSV.

(a) Read the lowest and highest true strength (column `truth_w_wt`) and the lowest and highest fresh acid command (column `fresh_cmd_kgh`) over the day. *Checkpoint:* 91.45 to 92.31 wt%, 3205 to 6472 kg/h

(b) Set the Soft sensor On, press Reset, run the day again and read the same four values. *Checkpoint:* 91.92 to 92.03 wt%, 3774 to 5774 kg/h

(c) Why does the loop on the titration alone swing so much more, on the same tuning?

10.8 App challenge · 6 min

Load case *5 · Withhold the acid: the runaway floor*, default seed. Before pressing Run, put the Strength loop on Man and set the Fresh acid to 0 kg/h. Run at 30 min/s. The run stops itself. Export the CSV.

(a) Read the times at which the true strength (`truth_w_wt`) first falls below 89 wt% and below 88 wt%, and the time of the ACID RUNAWAY stop, each within 1 min. *Checkpoint:* 2 h 48 min, 5 h 23 min, 7 h 30 min

(b) Read how many points of strength are lost in the first hour and in the last hour before the stop. *Checkpoint:* 0.35 and 0.52

(c) The case's completion rule asks for the same two moves after Run. If they are made at 5 min instead, what changes?

10.9 App challenge · 7 min

Load the challenge, *a day of olefins, water and heat*, default seed, and run all 24 h at 300 min/s as shipped (Fixed flows, Soft sensor and Moisture feedforward Off).

(a) Read the verdict: average product RON, lowest RON, lowest true strength and the cost of acid and power. Which targets fail? *Checkpoint:* 94.02, 93.52, 89.08 wt%, $18,738

(b) Press Reveal suggested tuning, then Reset, and run the day again. Read the same four values. *Checkpoint:* 94.91, 94.65, 91.45 wt%, $21,995

(c) Set both the Soft sensor and the Moisture feedforward Off, press Reset and run again. Does the day still pass, and what does the result say about whether either aid is necessary? *Checkpoint:* 94.95, 94.66, 91.37 wt%, $24,250

10.10 Application area · 4 min

Refiners run the contactor of a sulfuric acid unit cold.

(a) The chapter's octane law gives $b = 0.12\,\mathrm{K}^{-1}$. The cited patent reports about half an octane number gained over the 11.7 K from 49 to 28 °F. Compare the two slopes. *Checkpoint:* 0.043 K^{-1}, about a third

(b) Name two things that limit how cold a real contactor can run, from the chapter's history and model.

(c) In case 8 the supervisor places the contactor on its compressor row. What does a contactor 0.2 K warmer cost there, and why is this optimum on a limit?

10.11 Application area · 4 min

At design the isobutane balance closes as 289.09 kmol/h in the feed plus 322.90 of makeup against 525.53 alkylated, 85.46 in the side draw and 1.00 in the bottoms. In case 7 the makeup supply is lost.

(a) Suppose the alkylation and the feed's isobutane both scale with the olefin while the side draw and bottoms losses stay as they are. What isobutane deficit remains with the olefin cut to 60 %? *Checkpoint:* about 228 kmol/h

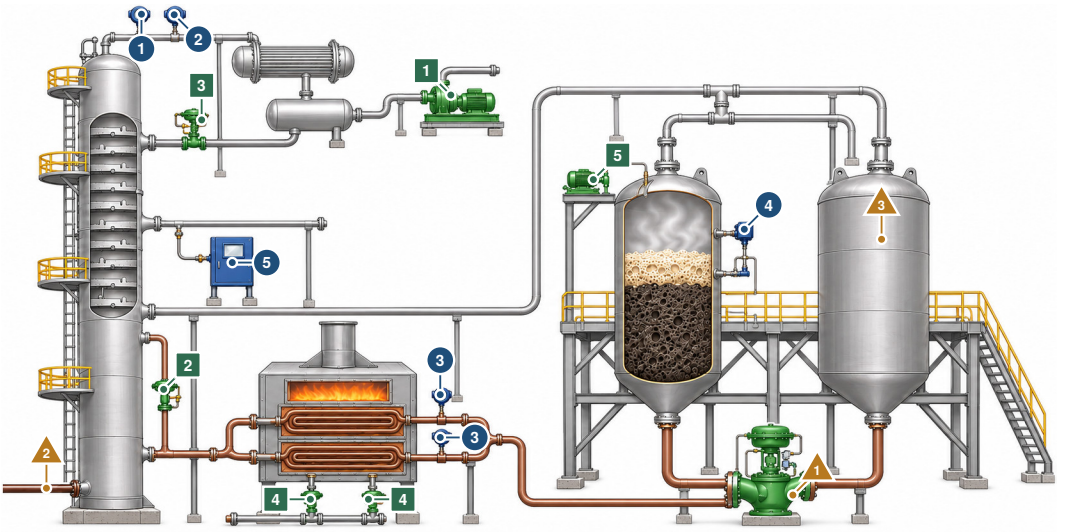
(b) Is there an olefin rate that holds the ratio without makeup? What does the app do when the true I/O stays below 5, and where would a real unit's setting for such an interlock come from?

10.12 Review the agent · 5 min

An AI assistant reads case 9, where the MPC scores an IAE of 1.20 against the supervisor's 6.18 wt% h on the titration alone, and writes: "The quadratic program plans four hours ahead, so it sees each titration coming. That plan is the source of the fivefold gain, and the same controller on a real unit, where the acid law is only roughly known, would hold the strength about five times closer than a PI loop on the laboratory." Find the flaw and say how to check it in the app.

CHAPTER 11

A Delayed Coker Whose Drum Switch Is a Disturbance Known in Advance



The controller moves 1 wet gas compressor 2 valve on the wash recycle 3 top reflux valve 4 fuel to the two furnace passes 5 antifoam pump

The controller reads 1 fractionator pressure 2 overhead temperature 3 two coil outlet temperatures 4 foam level on the live drum 5 analyzer on the heavy gas oil

What disturbs it 1 the scheduled drum switch 2 a heavier, foamier residue 3 a cold drum after a skipped warm-up

Notice the switch valve below the drums: at each switch it moves the charge from the cut-open live drum, where foam rides on the coke bed, to the closed drum beside it, which is colder than the vapor it will receive. Illustration, not to scale.

IN PLAIN TERMS Two Drums That Take Turns

The main idea. A delayed coker heats the heaviest part of crude oil and lets it break apart for hours in a tall drum. Solid coke stays behind, and the vapor goes on to a column that splits it into fuels.

The challenge. When a drum is full, the flow moves to an emptied drum that is still cooler than the oil. That drum soaks up heat and holds back vapor, and the column loses pressure unless its controls act on the known switch time.

Why it matters. The switch is the most predictable upset in the plant, yet only warming the new drum first saves the vapor.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the History box in Section 11.2, the Did you know? boxes in Sections 11.1 and 11.5, the Try it boxes in Section 11.3, and Key points in Section 11.8. No equations.
Study	45 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 11.5 with Figure 11.4, Section 11.6 and Section 11.7 with its Open challenge box.

The control question A delayed coker's two drums take turns on a fixed cycle while its furnace and fractionator run without a pause, and every switch sends the charge into a drum colder than its vapor. How much of that upset can control remove when the schedule that causes it is known hours ahead, and what can only the drum's warm-up prevent?

Learning objectives

- Compute the coke made, the rise of the bed and the outage left at the end of a fill from the yield correlation and the drum size.
- Calculate the vapor a cold drum holds back from its shell temperature, at the switch and in total.
- Predict the first compressor move of a feedforward built from the schedule and the shell thermocouples.
- Compare loops, schedule feedforward and MPC on a cold switch, and explain why grading the heavy gas oil cut at every instant separates them.
- Weigh the coil outlet temperature against coke yield and furnace run length, and the antifoam dose against foam-over.

Prerequisites Sections 1.4.2 (heaters and tube metal), 1.4.3 (cut points), 1.4.6 (analyzer delay), 1.6.2 (velocity form) and 1.6.4 (scenarios). Section 1.6.3 is needed only for the MPC. Chapter 5 makes the residue this unit cracks.

The app https://apmonitor.com/apps/delayed-coker/

See also Volume 3, Chapter 6 for feedforward from a measured load beside a backwash sequencer, Volume 3, Chapter 7 for a foam limit that moves with pressure and Volume 3, Chapter 4 for heater outlet control. Volume 8, Chapter 6 for a recipe the controller knows in advance and Volume 8, Chapter 13 for a continuous plant that a diversion sequence interrupts.

11.1 Introduction

The coke drum is a batch reactor that sits inside a continuous refinery unit. In the generic coker simulated here, 120 t/h of vacuum residue and 12.0 t/h of recycle leave a two-pass furnace at a coil outlet temperature (COT) of 495 °C, and the charge cracks in whichever of two drums is on line. About 36.6 t/h of coke stays behind, so the bed rises 1.12 m an hour, and after 16 h the switch valve sends the charge to the other drum, which has meanwhile been steamed, quenched, cut out and warmed. A drum whose shell sits 300 K below the vapor it receives holds back 16.2 t/h of vapor at the switch, and the fractionator, which sees only what leaves the drums, finds out first.

> ✦ **DID YOU KNOW? Coking Capacity in 2003**
>
> On January 1, 2003, the coking units of United States refineries could take 2,243,947 barrels per calendar day, 13.5 % of the country's crude distillation capacity [1]. The coker of this chapter takes 120 t/h of vacuum residue.

The simulator is built around one belief it is meant to correct, that a recurring disturbance belongs to feedback. A switch is the most predictable upset in the plant, because the sequencer sets its time and the shell thermocouples measure how cold the new drum is. On a switch onto a drum left cold by a skipped warm-up (case 8), the fractionator pressure falls 0.178 bar under its loops alone, while a schedule feedforward that moves the wet gas compressor with the predicted gas keeps the dip to 0.024 bar. A second belief follows from the first, that a feedforward which holds the pressure has fixed the switch. The cold drum keeps 15.4 t of vapor in the drum over four hours whatever the loops do, and only the warm-up moves that heat to a time the fractionator can spare it. The same shape, a batch step on a known schedule feeding a continuous section, appears in bed regenerations, filter backwashes, dryer and adsorber swaps and cyclic reactors, and the integration of production schedules with process control is reviewed as a field of its own [2]. Delayed coking itself is reviewed in [3, 4], and model predictive control of the main fractionator of a cracking unit in [5].

However, the published material on cokers separates the drum, whose kinetics and bed structure are modeled without a fractionator [6–8], from the furnace, whose coil fouling is treated as a run-length problem [9], and from the drum's mechanical life [10]. This chapter puts the drum, the furnace and the fractionator on one simulated coker whose sequencer, drum shells, foam, furnace coils and fractionator loops act on one another, and it measures what changes when the schedule is handed to the controllers and what does not.

The yield correlation, the velocity-form PI controller, the feedforward and the linear model predictive controller (MPC) are established methods, and the results are measurements on one synthetic unit. One of them carries the chapter. A pressure PI with five times the shipped gain also holds the challenge's pressure band, yet it leaves the heavy gas oil cut 9.9 K off at a cold switch, so the challenge, which grades that cut at every instant, needs the feedforward or the MPC. The model has 41 states, its charge balance over each drum cycle closes to 2.6×10^{-13} of the charge, and its controllers see instruments only, never the pitch, the true foam or the coil coke. Each of the nine cases credits its completion by a rule written against the plant state (Section 1.6.4).

Section 11.2 follows delayed coking and its drum cycle. Section 11.3 sets out the unit as modeled, from the yields and the live drum through the sequencer, the warm-up, the foam and the furnace coils to the fractionator, its pressure and the instruments. Section 11.4 describes the control layers, which are five PI loops, the schedule feedforward and the MPC. Section 11.5 reports the nine cases, from one fill by hand to two graded days on the unit, Section 11.6 sets the hardware of a real unit beside the model, Section 11.7 states what the model leaves out, and Section 11.8 draws the lessons together.

11.2 Delayed Coking and Its Drum Cycle

Delayed coking grew out of the first thermal cracking processes, as the History box below tells. The drums make the unit a batch process. A full drum leaves the line to be steamed, quenched, drained and emptied of its coke, and is warmed before it fills again, the sequence the simulator's sequencer runs in Section 11.3.

> ⧖ HISTORY **1913: From Unwanted Coke to Coke Drums**
>
> The first thermal cracking process was developed around 1913. Distillate fuels and heavy oils were heated under pressure in large drums until they cracked, and the method made large amounts of solid, unwanted coke. That process evolved into visbreaking, steam cracking and coking, which turns heavy residue into lighter products and by-product coke [11]. In delayed coking the furnace is designed and controlled to prevent premature coking in its tubes, and large coke drums give the long residence time the cracking needs. When the coke reaches a set level in one drum, the flow is diverted to another to keep the unit running, and the full drum is steamed, cooled with water and cut out by high-pressure water from a rotating cutter [11]. Fuel demand and ever heavier crudes renewed interest in the bottom of the barrel, and delayed coking, able to take virtually any refinery stream, gained in importance [3].

Yields came to be predicted from the feed. Refining texts give coke, gas and liquid yields as functions of the residue's Conradson carbon residue (CCR) [12], the residue's chemistry sets how it cracks [13], and an industry consortium studied the fundamentals of the process, from feed and operating conditions to coke morphology [14]. Inside the drum, kinetic models describe the deep thermal cracking [6, 7] and the bed that results is far from uniform [8]. Two limits outside the drum shape the cycle. The furnace coils foul with coke, so the furnace runs until its tube metal approaches its limit and is then decoked, and antifoulants are judged by laboratory tests against field experience [9]. The heater of Chapter 5, whose residue this unit takes, cokes its coil at a rate that doubles with every 10 K of film temperature, as the coils of Section 11.3.4 do. The drums themselves fatigue under the temperature swing of every cycle, so throughput is weighed against drum life [10].

Control followed the general history of refinery automation. Each draw, the furnace outlets and the overhead pressure received their own loops, and model predictive control reached fractionators of cracking units by the 1990s [5, 15]. The drum switch remained an operating event for the regulatory

Table 11.1 Variables and Their Roles at the Design Point

Variable	Role	Design	Range
P	CV, overhead pressure	1.000 bar g	SP 0.6–1.6
T_{95}	CV, HCGO 95 % point	480.0 °C	SP 460–500
T_{top}	CV, overhead temperature	130.0 °C	SP 110–150
COT_i	CV, coil outlets	495.0 °C	SP 470–515
h_{out}	CV (on foam level), outage	18.45 m	SP 2–10
u	MV, compressor	75.2 %	0–100
R	MV, recycle	10.0 % of F_f	0–30
L_r	MV, top reflux	40.0 t/h	0–150
Q_i	MV, absorbed duty per pass	10.10 MW	0–13
D	MV, antifoam	3 ppm	0–25
L	set, LCGO draw	22.4 t/h	0–40
F_f	DV, fresh residue	120 t/h	80–140
CCR, φ	DV, carbon, foaming	17.5 wt %, 1	12–26, 0.5–4
—	DV, compressor available	100 %	50–100
t_{cyc}, t_w, f_w	schedule	16 h, 4 h, 8 %	10–24 h
M_k, M_p, T_m, h_f	states, per drum	—	—
$Q_i, \mathrm{COT}_i, \delta_i$	states, per pass	—	—
T_N, T_L, T_R, T_{fz}	states, cut and flash zone	—	—

loops, while the textbook answer to a measured disturbance, feedforward [16], applies directly once the schedule is treated as a measurement, as Volume 3, Chapter 6 computes its coagulant dose from a raw water turbidimeter rather than wait for the settled water. Coordinating a schedule with the loops it disturbs is now studied in general [2], and the coker is one of its clearest cases because the schedule is fixed by the decoking crew and known to the minute.

11.3 The Unit as Modeled

The unit is generic and taken from no plant, as shown in Figure 11.1. Fresh residue F_f enters the fractionator's bottoms, where the wash adds the recycle R, and the charge $F_c = F_f + R$ is pumped through two furnace passes A and B, each taking half of it with no pass flow balancing. The passes join in the transfer line, the switch valve sends the charge to the live drum, and the drum's vapor goes to the flash zone of the main fractionator. The off-line drum follows its own sequence, and for the last 4 h of a 16 h cycle part of the live vapor is drawn through it to warm its shell. The fractionator washes the heaviest ends back into the bottoms as recycle and draws heavy and light coker gas oil (HCGO and LCGO), naphtha and the wet gas that the compressor takes. Twenty-nine physical states and twelve ledgers are integrated, the ledgers being the running totals of charge, vapor, condensate, coke laid and cut, products, recycle, antifoam, fuel, gas and liquid that let conservation be checked. The clock is in minutes, flows in t/h, duties in MW and pressures in bar g.

Table 11.1 classifies the variables and gives their values at the design point, case 3 at $t = 0$, four hours into drum A's fill.

11.3.1 Drum Yields and the Live Drum

The charge cracks in the live drum at yields set by a correlation of the published form [12], written on the residue's Conradson carbon, the COT and the drum pressure, as shown in Equation (11.1).

$$\begin{aligned} y_c &= 0.016\,\mathrm{CCR} - k_T\,\Delta\mathrm{COT} + k_P(p_d - 1.70) \\ y_g &= 0.078 + 0.00144\,\mathrm{CCR} + k_g\,\Delta\mathrm{COT} \\ C &= F_f y_c + R\,y_{c,R}, \qquad V_{gen} = F_c - C \end{aligned} \tag{11.1}$$

where $\Delta\mathrm{COT} = \mathrm{COT} - 495\,^\circ\mathrm{C}$, y_c and y_g are the coke and gas fractions of the fresh feed, clipped to 0.05–0.6 and 0.02–0.3, $k_T = 0.002$ and $k_g = 0.0008$ per K, $k_P = 0.015$ per bar, the recycle cracks a quarter to coke ($y_{c,R} = 0.25$) and 5 % to gas, and C and V_{gen} are the coke and the vapor the

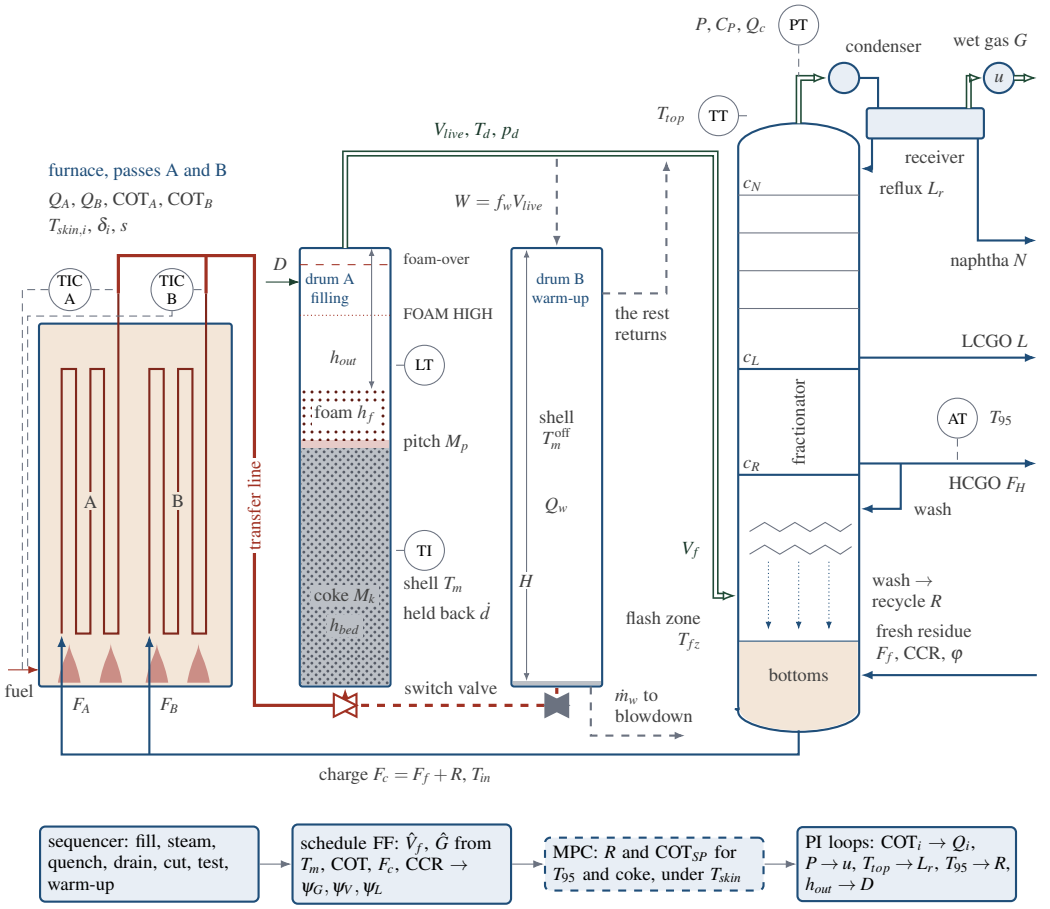


Figure 11.1 The two-drum delayed coker as modeled, redrawn from the Unit view of case 1 at $t = 10$ h. Drum A has filled 13 h and holds a coke bed of 14.6 m under 3.0 m of foam, leaving an outage of 8.4 m below the vapor line, with the FOAM HIGH (4 m) and foam-over (1 m) lines marked. Drum B is in its warm-up, its shell at 223 °C, condensing 4.2 t/h of a 7.6 t/h draw to blowdown while the rest rejoins the overhead line. The fractionator receives 91.3 t/h, its wash returns 12.1 t/h of recycle to the bottoms, and the analyzer AT reads the HCGO 95 % point. Drum heights are to scale, everything else is not. Liquid streams are thick solid lines (blue), the furnace effluent very thick (red, dashed on the shut leg to drum B), vapor and gas hollow double lines (green) and the warm-up gray dashed lines. The strip below lists the control layers, the MPC in a dashed box because it is an option.

charge makes. At 17.5 wt % and 495 °C the coke is 30.5 % of the fresh feed, the steady design value. The drum pressure is algebraic, $p_d = P + \Delta p_\ell (V/V_0)^2$, with $\Delta p_\ell = 0.7$ bar at the design vapor $V_0 = 95.4$ t/h.

A drum whose shell is colder than its vapor takes heat from its contents, and the charge that heat

would have cracked stays behind as pitch, as shown in Equation (11.2).

$$\begin{aligned} C_m \frac{\mathrm{d}T_m}{\mathrm{d}t} &= UA_d (T_{d0} - T_m) \\ \dot{d} &= UA_d \max(0,\, T_{d0} - T_m)/h_{eff} \\ \frac{\mathrm{d}M_p}{\mathrm{d}t} &= \dot{d} - \frac{M_p}{\tau_p}, \quad \frac{\mathrm{d}M_k}{\mathrm{d}t} = C + y_{c,p}\frac{M_p}{\tau_p} \\ V_{live} &= V_{gen} - \dot{d} + (1 - y_{c,p})\frac{M_p}{\tau_p} \end{aligned} \tag{11.2}$$

where $T_{d0} = \mathrm{COT} - 45\,\mathrm{K}$ is the drum vapor, $C_m = 60\,\mathrm{MJ/K}$, $UA_d = 18\,\mathrm{kW/K}$, $h_{eff} = 1200\,\mathrm{kJ/kg}$ is the cracking and vaporization heat per kilogram of vapor not made, $\dot{d}$ is capped at $0.6V_{gen}$, the pitch M_p cracks back with $\tau_p = 60\,\mathrm{min}$, 35 % of it to coke ($y_{c,p}$), and M_k is the coke. The shell's time constant on line is $C_m/UA_d = 55.6\,\mathrm{min}$, and each kelvin of shell below the vapor holds back 0.054 t/h. A quarter of the shell's heat is taken from the vapor's temperature, $T_d = T_{d0} - 0.25\,UA_d(T_{d0} - T_m)/(2.5\,V_{gen})$ with V_{gen} in kg/s, so the flash zone cools at a switch as well. The bed height is $M_k/(\rho_k A) + M_p/(0.95A)$ with $\rho_k = 0.90\,\mathrm{t/m^3}$ and $A = 36.3\,\mathrm{m^2}$ for a 6.8 m drum, 32.7 t of coke per meter.

🕮 **Pause.** Try Exercise 11.1 before reading on.

11.3.2 The Sequencer, the Off-Line Drum and the Warm-Up

The sequencer is plant state, advanced after each step, and exactly one drum is live at every instant. Off line a drum passes through steam-out (1.5 h), water quench (4), drain (1.5), cut and re-head (4) and steam test (1), twelve hours in all, then stands by until its warm-up, scheduled to end at the switch. In each phase its shell relaxes toward a declared temperature (380, 100, 95, 60 and 155 °C), the pitch left at the switch cracks during steam-out and is quenched into the coke, and the coke is cut out at a uniform rate. A switch falls due when the live drum has filled for t_{cyc}, and it is held until the other drum is ready, so a cycle under 12 h cannot be kept. During the warm-up a fraction f_w of the live vapor passes through the off-line drum, as shown in Equation (11.3).

$$\begin{aligned} W &= f_w V_{live}, \qquad Q_w = \varepsilon c\, W (T_d - T_m^{\mathrm{off}}) \\ C_m \frac{\mathrm{d}T_m^{\mathrm{off}}}{\mathrm{d}t} &= Q_w, \qquad \dot{m}_w = \min\left(W(1-g),\; Q_w/\lambda_w\right) \\ V_f &= V_{live} + V_{steam} - \dot{m}_w \end{aligned} \tag{11.3}$$

where $\varepsilon c = 2.16\,\mathrm{kJ/(kg\,K)}$, $\lambda_w = 900\,\mathrm{kJ/kg}$, g is the gas share of the vapor, V_{steam} is the steam-out's cracked pitch and V_f is the vapor reaching the fractionator. Only the condensate $\dot{m}_w$, never the gas, drains to blowdown, and the rest of the draw rejoins the overhead line, so the draw is subtracted from the fractionator once, as its condensate. In case 1 the warm-up starts at 9 h with the shell at 151 °C, draws 7.64 t/h and condenses 5.47 t/h, falling to 4.16 t/h at 10 h and 1.83 t/h by the switch, 13.3 t over the warm-up, and leaves the shell at 350.5 °C.

> **TRY IT IN THE APP** **1 · One fill, by hand**
>
> Load the case and press Run. Every station stays in Manual. Watch drum A: its bed climbs from 3.4 to 18.0 m with the foam riding on it, and the outage shrinks from 19.6 to 5.0 m. At 13.0 h the sequencer switches the charge to drum B, and with no loop to answer it the pressure falls to 0.907 bar g.

11.3.3 Foam and Outage

The foam on the live drum's bed approaches a steady thickness set by the vapor rate, the pressure, the residue's foaming tendency and the antifoam dose, as shown in Equation (11.4).

$$\begin{aligned} h_{f,ss} &= \frac{h_n \varphi \,(V_{live}/V_0)\,\pi_p}{1+\eta D}, \qquad \pi_p = \frac{p_{d0}+p_{atm}}{p_d+p_{atm}} \\ \tau_f \frac{\mathrm{d}h_f}{\mathrm{d}t} &= h_{f,ss} - h_f, \qquad h_{out} = H - h_{bed} - h_f \end{aligned} \tag{11.4}$$

where $h_n = 6\,\mathrm{m}$, $\eta = 0.33$ per ppm, $\tau_f = 20\,\mathrm{min}$, $p_{d0} = 1.70\,\mathrm{bar\,g}$ and $H = 26\,\mathrm{m}$ to the vapor line. At 3 ppm the design foam is 3.0 m, and a 16 h fill ends with 5.0 m of outage. FOAM HIGH below 4 m of outage and a latched FOAM-OVER at 1 m, which stops the run, are teaching values. Real alarm points come from the drum's radiation level detectors and the licensor. Through the factor π_p a fall in drum pressure swells the foam, the lever that Volume 3, Chapter 7 meets as a pressure margin against foaming at a filler.

> **TRY IT IN THE APP** **5 · Foam and antifoam**
>
> Load the case and press Run. The residue foams three times as much as the design residue, and the antifoam is fixed at 3 ppm. The outage reaches 1 m and the run stops with FOAM-OVER at 11.2 h. Set Antifoam to On foam level and press Reset. The loop then holds the outage within 4.39 to 5.44 m and uses 6.2 kg.

11.3.4 The Furnace and Its Coils

Each pass carries an absorbed duty that follows its fuel command through the firebox and a coil outlet that lags its heat balance, as shown in Equation (11.5).

$$\begin{aligned} \tau_{fb} \frac{\mathrm{d}Q_i}{\mathrm{d}t} &= Q_{i,cmd} - Q_i \\ \tau_c \frac{\mathrm{d}\mathrm{COT}_i}{\mathrm{d}t} &= T_{in} + \frac{Q_i}{F_i c} - \mathrm{COT}_i \\ T_{skin,i} &= \mathrm{COT}_i + \theta_f \phi_i \frac{Q_i}{Q_0} + \beta \delta_i \\ \frac{\mathrm{d}\delta_i}{\mathrm{d}t} &= k_f \, 2^{(T_{film,i} - T_{film,0})/10\,\mathrm{K}} \sqrt{F_0/F_i}\; s^{-0.4} \end{aligned} \tag{11.5}$$

where $\tau_{fb} = 2\,\mathrm{min}$, $\tau_c = 3\,\mathrm{min}$, $Q_{i,cmd} \le 13\,\mathrm{MW}$, $F_i = F_c/2$, $c = 3.0\,\mathrm{kJ/(kg\,K)}$ with the coil's vaporization, $\theta_f = 55\,\mathrm{K}$, $Q_0 = 10.10\,\mathrm{MW}$, $\phi_A = 1$ and $\phi_B = 1.06$ because pass B sits nearer the burners, $\beta = 8\,\mathrm{K/mm}$, $k_f = 0.267\,\mathrm{mm/day}$, $T_{film,i} = \mathrm{COT}_i + 0.55\theta_f \phi_i Q_i/Q_0$ with $T_{film,0} = 525.25\,^\circ\mathrm{C}$, $F_0 = 66\,\mathrm{t/h}$ and s the velocity steam in wt % of the charge. The skin is the tube metal temperature of Equation (1.3) with a coke term that grows by the coking law. The charge enters at T_{in}, which lags the mix of fresh residue at 300 °C and recycle at $T_{fz} - 25\,\mathrm{K}$ with 10 min, 311.4 °C at design. On coil day 2 the skins are 554.3 and 558.2 °C. A compression parameter multiplies the coil coking rate so that case 7 runs 100 coil days per simulated day, and the skin limit of 650 °C is a teaching value for the coil alloy's design metal temperature.

11.3.5 The Fractionator and Its Pressure

The vapor reaching the flash zone is cut on its own pseudo true-boiling-point curve, through 15 points from −160 °C at zero to 200 °C at 0.354, 345 °C at 0.589, 495 °C at 0.874 and 620 °C at 1, as shown in Equation (11.6).

$$
\begin{aligned}
c_N &= g_F + n_0\left(\frac{\rho_0}{L_r/V_f}\right)^{0.45}, \quad c_L = c_N + \frac{L}{V_f} \\
c_R &= 1 - \frac{\min(R_w,\, 0.3V_f)}{V_f} \\
F_H &= (c_R - c_L)V_f, \qquad T_{95} = T_R - 15\,\mathrm{K} \\
\tau_N \frac{\mathrm{d}T_N}{\mathrm{d}t} &= \mathrm{TBP}(c_N) - T_N \quad \text{(likewise } T_L,\, T_R\text{)} \\
C_P \frac{\mathrm{d}P}{\mathrm{d}t} &= G - u\,Q_c \frac{P + p_{atm}}{P_0 + p_{atm}}
\end{aligned}
\tag{11.6}
$$

where g_F is the gas share of V_f (all the gas arrives), $n_0 = 0.2175$, $\rho_0 = 0.42$, R_w is the wash's recycle command after a 0.5 min valve lag, the LCGO draw is limited to $(c_R - c_N - 0.02)V_f$, and the cut temperatures lag with 4, 8 and 10 min. The HCGO F_H is the balancing stream between the LCGO and the recycle, so its 95 % point is the HCGO and recycle cut. At design one more percent of recycle lowers it 7.29 K and adds 0.3 points of coke. The overhead temperature is $T_{top} = T_N - 70\,\mathrm{K} + 25\,\mathrm{K/bar}\,(P - P_0)$, the flash zone lags the drum vapor with 5 min, and the recycle reaches the charge through a 15 min bottoms lag. In the pressure balance $G = g_F V_f$ is the wet gas that the compressor takes, $C_P = 0.645\,\mathrm{t/bar}$, $Q_c = 17.3\,\mathrm{t/h}$ and $P_0 = 1.0\,\mathrm{bar\,g}$, which gives an overhead time constant of 6.0 min with the compressor at 75.2 % of its capacity. The condensable vapor is taken as condensed at once.

11.3.6 Instruments, Initialization and Numerics

The controllers read a measurement packet and nothing else. The coil outlets carry a 0.2 min lag and 0.3 K of noise, the skins 1 min and 1 K, the overhead 0.3 min and 0.2 K, the drum shells 2 min and 1 K, the overhead pressure 0.05 min and 0.004 bar, the flows 0.3 %, and the radiation level detectors read each drum's foam top with 1 min and 0.15 m. An online distillation analyzer draws a sample every 10 min, analyzes it for 10 min behind a 5 min sample line and reports T_{95} with 0.7 K of noise, a sampled analyzer whose effective delay follows Equation (1.12). The feed tank's assay gives the CCR, and the sequencer's phases and clocks are known exactly because the control system runs them. Every case starts lined out at its setpoints, with the drums pre-run through two cycles to the hour of drum A's fill the case names. The engine of Section 1.6.1 integrates by fourth-order Runge–Kutta with a 3 s step, the controllers execute every 6 s and the historian logs every 12 s. The exported log sets every reading beside the simulated quantity behind it, the pitch, the true foam, the held-back vapor, the coil coke and the ledgers, in columns whose names begin with `truth_`. Over every full cycle in the challenge and in case 8, the charge equals the vapor to the fractionator, the warm-up condensate and the coke laid to within 2.6×10^{-13} of the charge, and at every step $V_f = V_{live} + V_{steam} - \dot{m}_w$ exactly. The parameters are collected in Table 11.2.

11.3.7 Assumptions

The yields are teaching correlations of a published form [12], and real cokers are characterized on their own residues [13, 14]. A drum is one mass of coke, one pool of pitch and one shell temperature, with no bed structure [8], no liquid level inside the bed and no thermal stress, which is the real limit on how fast a cycle may run [10]. The new drum's deficit is written as vapor not made, while a real drum also condenses vapor on its walls. Both hold material in the drum for a while and both depend on the shell's temperature. The foam is a steady thickness with no ceiling, so on a foaming residue with little antifoam it stands as tall as 18 m, which a real unit would not allow. The fractionator has

Table 11.2 Parameters. A: Adjustable Under Parameters in Show Model. V: Fixed and Stated in the App. H: Fixed and Not Displayed

Symbol	Meaning	Value	
H, D_d, ρ_k	drum height, diameter, bed density	26 m, 6.8 m, 0.90 t/m^3	V
C_m, UA_d	shell capacity, contents to shell	60 MJ/K, 18 kW/K	V
h_{eff}, τ_p, $y_{c,p}$	vapor not made, pitch	1200 kJ/kg, 60 min, 0.35	V
εc, λ_w	warm-up exchange, condensing	2.16 kJ/(kg K), 900 kJ/kg	V
k_T, k_g, k_P	yield slopes, points of coke or gas	0.20/K, 0.08/K, 1.5/bar	V
h_n, τ_f	foam at design vapor, lag	6 m, 20 min	V
η	antifoam efficiency	0.33 per ppm	A
—	FOAM HIGH, FOAM-OVER outage	4 m, 1 m	A
—	antifoam pump capacity	25 ppm	V
Q_{max}	duty limit per pass	13 MW	A
τ_{fb}	firebox	2 min	H
τ_c	coil outlet	3 min	A
c	coil heat capacity	3.0 kJ/(kg K)	V
η_F	furnace efficiency (fuel ledger)	0.85	H
θ_f, β, ϕ_B	film, coke on skin, pass B flux	55 K, 8 K/mm, 1.06	V
k_f	coil coking at design film	0.267 mm/day	V
$T_{skin,max}$	skin limit	650 °C	A
—	coil day at start, compression	2 days, 1×	A
C_P, Q_c, Δp_ℓ	overhead, compressor, drum line	0.645 t/bar, 17.3 t/h, 0.7 bar	V
τ_N, τ_L	naphtha and LCGO cut lags	4, 8 min	V
τ_R	HCGO and recycle cut lag	10 min	A
τ_B, τ_{fz}, τ_{in}, τ_v	bottoms, flash zone, charge, valves	15, 5, 10, 0.5 min	H
—	analyzer period, sample line	10 min, 5 min	A
—	feedforward model error in UA_d	0 %	A

no trays, pumparounds or wash-bed coking, and the condensable vapor condenses at once. Coil coke raises only the skin, without the pressure drop it adds. Every size, yield, rate, alarm and limit is a teaching value, and the real skin limit comes from the coil alloy's design metal temperature under API Standard 530.

11.4 The Control Layers as Implemented

The sequencer owns the drums and no controller moves the switch. A global Manual and Auto switch is the master. In Manual every station drives its actuator and every loop tracks it, as Section 1.6.2 describes, so the return to Auto is bumpless. Each execution, every 6 s, builds the measurement packet, steps the feedforward's model, runs the MPC when it is due, then the loops, and writes the actuator commands.

11.4.1 Five PI Loops

All five loops run the velocity-form routine of Section 1.6.2 with the derivative term set to zero, written in this chapter's symbols in Equation (11.7).

$$\begin{aligned} e_k &= a(\mathrm{SP}_k - \mathrm{PV}_k), \qquad & S_k &= S_{k-1} + K_c(e_k - e_{k-1}) \\ \hat{v}_k &= v_0 + S_k + I_{k-1}, & v_k &= \min(v_{max}, \max(v_{min}, \hat{v}_k)) \\ I_k &= I_{k-1} + \left[(K_c/\tau_I)\, e_k + (v_k - \hat{v}_k)/\tau_I\right]\Delta t \end{aligned} \tag{11.7}$$

where $a = \pm 1$ is the direction, S the unclamped running sum, v_0 the output at the start and the back-calculation tracking time equals τ_I. In the symbols of Equations (1.18) and (1.19), v_0 is u_b, $\hat{v}_k$ is v_k and v_k is u_k. Setting an integral time at or above 600 min leaves a loop with proportional action alone. The COT loops ($a = +1$) move each pass's duty at 0.06 MW/K and 5 min within 0–13 MW. The pressure loop ($a = -1$) moves the compressor at 20 % per bar and 20 min within 0–100 %. The overhead loop ($a = -1$) moves the top reflux at 0.6 t/h per K and 6 min. The HCGO loop ($a = -1$) moves the recycle at 0.04 % of the fresh feed per K and 15 min within 0–30 %, and it executes only

when the analyzer reports, with Δt the time since the last report, holding while the analyzer is stale. Antifoam runs either at a fixed ratio to the charge, 3 ppm as shipped, or on foam level, where a PI ($a = +1$) holds the measured outage at 5 m at 2 ppm per m and 20 min within 0–25 ppm. The LCGO draw has no loop.

11.4.2 Schedule Feedforward

The feedforward carries its own copy of the drum constants, as commissioned, and predicts the vapor and gas the fractionator will get from the measurements and the schedule, as shown in Equation (11.8).

$$\begin{aligned} \hat{V}_f &= \hat{V}_{gen} - \hat{d} + (1 - y_{c,p})\frac{\hat{M}_p}{\tau_p} + \hat{V}_{steam} - \hat{m}_w \\ \hat{G} &= \hat{g}\,(\hat{V}_f + \hat{m}_w), \quad \psi_G = \hat{G}/G_0, \quad \psi_V = \hat{V}_f/V_0 \\ \psi_L &= \psi_V \left(n_0/\hat{n}_v\right)^{1/0.45}, \quad u = \psi_G\, u_{PI}, \quad L_r = \psi_L L_{r,PI} \\ R &= \psi_V (F_{ref}/F_f)\, R_{PI}, \quad L = \psi_V L_{set} \end{aligned} \tag{11.8}$$

where $\hat{V}_{gen}$ and $\hat{g}$ come from the yield correlation at the measured coil outlets, charge and fresh feed and at the assay lagged by its 15 min transport, $\hat{d}$ uses the measured live shell, $\hat{M}_p$ is the model's own pitch, $\hat{m}_w$ uses the off-line shell during the warm-up, $\hat{n}_v = \max(0.03,\ n_0 - (\hat{g}_F - g_0))$ corrects the reflux for the gas share, G_0, V_0 and g_0 are design values and $F_{ref} = 120\,\mathrm{t/h}$. The measured inputs are filtered over 1 min, while the schedule's steps pass at once. Each PI loop keeps running on its own error, with its output limit divided by its multiplier, and turning the feedforward on or off re-bases the loops so the actuators do not jump. In the last 2 h of a fill it adds 4 ppm of antifoam. A parameter scales the model's UA_d to give the feedforward a known model error, zero as shipped.

11.4.3 Model Predictive Control of the Recycle and the COT

The MPC, the option of Section 1.6.3 built here on top of the feedforward, plans the recycle base u_1 (the HCGO loop's output, % of the fresh feed at the design vapor) and the COT setpoint u_2 every 2 min over 120 steps (4 h) in graduated blocks. Its model is relinearized at each solve from the measurements, as shown in Equation (11.9).

$$\begin{aligned} \tau_R \dot{x}_1 &= k_r x_3 + b - x_1, \quad 6\dot{x}_2 = u_2 - x_2, \quad \dot{x}_3 = u_1 - x_3 \\ y_2 &= y_{c0} - k_T(x_2 - \mathrm{COT}_0) + k_{rc}(x_3 - r_0) \\ y_3 &= s_0 + k_s(x_2 - \mathrm{COT}_0) \\ y_4 &= \ell_0 + k_\ell(x_2 - \mathrm{COT}_0) \\ \min\ & \textstyle\sum_k w_{95}(x_{1,k} - T_{95,SP})^2 + w_c(y_{2,k} - y_{c,SP})^2 \\ & \quad + \Delta u_k^\top S\, \Delta u_k \\ \text{s.t. } & y_3 \le T_{skin,max} - 5 + \sigma_1, \ \ y_4 \le \bar{u}_c - 2 + \sigma_2 \\ & 0 \le u_1 \le 30, \ \ 480 \le u_2 \le 510 \\ & |\Delta u_1| \le 2, \ \ |\Delta u_2| \le 3 \end{aligned} \tag{11.9}$$

where x_1 is the HCGO 95 % point with the analyzer bias, moved halfway toward each new analyzer result, compared with what the model held at the instant that sample was taken, x_2 the COT, x_3 the delivered recycle, k_r the slope of the curve at the cut, b the term that places the steady value at the present 95 % point, y_2 the coke yield in percent of the fresh feed with the assay as a measured disturbance, y_3 the hotter skin, y_4 the compressor load against its available capacity $\bar{u}_c$, y_{c0}, s_0 and ℓ_0 the present coke yield, skin and load at the present COT COT_0 and recycle r_0, k_{rc}, k_s and k_ℓ their slopes, $w_{95} = w_c = 1$, $y_{c,SP} = 30.5\,\%$, $S = \mathrm{diag}(4, 2)$, and the slacks σ are priced $10^3\sigma^2 + 10\sigma$. The framework's linear MPC module solves each quadratic program by the Goldfarb–Idnani dual algorithm [17], as industrial MPC does [15], and a solution that is infeasible, non-finite or out of range holds the last decisions. Over the challenge's two days it solves 1440 times with no fallback.

Table 11.3 The Nine Cases

#	Case and events	Teaching point
1	Manual, drum A 3 h into its fill	one fill, the warm-up, the switch
2	Loops, drum B warming, switch by hand	the switch transient
3	COT 495 to 505 °C	yield against coil coking
4	Loops, then schedule FF, switch at 2 h	feedforward from a schedule
5	Residue foaming 3×, fixed 3 ppm	antifoam on foam level
6	CCR 22 wt %, foaming 1.6×	the cycle must shorten
7	Coil clock 100×, from day 2	skin limit, run length
8	Switch warmed at 2 h, warm-up skipped, cold at 18 h	what only the warm-up does
9	Challenge, 48 h: cold switch at 28 h, CCR 20 at 20 h, compressor 80 % at 30 h	the whole unit

11.4.4 Three Models and Six Design Choices

The plant is Section 11.3. A second, identified model is the one a reader builds from the trends, for instance the 6 min overhead or the 56 min shell. The controllers' models are the feedforward's commissioned copy of the drum and the MPC's linearization, neither of which sees the pitch, the true foam or the coil coke.

Six choices depart from the app's original specification and are design decisions of the simulator. First, the HCGO 95 % point is held by the wash, the recycle, with the HCGO draw as the balancing stream. On one boiling curve the 95 % point is the HCGO and recycle cut, so the specification's two loops, one on the HCGO draw and one on the recycle ratio, would act on the same cut point, since each draw moves every cut below it (Section 1.4.3), the interaction that Chapter 4 measures on a crude tower. Second, the challenge holds three switches, the second onto a cold drum, and judges the pressure over the whole run, because with warmed switches only the shipped loops already stay within 0.1 bar (a 0.060 bar dip in case 8). Third, only the warm-up's condensate is taken from the fractionator. Subtracting the whole draw made the start of the warm-up a larger step in gas than the switch itself. Fourth, there is no pass flow balancing, so pass B's higher flux makes it the pass that ends the furnace run. Fifth, the challenge judges the HCGO 95 % point within ±5 K at every instant. An earlier rule asked for the band over 95 % of the run, and a pressure PI alone at 100 % per bar passed it (Section 11.5). Sixth, the foam has no ceiling, the documented simplification above.

11.5 Nine Cases, Measured

The cases appear in Table 11.3 and the principal measurements in Table 11.4. Every number comes from the engine run without a browser at the default seed with the browser's own step, any change a case asks for is made after the run begins unless the text says Reset, and seeds 1 to 3 and a halved step were used to check the challenge. Pressures are the logged transmitter values unless called true.

11.5.1 One Fill and One Switch

Case 1 holds every station still in Manual from 3 h into drum A's fill. The charge leaves 36.6 t/h in the drum, 30.5 % of the fresh feed, and sends 95.7 t/h of vapor on, so the bed climbs steadily from 3.4 to 18.0 m by the switch at 13.0 h, with the foam riding on it at 3.0 m and the outage shrinking from 19.6 to 5.0 m. The warm-up from 9 h takes the fractionator's vapor from 95.46 to 89.98 t/h, by its condensate, and lightens the HCGO to 475.7 °C with the recycle fixed. The pressure barely moves at the warm-up start (0.982 to 1.006 bar g) because the condensate carries no gas, and at the switch, with no loop, it falls to 0.907 bar g.

In case 2 the loops run without the schedule. Switched by hand at 10 min, with the new shell at 325 °C after 3 h 10 min of its warm-up, the drum holds back 6.73 t/h, the pressure falls to 0.923 bar g after 8.8 min, the overhead to 126.2 °C, and the pressure is back above 0.99 bar g for good 39 min after the switch, then overshoots to 1.027 bar g at 72 min as the pitch cracks back. Left to the schedule, the shell reaches 350.5 °C and the dip is 0.941 bar g with 5.37 t/h held back.

Table 11.4 Case Outcomes Measured on the Engine

Case and condition	Result
1, the fill	bed 3.4 to 18.0 m, outage 19.6 to 5.0 m
1, the switch at 13.0 h	P to 0.907 bar g, no loop
2, by hand at 10 min (325 °C)	$\dot{d}$ 6.73 t/h, P 0.923 bar g
2, scheduled at 1 h (350 °C)	$\dot{d}$ 5.37 t/h, P 0.941 bar g
3, COT +10 K	coke 30.60 to 28.72 %, coking ×2.23/2.25
4, loops	P 0.940 bar g at +9 min
4, schedule FF	P 0.988–1.014 bar g
5, fixed 3 ppm	FOAM-OVER at 11.2 h, 4.4 kg
5, fixed 12 ppm	outage ≥ 4.42 m, 20.6 kg to the switch
5, on foam level	outage ≥ 4.39 m, 6.2 kg
6, cycle 16 / 15 / 14 h	outage 2.97 / 3.74 / 4.26 m
6, cycle 14 / 12 h	new shell 268 / 151 °C
7, COT 495 / 490 / 485 °C	skin limit on day 40.0 / 63.7 / 101
8, warmed / cold, loops	dip 0.060 / 0.178 bar, 4.9 / 15.4 t held
8, cold, schedule FF	dip 0.024 bar, T_{95} 479.4–480.9 °C

11.5.2 What the Schedule Is Worth

Case 4 runs the same scheduled switch twice. With the loops, the pressure falls to 0.940 bar g 9 min after the switch while the compressor reaches its lowest, 73.4 %, only after 17 min. With the schedule feedforward the compressor steps with the switch, from 75.0 to 74.3 % in the first 12 s and 70.8 % at 2 min, and the logged pressure stays within 0.988–1.014 bar g, 0.993–1.003 bar g in truth against the transmitter's 0.004 bar noise. Case 8 puts a warmed switch and a cold one on one axis, as shown in Figure 11.2. The cold shell, at 137 °C, holds back 16.9 t/h against 5.38 t/h for the warmed one and 15.4 t against 4.9 t over four hours, 3.1 times as much, and with the loops the dip is 0.178 bar against 0.060 bar, while the true HCGO 95 % point swings from 471.1 to 484.9 °C. The feedforward cuts the cold dip to 0.024 bar and holds the 95 % point within 479.4–480.9 °C, but the vapor lost to the fractionator does not change, because the heat to warm 60 MJ/K of steel by 300 K must come from the charge.

🕮 **Pause.** Try Exercise 11.2 before reading on.

The feedforward is only as good as its copy of UA_d. With that copy 20 % low the cold switch of case 8 dips to 0.959 bar g, with it 20 % high the pressure overshoots to 1.056 bar g, against 0.976–1.018 bar g with the exact copy, and the challenge still passes at 0.040 and 0.052 bar with the 95 % point within 1.5 and 3.4 K. At 50 % low it passes at 0.095 bar with the 95 % point 4.6 K off, so close to its band that seed 1 fails it at 5.3 K, and at 50 % high it fails at 0.121 bar with the 95 % point 7.3 K off. The degradation is graceful and roughly in proportion to the error, and an error that overpredicts the held-back vapor is the more harmful one because it pulls the compressor down further than the gas falls.

11.5.3 Coke Against Tube Life, Foam and Heavier Feed

Raising the COT to 505 °C in case 3 lowers the coke from 30.60 to 28.72 % of the fresh feed and raises the liquid from 58.71 to 59.72 % and the gas from 10.83 to 11.64 %, which loads the compressor from 75.1 to 80.2 %. The duty rises by 1.03 MW, the hotter skin from 558.1 to 571.6 °C, and the coil coking rate by 2.23 and 2.25 times in passes A and B, so the run left falls from 38.1 to 14.5 days. Case 7 follows that trade to its end at 100 coil days per simulated day. Pass B reaches 650 °C on coil day 40.0 at 495 °C, day 63.7 at 490 and day 101 at 485, and the cooler run costs coke, 31.55 against 30.61 % at the start.

In case 5 a residue that foams three times as much stands 9.05 m of foam on the fixed 3 ppm, and as the bed rises under it the outage reaches 1 m at 11.2 h, 14.2 h into the fill, after 4.44 kg of antifoam. A fixed ratio that reaches the switch needs 6 ppm for 2.01 m of outage, under the FOAM HIGH line, and 12 ppm for 4.42 m at 20.6 kg. On foam level the loop doses 0.24 ppm at 1 h, 3.49 at 8 h and 12.9

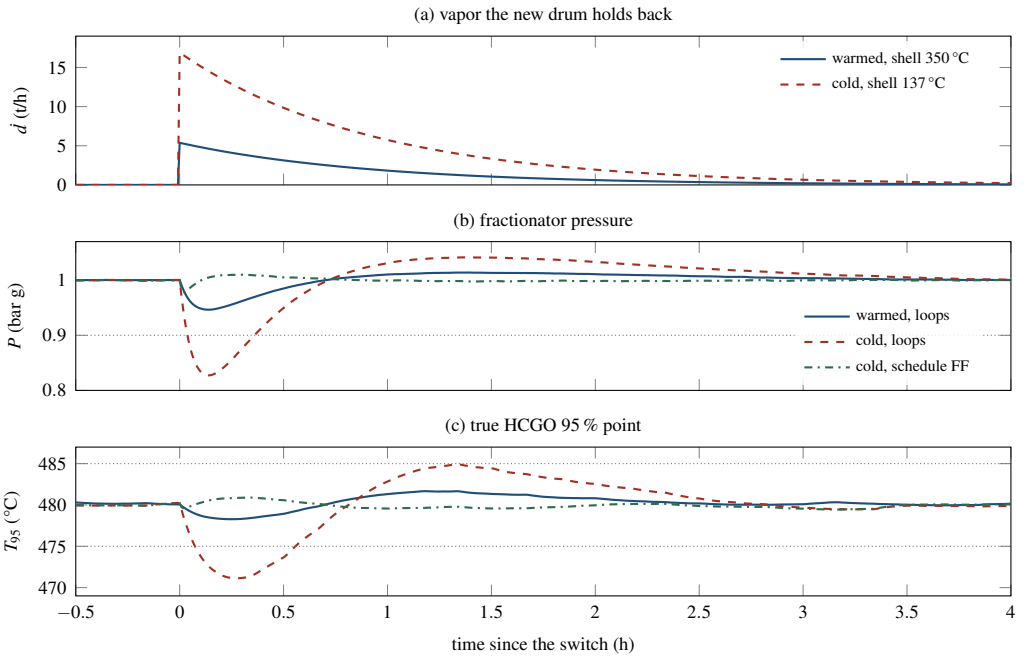


Figure 11.2 The two switches of case 8 on a common axis of time since the switch, true values every 30 s. The warmed switch is drawn solid, the cold switch on the loops dashed and the cold switch with the schedule feedforward dash-dotted. (a) The cold shell holds back 3.1 times the vapor. (b) With the loops the cold switch dips 0.17 bar in truth, and the schedule feedforward keeps it within 0.978–1.010 bar g. The gray dotted line is the challenge's lower band. (c) The HCGO 95 % point leaves its ±5 K band, the gray dotted lines, with the loops and not with the feedforward.

just before the switch, holds the outage within 4.39–5.44 m and uses 6.2 kg. Early in the fill the same loop lets the foam stand 16.1 m tall, and 17.9 m on the new drum after the switch, the consequence of a foam law with no ceiling. In case 6 a residue at 22 wt % CCR leaves 44.7 t/h of coke, the bed rises 1.37 m/h, and on the 16 h cycle the outage falls below 4 m 14.6 h into the fill and reaches 2.97 m at the switch with the antifoam at its 25 ppm limit. Shortening the cycle to 14 h keeps 4.26 m, but the shorter warm-up puts the next drum on line at 268 °C instead of 339 °C, and at 12 h at 151 °C. With 26 wt % at 135 t/h no cycle the decoking allows keeps the outage above 4 m (2.99 m at 12 h), so past that floor only the feed rate is left.

🕮 **Pause.** Try Exercise 11.3 before reading on.

11.5.4 Two Days on the Coker

The challenge runs 48 h from coil day 30, with switches at 12, 28 and 44 h, the second onto a cold drum because its warm-up is skipped at 23 h, a residue at 20 wt % CCR and foaming 1.5 times from 20 h, and the compressor limited to 80 % from 30 h. It is met with the true HCGO 95 % point within ±5 K and the true pressure within ±0.1 bar throughout, no foam-over, the skin under 650 °C and at most 15 kg of antifoam, judged on the truth at every step. The 95 % point is graded on the simulator's value rather than on the analyzer's reports, which arrive 15 to 25 min late. The configurations are compared in Table 11.5, and Figure 11.4 draws three of its columns against their limits. As shipped the pressure leaves its band first, at the cold switch at 28.0 h, the 95 % point 1.9 min later, and the fixed 3 ppm passes the budget at 38.0 h. The feedforward and antifoam on foam level each fix one of those failures, and together they pass at all four seeds and with the step halved (0.024–0.025 bar,

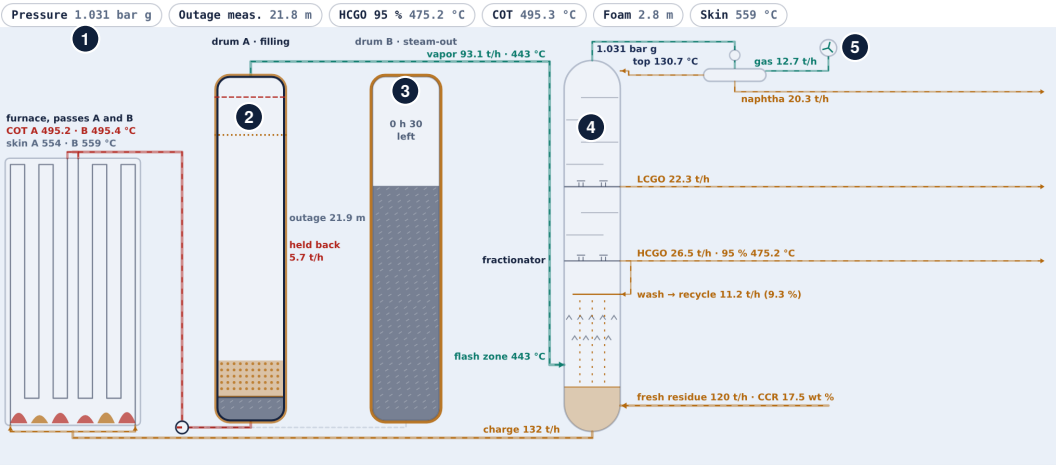


Figure 11.3 Case 8 on the loops alone at $t = 19$ h, one hour after the switch onto the cold drum, on the Unit view, cropped to the plant with numbered callouts. The badges (1) read 1.031 bar g as the pressure overshoots on the pitch cracking back, 21.8 m of measured outage, a COT of 495.3 °C, 2.8 m of foam and a 559 °C skin, and the HCGO analyzer reports 475.2 °C. Drum A (2) is filling, its shell still holding back 5.7 t/h, and drum B (3), full, is in steam-out with 30 min left. The vapor reaches the fractionator (4) at 93.1 t/h and 443 °C, the wash returns 11.2 t/h, and the wet gas compressor (5) takes the gas from the overhead.

Table 11.5 Challenge Configurations Scored at 48 h

Configuration	P worst (bar)	T_{95} in band	T_{95} off (K)	Antifoam (kg)	Verdict
shipped (loops, 3 ppm)	0.186	98.7 %	9.7	18.9	fails
loops, on foam level	0.186	98.7 %	9.7	6.0	fails
schedule FF, 3 ppm	0.024	100 %	1.0	22.1	fails
schedule FF, 2 ppm	0.024	100 %	1.0	15.8	fails
PI 70 %/bar, 6 min, foam level	0.103	98.2 %	9.9	6.0	fails
PI 100 %/bar, 10 min, foam level	0.089	98.2 %	9.9	6.0	fails
suggested (FF, foam level)	0.024	100 %	1.0	6.8	passes
MPC, foam level	0.049	100 %	1.2	6.1	passes

the 95 % point within 1.0–1.3 K, 6.81–6.84 kg). The MPC with antifoam on foam level also passes, raising the COT as high as 508.4 °C to win back liquid on the heavier residue until the skins, 646 °C against its 645 °C soft limit, and the compressor stop it.

> ✦ DID YOU KNOW? **The Tightest Pressure Still Fails**
>
> From the simulator (Table 11.5). On a fixed antifoam ratio of 3 ppm, the schedule feedforward holds the pressure within 0.024 bar and the heavy gas oil cut within 1.0 K, yet it fails the challenge. It spends 22.1 kg of antifoam against a budget of 15 kg. On foam level the same feedforward uses 6.8 kg and passes.

The pressure band by itself does not need the feedforward. Without it a pressure PI at 100 % per bar holds 0.077, 0.089 and 0.092 bar at integral times of 2, 10 and 20 min, and the 100 % per bar, 10 min setting keeps the pressure inside its band at all four seeds and with the step halved (0.089–0.090 bar). That setting still fails the challenge, on the heavy gas oil. It leaves the true HCGO 95 % point 9.6–10.5 K off at the cold switch over the four seeds, against 1.0–1.3 K for the suggested settings and 1.2–1.6 K for the MPC. The excursion lasts under 2 % of the run, so the earlier rule,

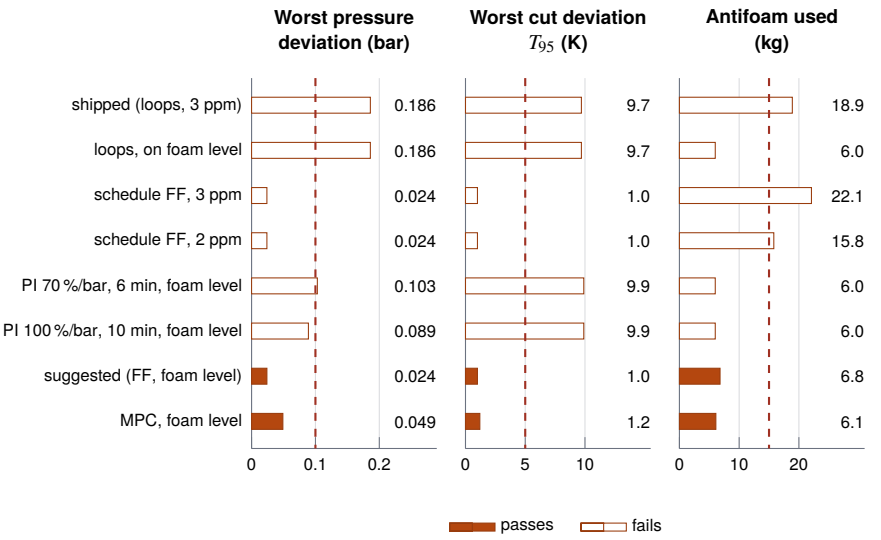


Figure 11.4 Three columns of Table 11.5 against the challenge's limits (dashed): the worst pressure deviation against 0.1 bar, the worst deviation of the true HCGO 95 % point against 5 K and the antifoam against 15 kg. Solid bars pass the challenge, white bars fail it. Notice that the pressure PI at 100 % per bar reaches the pressure band but not the cut, and that the feedforward on a fixed ratio fails only on antifoam.

the band over 95 % of the run, passed it, and only grading the cut at every instant shows what the feedforward adds. Retuning the HCGO PI does not rescue the loops. With the pressure PI at 100 % per bar, gains of 0.02 to 0.5 % per K and integral times of 5 to 30 min leave a worst deviation of 9.7 K at best, and the settings from 0.2 % per K, or with a 5 min integral time from 0.04 % per K, swing 35 to 116 K off on reports that arrive 15 to 25 min late. The PI alone also moves the compressor 13 041 % of its range over the two days against 2930 % for the feedforward, 4.5 times as far. That second advantage belongs to the transmitter's noise. With the noise switched off the PI at 100 % per bar travels 48 % and the feedforward 1149 %, so the high-gain loop's travel is noise passed to the compressor, while the feedforward's is the schedule it follows.

11.6 Technology in Practice

Figure 11.5 draws the loops of Section 11.4 with the hardware a refinery installs.

What Measures It

The feedforward rests on the drum's skin thermocouples. In refiners' practice they sit around the skirt and the lowest circumferential welds, the operators read them beside the quench water flow and the drum pressure, and strain gauges in the axial and hoop directions next to them show the stress in the wall, which is largest during the water quench [18]. Skin temperatures alone do not give what the gauges give, but they help ensure an adequate warm-up [18]. The 2 min lag and 1 K of noise of Section 11.3 stand for such a thermocouple. The heavy gas oil cut comes from an online distillation analyzer that reports every 10 min behind a 5 min sample line, and the outage comes from detectors that see through the drum's wall.

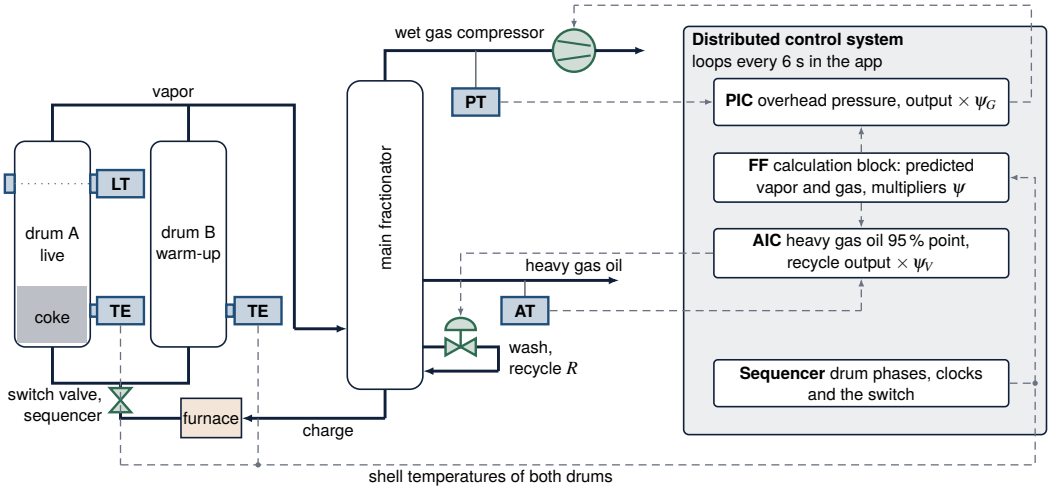


Figure 11.5 The fractionator's loops and the schedule feedforward with their hardware. Lettered boxes are sensing devices: the overhead pressure transmitter (PT), the online distillation analyzer on the heavy gas oil (AT), the drum shell thermocouples (TE) and the radiation level detectors across the live drum (LT). Valve symbols and the compressor are final elements, and dashed arrows are signals. Notice that the feedforward reads the sequencer and the shells, not the fractionator, and writes only multipliers on the outputs of the loops beside it. The coil outlet loops, the overhead temperature loop and the antifoam are left out. Not to scale, and not a piping diagram.

TECHNOLOGY **Sensor: Radiation Level Detectors on the Drum**

Gamma rays pass through a vessel's wall and are weakened by the material inside, so a source on one side and a detector on the other give an approximate level. The detector counts pulses in a gas-filled tube. A cobalt source has a half-life of 5.3 years and a cesium source 30 years, so cesium drifts less. Accuracy varies with the density of the contents, coating on the wall, source decay and detector drift, and such gauges are used where no other instrument is practical, delayed coking vessels among them, since the coke is cut out by a high-pressure water jet [19]. In a pilot coker a scanned gamma densitometer told gas, foam, liquid and coke apart by density [14]. The chapter's detectors read each drum's foam top with a 1 min lag and 0.15 m of noise.

What Moves It

The wet gas compressor is the handle of the pressure loop and of the feedforward. The coil outlets are held by the firing of each pass, the recycle by the wash and the foam by an antifoam injection pump. Foam is the reason the pump exists: in a pilot coker, runs without antifoam foamed over, plugged the overhead line and ended the run [14], the event the latched FOAM-OVER of Section 11.3 stands for.

TECHNOLOGY Final Element: Wet Gas Compressor Speed

A compressor's gas flow can be set by its speed instead of by a valve. Varying the speed uses less energy than throttling the discharge or recycling gas to the suction, but a variable-speed machine often cannot change flow as fast as a control valve and has limited options for a fail-safe state. A centrifugal compressor surges, violently, when the pressure across it is too high and its flow too low, and a recycle valve from discharge to suction prevents it [19]. One refiner answers a pressure drop after a drum switch by raising the compressor's suction pressure setpoint, which lowers its speed [20]. The chapter's compressor is the fraction u of its capacity in Equation (11.6), with no surge term.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level oversees loops that are sometimes nested, with cycle times from milliseconds to minutes [21]. The app executes its loops every 6 s and its MPC every 2 min, inside that range. The five loops are PI blocks. The sequencer is sequence logic that steps each drum through its phases and moves the switch valve, and the feedforward is a calculation block that multiplies the outputs of the compressor, reflux and recycle loops, as Equation (11.8) writes it. The sequencer's phases and clocks are known exactly because the control system runs them, and that is what lets the feedforward act at the switch rather than after it.

IN PRACTICE Warm the Drum, Then Switch

Refiners describe the switch as this chapter models it. In a 2011 panel, one refiner explained that a switch onto a relatively cold drum lowers the system pressure until the new drum reaches its normal overhead temperature [20]. Its practice was to watch the drum's skin temperatures to avoid a switch to a cold drum, to start the antifoam about four hours before the switch, and, if the pressure still fell, to raise the compressor's suction pressure setpoint. In a 2010 panel a refiner set the minimum preheat before a switch from strain studies, and an engineering contractor described how limits on the skin temperature's rate of change can enter the warm-up and switch controls of the distributed control system [18].

What Is Modernizing It

Three developments touch this unit. The first is drum life. Strain gauges at high temperature and laser inspection of bulges join the internal inspection that is made only at major turnarounds [22], and one refiner recommends strain gauges on new drums and fits them to old ones [18]. The second is a view into the drum. A scanned gamma densitometer in a pilot coker detected foam forming and collapsing, by itself and under antifoam, and classified the coke morphology [14], where the chapter's detectors read one foam top. The third is the cut between analyzer reports. Predicting a quality that is sampled slowly from fast measurements is the original and still dominant use of soft sensors [23], and model predictive control, with more than 4600 reported applications, is most common in refining [15]. The chapter's MPC moves its analyzer bias only when a result arrives.

11.7 Limitations

Every size, yield, rate and limit is a teaching value fitted to no operating coker. The yield correlation has the published form on the CCR, with teaching slopes for the COT and the drum pressure, and the pseudo boiling curve is the app's own. A drum is lumped, so the bed's structure, the liquid inside it and the drum's thermal fatigue, the real reason a cycle is not shortened at will [10], are absent, and the new drum's deficit is written as cracking deferred rather than condensation on the walls. The foam has no ceiling and stands 16 m to 18 m tall on case 5's residue while the outage loop doses little, a height a real unit would not tolerate because the foam's top would move with every change of vapor or pressure. The fractionator has no trays, pumparounds or wash-bed coking, and its pressure model condenses all condensable vapor at once. A shell that warms at exactly one declared rate and a switch on the stroke of the schedule would not occur in a real unit, but each isolates the one mechanism its case is about.

The control layers carry simplifications of their own. There is no pass flow balancing, so the run length is set by pass B. The feedforward's model is exact as shipped, and only its UA_d can be given an error. A high-gain pressure PI alone meets the challenge's pressure band, so the grade that calls for the feedforward is the HCGO cut, and it is taken on the simulator's true 95 % point, which no instrument in a real unit reports without the analyzer's delay. The MPC has no run-length term and spends the coils for yield up to its soft skin limit. The detector outage differs from the truth by up to 0.68 m, mostly the detectors' noise.

OPEN CHALLENGE **Drums Not Built for Shorter Cycles**

The model's drum has no wall stress, yet the wall limits the cycle. Coke drums were designed to ASME Section VIII, Division 1, and not at first for cyclic loads. As some plants cut the cycle from 24 to 12 hours or less, through-wall cracks came more often and earlier [22]. Hot and cold spots in the wall raise bulges, and the internal inspection that finds them happens only at major turnarounds, about every four years. Most drum failures come from high cyclic thermal stresses, and they grow more frequent as throughput rises and cycles shorten [10, 22]. Progress would be a sequencer that plans the warm-up, the switch and the quench against the measured wall stress.

11.8 Conclusion

A delayed coker hosts a batch step inside a continuous plant, and this chapter put the drum cycle, the furnace coils and the fractionator's loops on one clock. A switch onto a drum left cold holds back 16.9 t/h of vapor, 3.1 times a warmed one, and dips the fractionator pressure 0.178 bar on the loops alone. A feedforward built from the schedule and the shell thermocouples keeps the dip to 0.024 bar and the HCGO 95 % point within 1 K, degrades gracefully with a model error of 20 %, and cannot make the vapor the steel takes, which only the warm-up saves.

Ten degrees of COT trade 1.9 points of coke for coils that coke 2.2 times as fast, antifoam on the drum's outage uses 6.2 kg where a safe fixed ratio uses 20.6 kg, and a heavier residue shortens the cycle until the decoking floor of 12 h leaves only the feed rate.

The drum and foam models are lumped and declared rather than fitted, and a high-gain PI also holds the graded pressure band, so only the HCGO cut, graded at every instant, separates it from the feedforward. A fuller model would add a foam ceiling, wall condensation and thermal stress in the drum, and pass flow balancing. The coker is a plant on which to practice feedforward from a known schedule.

KEY POINTS

- A delayed coker cracks residue in one of two drums that take turns, and each switch sends the charge into a drum colder than its vapor.
- A switch onto a cold drum holds back 16.9 t/h of vapor, 3.1 times a warmed one, and dips the pressure 0.178 bar on the loops alone.
- A schedule feedforward keeps the dip to 0.024 bar but cannot make the vapor the steel takes, which only the warm-up saves.
- A high-gain pressure PI also holds the pressure band, so only the heavy gas oil cut, graded at every instant, separates it from the feedforward.

Where else this applies. The backwash sequencer of Volume 3, Chapter 6 and the diversion sequence of Volume 8, Chapter 13 are batch steps on a known schedule beside a continuous plant.

References

[1] E. Worrell and C. Galitsky, "Energy efficiency improvement and cost saving opportunities for petroleum refineries: An ENERGY STAR guide for energy and plant managers," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-56183, 2005.

[2] M. Baldea and I. Harjunkoski, “Integrated production scheduling and process control: A systematic review,” *Computers & Chemical Engineering*, vol. 71, pp. 377–390, 2014.

[3] A. N. Sawarkar, A. B. Pandit, S. D. Samant, and J. B. Joshi, “Petroleum residue upgrading via delayed coking: A review,” *The Canadian Journal of Chemical Engineering*, vol. 85, no. 1, pp. 1–24, 2007.

[4] K. Wisecarver, “Delayed coking,” in *Springer Handbook of Petroleum Technology*, ser. Springer Handbooks. Cham: Springer International Publishing, 2017, pp. 903–913.

[5] R. Dittmar and G. Reinig, “Model predictive FCCU main fractionator control – robustness and adaptation aspects,” *IFAC Proceedings Volumes*, vol. 30, no. 6, pp. 1201–1206, 1997.

[6] J. Yang, G. Shen, P. Li, W. Wei, C. Zhang, and J. Xiao, “Kinetic model for the deep-severity thermal reaction in the coke drum of delayed coking,” *Energy & Fuels*, vol. 31, no. 3, pp. 2681–2686, 2017.

[7] X.-L. Zhou, X. Di, G.-X. Yu, R.-X. Lu, and C.-L. Li, “Simulation of delayed coking reaction in coke drum,” *Petroleum Science and Technology*, vol. 28, no. 3, pp. 277–285, 2010.

[8] A. C. Silva, C. McGreavy, and M. F. Sugaya, “Coke bed structure in a delayed coker,” *Carbon*, vol. 38, no. 15, pp. 2061–2068, 2000.

[9] P. E. Eaton and J. Williams, “Delayed coker furnace fouling control – laboratory correlation to field experience,” in *CORROSION 2008*. New Orleans, LA: NACE International, 2008, pp. 1–7.

[10] K. Kirkpatrick, L. Antalffy, B. Millet, and G. Miller, “Extending drum life while increasing throughput in a delayed coking unit,” in *ASME 2017 Pressure Vessels and Piping Conference, Volume 3A: Design and Analysis*. Waikoloa, Hawaii, USA: American Society of Mechanical Engineers, 2017, paper V03AT03A006.

[11] Occupational Safety and Health Administration, “OSHA technical manual, section IV, chapter 2: Petroleum refining process,” U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[12] J. H. Gary and G. E. Handwerk, *Petroleum Refining*. CRC Press, 2001.

[13] J. G. Speight, *The Chemistry and Technology of Petroleum*. CRC Press, 2014.

[14] M. Volk, Jr., K. D. Wisecarver, and C. M. Sheppard, “Fundamentals of delayed coking joint industry project,” Office of Scientific and Technical Information (OSTI), Tech. Rep., 2003.

[15] S. J. Qin and T. A. Badgwell, “A survey of industrial model predictive control technology,” *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[16] D. E. Seborg, T. F. Edgar, D. A. Mellichamp, and F. J. Doyle, III, *Process Dynamics and Control*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2016.

[17] D. Goldfarb and A. Idnani, “A numerically stable dual method for solving strictly convex quadratic programs,” *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[18] American Fuel & Petrochemical Manufacturers, “Question 53: During the coking cycle, how are drum skin temperatures used to monitor drum wall condition?” AFPM Q&A Database, panel answers by two refiners and an engineering contractor, 2010, afpm.org, question and answer search.

[19] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[20] American Fuel & Petrochemical Manufacturers, “Question 51: What are some steps that can be taken to minimize re-foaming while steaming to the fractionator?” AFPM Q&A Database, panel answers by two refiners and an engineering contractor, 2011, afpm.org, question and answer search.

[21] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, “Guide to operational technology (OT) security,” National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[22] J. Chen, “Fatigue behaviour of coke drum materials and its application to extend the service lives of coke drums,” Ph.D. dissertation, University of Alberta, 2014.

[23] P. Kadlec, B. Gabrys, and S. Strandt, “Data-driven soft sensors in the process industry,” *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

11.9 Exercises

11.1 Concept check · 2 min

At the scheduled switch of case 2 the fractionator pressure falls from 1.00 to 0.941 bar g, although the charge, the coil outlet temperature and the compressor's capacity have not changed. What takes the gas away?

(a) The switch valve shuts for a moment, and the charge reaches neither drum

(b) The new drum stands at a lower pressure, and the overhead gas flows back into it

(c) The new drum's cooler shell takes heat from its contents, so less of the charge cracks

(d) The warm-up condensate drains to blowdown and carries part of the wet gas with it

11.2 Concept check · 2 min

On the cold switch of case 8 the schedule feedforward cuts the pressure dip from 0.178 to 0.024 bar. Which quantity does it leave unchanged?

(a) The time at which the wet gas compressor starts to move

(b) The swing of the true HCGO 95 % point after the switch

(c) The top reflux that the overhead loop and its multiplier send

(d) The vapor that the cold drum holds back over the next four hours

11.3 Concept check · 1 min

In case 5 a residue that foams three times as much runs on a fixed antifoam ratio of 3 ppm, and it foams over 14.2 h into a 16 h fill. Why near the end of the fill and not at its start?

(a) The foam stays about 9 m thick while the bed rises under it, so the outage shrinks

(b) The vapor rate rises as the drum fills, so the steady foam grows through the fill

(c) The antifoam is used up over the fill, so the dose falls as the switch comes nearer

(d) The drum pressure rises as the drum fills, and the pressure factor swells the foam

11.4 Theory and calculation · 6 min

The design residue has a Conradson carbon of 17.5 wt%. The fresh feed is 120 t/h, the recycle 12.0 t/h, the coil outlet 495 °C and the drum pressure 1.70 bar g. A drum is 6.8 m across and 26 m to the vapor line, and the coke bed has a density of 0.90 t/m^3.

(a) Use Equation (11.1) to find the coke the charge leaves in the drum, in t/h and as a share of the fresh feed. *Checkpoint:* 36.6 t/h, 30.5 % of the fresh feed.

(b) Find the coke per meter of bed and the rate at which the bed rises. *Checkpoint:* 32.7 t/m and 1.12 m/h.

(c) A fill starts with an empty drum and the foam rides on the bed at 3.0 m. Find the outage at the end of a 16 h fill, and compare it with the FOAM HIGH line at 4 m. *Checkpoint:* 5.1 m, 1.1 m above the FOAM HIGH line.

11.5 Theory and calculation · 7 min

Take the drum constants of Equation (11.2): $C_m = 60\,\mathrm{MJ/K}$, $UA_d = 18\,\mathrm{kW/K}$ and $h_{eff} = 1200\,\mathrm{kJ/kg}$, with the drum vapor at $T_{d0} = \mathrm{COT} - 45\,\mathrm{K}$ and the coil outlet at 495 °C.

(a) Find the vapor held back for each kelvin of shell below the vapor, in t/h per K, and the shell's time constant on line. *Checkpoint:* 0.054 t/h per K and 55.6 min.

(b) In case 8 the warmed drum goes on line with its shell at 350.4 °C and the cold drum at 137.1 °C. Find the vapor each holds back at the switch, and their ratio. *Checkpoint:* 5.38 t/h and 16.9 t/h, a ratio of 3.14.

(c) Let the shell approach the vapor with the time constant of (a) and ignore the pitch that cracks back. Find the total vapor each drum holds back and the part of it held in the first four hours. Show that the total equals $C_m(T_{d0} - T_m)/h_{eff}$. *Checkpoint:* 15.6 t and 5.0 t in total, 15.4 t and 4.9 t in four hours.

11.6 Theory and calculation · 6 min

The overhead pressure follows the balance of Equation (11.6), $C_P\,dP/dt = G - u\,Q_c(P + p_{atm})/(P_0 + p_{atm})$, with $C_P = 0.645\,\text{t/bar}$, $Q_c = 17.3\,\text{t/h}$, $P_0 = 1.0\,\text{bar g}$ and $p_{atm} = 1.013\,\text{bar}$. At design the compressor stands at 75.2 % and the fractionator receives 95.4 t/h of vapor.

(a) Find the wet gas at design and the time constant of the overhead pressure. *Checkpoint:* 13.0 t/h and 6.0 min.

(b) The warmed switch of case 4 holds back 5.37 t/h of vapor with the gas share of the design vapor. Find the fall in wet gas, and the compressor position, from 75.0 %, that holds the pressure at P_0, which is what the multiplier ψ_G of Equation (11.8) does. *Checkpoint:* 0.73 t/h and 70.8 %.

(c) If nothing moved the compressor, how far would the pressure settle below P_0 for that fall in gas? Compare with the 0.907 bar g that case 1 reaches at the same kind of switch with no loop. *Checkpoint:* 0.11 bar, to about 0.89 bar g.

11.7 App challenge · 8 min

Load scenario 4, Pressure at the switch: loops or schedule feedforward. The sequencer switches the drums at 120 min, onto a drum at the end of its warm-up. Use the default seed and the default speed of 60 min per second, and export the CSV file after each run.

(a) Run as shipped (Loops) to 240 min. Find the lowest `pressure_bar_g` and its time, and the lowest `comp_pct` and its time. *Checkpoint:* 0.940 bar g at 129.2 min and 73.4 % at 137.4 min, within 0.002 bar, 0.1 % and 0.2 min.

(b) Choose Loops + schedule FF, press Reset and run to 240 min again. Find the range of `pressure_bar_g`, and the lowest `comp_pct` and its time. *Checkpoint:* 0.988 to 1.014 bar g, and 70.8 % at 122.0 min.

(c) In each run, which moved first, the pressure or the compressor? Say what the feedforward knew at 120 min that the pressure loop did not.

11.8 App challenge · 9 min

Load scenario 8, A skipped warm-up. Drum B goes on line at 2 h after its warm-up, and drum A, whose warm-up is skipped, goes on line cold at 18 h. Use the default seed and a speed of 300 min per second, run to 1260 min and export the CSV file.

(a) Read each new drum's shell temperature from the note the switch posts, and the peak of `truth_deficit_tph` after each switch. *Checkpoint:* 350 °C and 137 °C. 5.38 t/h and 16.90 t/h.

(b) Find the lowest `pressure_bar_g` in the three hours after each switch, and its dip below the setpoint of 1.00 bar g. *Checkpoint:* 0.940 bar g and 0.822 bar g, dips of 0.060 and 0.178 bar, within 0.002 bar.

(c) Choose Loops + schedule FF, press Reset and run to 1260 min again. Find the lowest pressure after the cold switch and the peak of `truth_deficit_tph` after it. *Checkpoint:* 0.976 bar g and 16.90 t/h.

Explain the ratio of the two peaks in (a), and why (c) changes the dip but not the peak.

11.9 App challenge · 9 min

Load scenario 9, Challenge, two days on the coker. Use the default seed and a speed of 600 min per second, so that each 2880 min run takes about five seconds. Read the So far tile when the run stops at 48 h.

(a) Run as shipped. Read the worst HCGO deviation, the worst pressure deviation and the antifoam used, and the verdict. *Checkpoint:* 9.7 °C, 0.186 bar, 18.9 kg. The target is missed.

(b) Choose Loops + schedule FF and set Antifoam to On foam level, press Reset and run the two days again. *Checkpoint:* 1.0 °C, 0.024 bar, 6.8 kg. The target is met.

(c) Use Table 11.5 to say which of the two changes fixes which failure of (a), and name the event at which the pressure and the HCGO left their bands in (a).

11.10 Application area · 4 min

In case 7 pass B reaches the 650 °C skin limit on coil day 40.0 at a coil outlet of 495 °C, and on day 63.7 at 490 °C. At the start of the run the cooler outlet makes 31.55 against 30.61 % coke on 120 t/h of fresh feed.

(a) Treat each run as lasting from a clean coil to the limit. How many decokes a year does each outlet need, and how much more coke a year does the cooler outlet make? *Checkpoint:* 9.1 against 5.7 decokes, 3.4 fewer. 1.13 t/h of coke, about 9900 t a year.

(b) Name what a planner must know to choose between the two, and two things the simulator leaves out of this trade.

11.11 Application area · 3 min

In case 6 a residue at 22 wt% CCR fills the drum faster. On the 16 h cycle the outage reaches 2.97 m at the switch. A 14 h cycle keeps 4.26 m, but the next drum goes on line at 268 °C instead of 339 °C, and a 12 h cycle at 151 °C.

(a) Why does a shorter cycle make a colder switch? Use the 0.054 t/h per kelvin of Exercise 11.5 to find the vapor held back at the switch for the 16, 14 and 12 h cycles. *Checkpoint:* 6.0, 9.8 and 16.1 t/h.

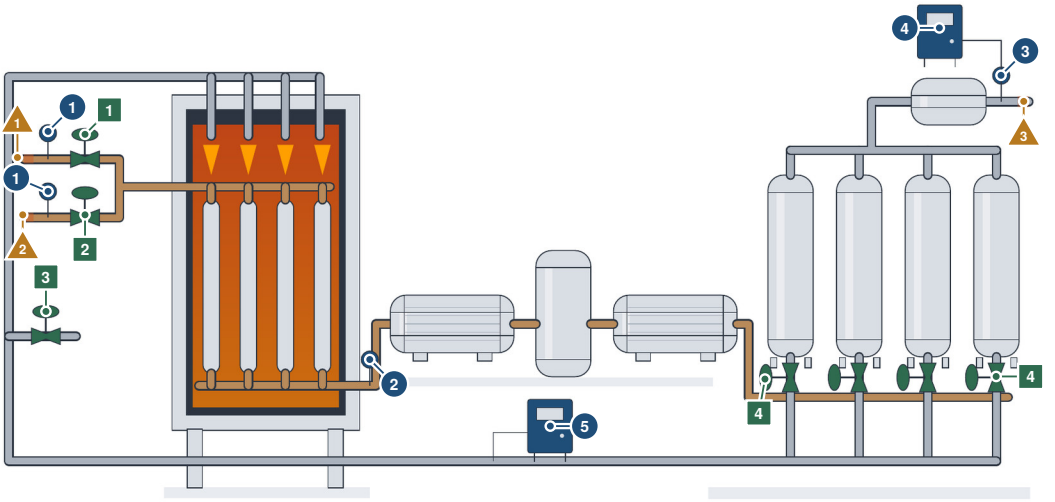
(b) At 26 wt% and 135 t/h no cycle keeps the outage above 4 m. Which handle is left, and what sets the shortest cycle?

11.12 Review the agent · 5 min

An AI assistant reads Table 11.5 and writes: “The schedule feedforward is unnecessary. A pressure PI at 100 % per bar with a 10 min integral time keeps the pressure within 0.089 bar over the two days, inside the 0.1 bar band at all four seeds, and it needs no model of the drum. Use this tuning on the loops and keep antifoam on foam level.” Find the flaw in the proposal and say how you would check it in the app.

CHAPTER 12

A Hydrogen Plant Whose Steam Must Lead Its Feed and Whose Waste Is Its Fuel



The controller moves 1 natural gas feed valve 2 process steam valve 3 makeup fuel valve to the burners 4 adsorber switching valves, stepped on the cycle

The controller reads 1 natural gas and steam flow meters 2 reformer outlet temperature 3 hydrogen header pressure 4 hydrogen purity analyzer 5 heating value of the tail gas

What disturbs it 1 more ethane in the natural gas 2 a change in steam supply pressure 3 the users' hydrogen demand

Notice that the tail gas from the four adsorbers runs back along the bottom to the burners on top of the reformer, so the hydrogen the beds do not recover becomes the furnace's fuel. Illustration, not to scale.

IN PLAIN TERMS Steam First, and the Waste Becomes Fuel

The main idea. A hydrogen plant heats natural gas and steam in tubes filled with a catalyst, a material that speeds up the change into hydrogen. Beds of special grains then hold back the other gases and let clean hydrogen through.

The challenge. The gas must never outrun the steam, or carbon builds up on the catalyst, yet the steam can only speed up slowly. The gases the beds give back are burned to heat the tubes, so keeping more hydrogen means buying more fuel.

Why it matters. The order in which the gas and the steam move matters as much as how much of each the plant uses.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the Did you know? box in Section 12.1, the History boxes in Section 12.2, the Try it boxes and the second Did you know? box in Section 12.5, and Key points in Section 12.9. No equations.
Study	55 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Section 12.5 with Figure 12.3, Section 12.6 and Section 12.7 with its Open challenge box.

The control question A hydrogen plant must keep its slow steam ahead of a fast feed, and it burns whatever hydrogen its adsorber does not recover. Which structure keeps the steam-to-carbon ratio above its floor through a rate change, and how does the adsorber's recovery reach the furnace?

Learning objectives

- Compute the steam a feed needs, the ratio left after a feed step with the steam held, and the valve position that restores it.
- Explain why a plain ratio to the measured feed falls through its floor when the steam ramps slowly, and how a lead-lag on the rate demand prevents it.
- Predict the adsorber's recovery and purity from its step time and its feed with the declared correlation.
- Trace a change in recovery through the tail gas to the makeup fuel and to the natural gas used per unit of hydrogen.
- Configure the header, ratio, adsorber and firing layers to pass the eight-hour challenge, and name the target each change protects.

Prerequisites Sections 1.4.2 (heat-balance feedforward and tube metal), 1.4.5 (heat fed back), 1.4.6 (bias update), 1.6.1 and 1.6.2. Section 1.6.3 is needed only for the predictive controller.

The app https://apmonitor.com/apps/hydrogen-plant/

See also Volume 3, Chapter 4 for cross-limiting fuel and air and an override with external reset feedback. Volume 4, Chapter 3 for limiters selected on one fuel valve. Volume 6, Chapter 7 for a recycle that amplifies a production change. Chapter 14 for a header pressure that nobody schedules.

12.1 Introduction

Hydrogen for a refinery's hydrotreaters and hydrocracker is made, in the generic plant simulated here, from 18 700 Nm^3/h of natural gas and 46.9 t/h of steam, three moles of steam for every mole of carbon. The reformer's tubes leave at 860 °C and 25 bar a with 5.37 % of the dry gas still methane, a shift bed converts most of the carbon monoxide, and a four-bed adsorber delivers 50 004 Nm^3/h of hydrogen at 99.96 % to a header at 20 bar g. The adsorber keeps 85.0 % of the hydrogen it is fed. The rest, with the methane and the carbon oxides, returns to the furnace as tail gas, and at the design point it supplies 73.92 MW of the 103.43 MW fired, 71.46 % of the fuel.

> **DID YOU KNOW? Refinery Hydrogen From 2008 to 2014**
>
> From 2008 to 2014 the hydrogen used by United States refineries rose 60 %. The hydrogen they made on site changed by less than 1 %, and the hydrogen bought from industrial gas suppliers rose 135 %. Made on site, the extra hydrogen would have doubled the refineries' natural gas feedstock [1].

The simulator is aimed at two beliefs. The first is that a ratio loop protects its ratio. A steam flow loop in ratio to the measured feed can only follow the feed, and when the steam supply can move only 4 % of its valve a minute while the feed answers in seconds, a plain ratio drops the steam-to-carbon ratio from 3.00 to 2.729 on a 15 % feed rise, below the 2.8 floor at which the catalyst risks carbon. Over 64 tunings of the steam loop the lowest ratio stays between 2.629 and 2.763, so the remedy is in the structure and not in the gains. The second belief is that recovering more hydrogen saves natural gas. Lengthening the adsorber's cycle from 10 to 11 min raises its recovery from 85.0 to 86.4 %, the tail gas carries less heat, and the furnace needs 8.4 % more makeup natural gas. Losing a bed does the reverse and the makeup falls by 16.8 %, while the natural gas per normal cubic meter of hydrogen rises from 0.4323 to 0.4403.

Both shapes recur well beyond reforming. A safety ratio between two streams of different speed appears in fuel and air on every burner, for which Volume 3, Chapter 4 offers cross-limiting, and in reactant feeds to exothermic reactors. A process whose byproduct is its own fuel or feed has the same coupling as a recycle, studied plantwide in Volume 6, Chapter 7, and refinery hydrogen itself is

analyzed as a distribution network between the plants that make it and the units that consume it [2]. However, the published treatments divide this plant among separate specialties. Reformer models are written for the tube and the furnace [3, 4], adsorber control is studied on a cycle with its own optimization [5], and ratio control is taught as a chapter of its own [6].

This chapter works with a complete steam methane reforming hydrogen plant in one browser simulation in which the steam-to-carbon ratio, the reformer outlet, the adsorber's cycle, the header and the furnace's fuel are coupled as they are on a plant. It measures what a lead-lag, a tail-gas feedforward, a cycle that follows the feed and a predictive controller do to that plant. The equilibrium, the radiant firebox, the adsorption correlation and the control laws are all established methods. The findings are measurements on one generic plant at one seed, checked at four more, among them that the lead-lag works by holding the feed back rather than by speeding the steam. The model has 46 states. Its carbon, hydrogen and oxygen balances close to 2.6×10^{-16} in every cell, its fuel balance closes to rounding, its equilibrium constants come from van 't Hoff fitted to the NIST-JANAF tables, and its boiler uses the IAPWS-IF97 module shared by the framework apps. The model is written out in full so that the plant can be rebuilt. Nine cases are provided, the ninth an eight-hour challenge, and each is graded on the learner's own moves.

Section 12.2 follows the domain, from reformer chemistry and the steam floor to the adsorber and the hydrogen header. Section 12.3 sets out the plant as simulated: the equilibrium cells, the constants, the tubes and firebox, the fuel balance, the boiler and shift, the adsorber, the drums and valves, the instruments and the protection. Section 12.4 describes the control layers, from the header loop through the ratio and its lead-lag to the outlet loop, its override and feedforward, and the predictive controller. Section 12.5 measures the nine cases and the challenge. Section 12.6 sets the hardware of a real plant beside the model, Section 12.7 states the limitations, Section 12.8 records where the boiler's property module came from, and Section 12.9 draws the lessons together. Section 12.10 holds the exercises.

12.2 Hydrogen From Natural Gas, and the Adsorber Behind It

Large-scale hydrogen is made from natural gas by catalytic steam reforming, and the engineering of the reformer, its catalyst and its furnace has its own literature [7, 8]. The chemistry is two reactions on nickel, the strongly endothermic reforming of methane and the mildly exothermic water-gas shift. A 1989 study derived intrinsic rate equations for both on a nickel catalyst by discriminating among twenty-one sets of mechanisms [9]. The simulator of this chapter sets kinetics aside and holds each cell at equilibrium, so the methane that slips through follows the outlet temperature, the pressure and the steam-to-carbon ratio through the equilibrium constants. An industrial steam-methane reformer has been modeled for on-line deployment [3], and a reforming reactor has been modeled by computational fluid dynamics with its control designed on the result [4].

> ⧖ HISTORY **1954: Refineries Begin to Need Hydrogen**
>
> After the Second World War, refining processes began to use catalysts and hydrogen to change molecules and remove sulfur. A federal manual dates hydrodesulfurization to 1954 and hydrocracking to 1960. Such units need hydrogen of 95 to 99 % purity, and the hydrogen recovered from the gases of catalytic reformers is often not enough, so refineries make more. The manual describes how. Desulfurized gas is mixed with superheated steam and reformed in tubes filled with a nickel catalyst, then passed through converters with an iron catalyst, where carbon monoxide and steam form carbon dioxide and more hydrogen. Amine washing removes the carbon dioxide, and the carbon monoxide that remains is converted to methane [10]. The plant of this chapter purifies its gas in adsorbers instead.

The steam-to-carbon ratio is a constraint before it is a setpoint. Steam in excess of the reforming stoichiometry is what keeps carbon from forming on the catalyst, and coke formation during steam reforming and its minimization are a subject of their own [11]. Every plant therefore carries a floor on

the ratio that its catalyst supplier sets for its feed, and a feed change becomes a sequencing problem, steam first on the way up and steam last on the way down. In the plant simulated here the steam is also the slower stream, because it comes from a steam system whose response is limited, so the order has to be imposed on the demand before either stream moves.

The adsorber is the plant's purifier. Pressure-swing adsorption separates by holding impurities on an adsorbent at high pressure and releasing them at low pressure. It has a book-length treatment [12] and an industrial commentary on its development [13]. Commercial adsorbers produce high purity hydrogen from steam reformer off-gas and refinery off-gases in several cycle variants, and ideas to raise their recovery by coupling them to membranes or to further adsorbers have been reviewed [14]. The adsorber is a cyclic unit feeding a continuous user. Its beds are switched on a schedule, its product and its tail gas both pulse, and the performance it delivers depends on its steps and its pressure profiles, which is why its design and control have been optimized together and closed with predictive controllers [5]. In a hydrogen plant the tail gas goes to the reformer's burners, so the adsorber's recovery is also a term in the furnace's fuel balance.

> ⧖ HISTORY **1960 and 1969: Beds That Swing in Pressure**
>
> A patent granted in 1960 dried air in two beds of adsorbent that took turns. One adsorbed at pressure while the other was opened to a lower pressure and backwashed with part of the product, without heat from an outside source [15]. A patent filed in 1967 and granted in 1969 purified hydrogen in at least four beds. Each bed stops adsorbing while its front is still inside it, passes the gas left between its grains straight to a cleaner bed until their pressures equalize, is blown down through its inlet and purged, and is brought back up last with product. The patent states that earlier processes, at 70 to 80 % recovery, gave hydrogen of only 95 to 98 % purity, and it names steam reformed hydrocarbons among its feeds [16]. The four beds of this chapter step through the same kinds of steps.

The plant's place in the refinery explains its rate loop. Stricter fuel specifications raised hydrotreating and hydrocracking capacity and with it the refinery's hydrogen consumption, leaving many refineries with a tight hydrogen balance [2]. The users draw on a header and nobody schedules the hydrogen plant for them, so the plant reads the header pressure and sets its own rate, as the fuel gas header of Chapter 14 integrates sources and users that nobody schedules for its sake. The carbon in the feed leaves either in the flue gas or, where it is captured, at one of the points of the reforming process at which capture has been assessed [17], and natural gas per unit of hydrogen is the figure that carries both the fuel bill and the carbon.

12.3 The Plant as Simulated

No plant stands behind the simulation, and every size, flow, correlation and threshold is a generic teaching value for a plant of 50 000 Nm^3/h. The layout, the symbols used in the equations, and the values that case 4 shows at 11.3 min are shown in Figure 12.1. Natural gas of 94 % methane, 5 % ethane and 1 % nitrogen (carbon number $n_C = 1.040$) is mixed with steam and enters a representative reformer tube at 600 °C. The reformed gas passes a waste-heat boiler, a high-temperature shift bed and a cooler, and a four-bed adsorber sends its product through a surge drum to the header and its tail gas through a tail-gas drum to the burners, where natural gas makeup completes the fuel.

The model integrates 46 states, listed by role in Table 12.1. Thirty-two are physical: the catalyst temperature T_i and the tube-metal temperature $T_{w,i}$ of each of eight cells, the bridge-wall temperature T_{bw}, the shift bed temperature T_{HTS}, the feed, makeup and tail-gas flows behind their flow loops, the steam and vent valve positions, the moles in the surge drum, its impurities and the moles in the header, and the moles of each of six species in the tail-gas drum. Fourteen are ledgers that integrate what crosses the boundary (feed, makeup, product, vent, demand, tail gas, adsorber feed, flue carbon dioxide, fired energy, boiler steam, process steam, the drum's inflow and the tail gas burned) and the firebox's energy balance, so that every total can be checked. The clock is the minute, temperatures

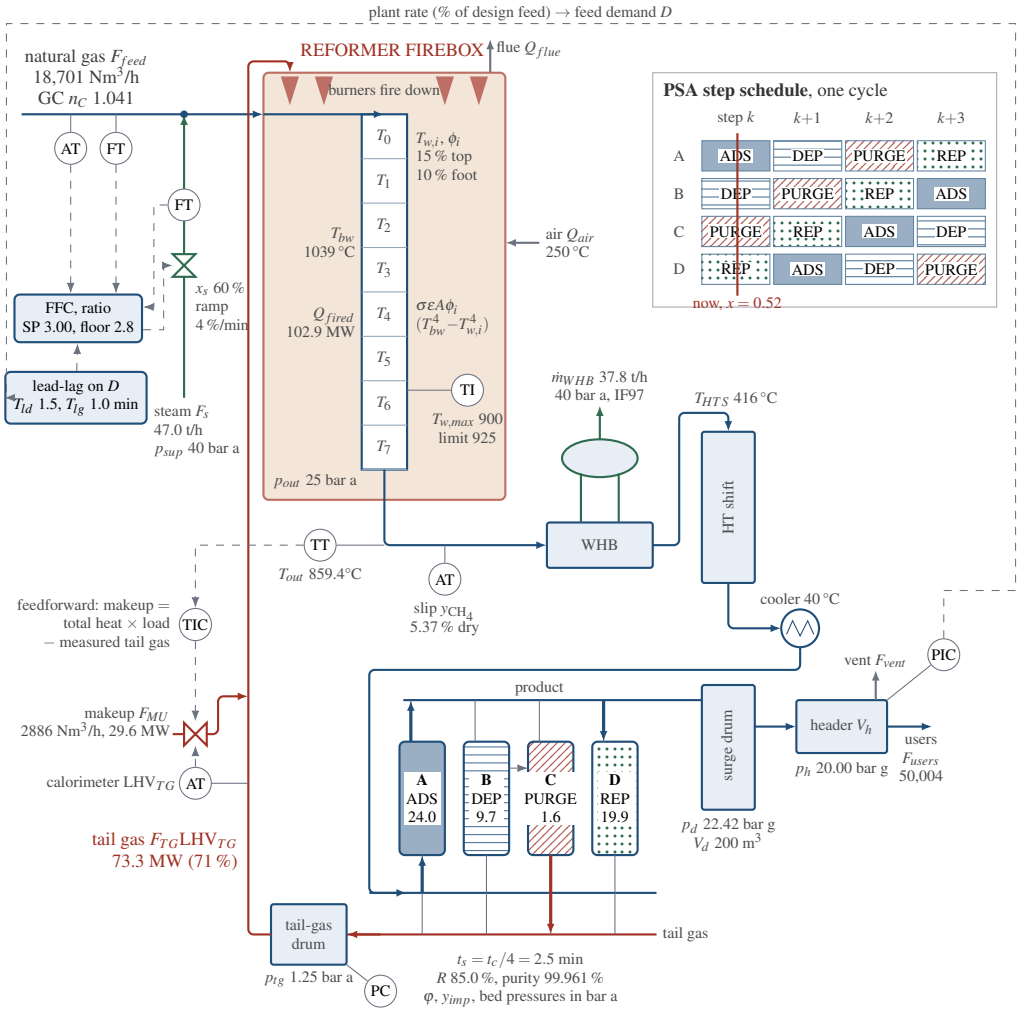


Figure 12.1 The hydrogen plant as modeled, drawn after the Plant view of case 4 at $t = 11.3$ min, when the adsorption step is 52 % done and all four phases are distinct: bed A adsorbs at 24.0 bar a (solid), B depressurizes at 9.7 (ruled) and provides purge gas, C is blown down and purged at 1.6 (hatched) and D is repressurized with product at 19.9 (dotted). Feed and steam meet at the tee under the ratio station, whose lead-lag acts on the feed demand D the header loop sets. The eight cells T_0 to T_7 sit in a top-fired box at T_{bw}, the boiler raises 37.8 t/h at 40 bar a, and the tail gas, 71 % of the 102.9 MW fired by the meters, returns to the burners with the makeup. The adsorber, the surge drum, the header and the tail-gas drum form the lower row, fed by the cooled gas that runs back under the reformer. The inset is the step schedule of one cycle with the present instant marked. Process model only, not to scale.

Table 12.1 Variables by Role at the Design Point

Symbol	Role	Design	Range
p_h	CV, header (bar g)	20.00	SP 18–22
S/C	CV, steam-to-carbon (mol/mol)	3.000	SP 2.5–4, floor 2.8
T_{out}	CV, reformer outlet (°C)	860.0	SP 820–900
r	MV, plant rate (% of design feed)	100	40–140
x_s	MV, steam valve (%)	60.0	0–100
F_{MU}	MV, makeup fuel (Nm^3/h)	2885	0–12,000
t_c	MV, adsorber cycle (min)	10	6–16
d	DV, hydrogen demand (% of design)	100	60–140
$x_{C_2H_6}$	DV, ethane in the feed (mol %)	5	0–20
p_{sup}	DV, steam supply (% of 40 bar a)	100	60–110
—	DV, makeup heating value (%)	100	70–100
T_i, $T_{w,i}$	states, cells, walls (°C)	602–860, 713–900	limit 925
T_{bw}	state, bridge wall (°C)	1040.9	limit 1120
T_{HTS}	state, shift bed (°C)	416.8	—
n_d, n_h	states, drum and header (kmol)	—	—
g_k	states, tail-gas drum species (kmol)	—	p_{tg} 1.25 bar a
y_{CH_4}	output, slip (% dry)	5.367	—
R, x_{H_2}	output, recovery, purity (%)	85.00, 99.960	spec 99.9
Q_{fired}	output, fired heat (MW)	103.43	burners 170
$\dot{m}_{WHB}$	output, boiler steam (t/h)	37.80	—

are in °C, flows in kmol/h inside and Nm^3/h (0 °C, 1.013 25 bar a) on the screen, header and drum pressures in bar gauge, process pressures in bar absolute, and heat in MW on the lower heating value.

12.3.1 Equilibrium in Eight Cells

The gas crosses a cell in a fraction of a second while the catalyst holds its heat, so a cell's composition is solved algebraically while its temperature is integrated. Every cell holds the equilibrium of the tube's inlet elements, reforming evaluated ΔT_{app} colder than the cell and the shift at the cell's own temperature, as shown in Equation (12.1).

$$K_1(T_i - \Delta T_{app}) = \frac{y_{CO}\, y_{H_2}^3}{y_{CH_4}\, y_{H_2O}}\, p_i^2, \qquad K_2(T_i) = \frac{y_{CO_2}\, y_{H_2}}{y_{CO}\, y_{H_2O}}$$
$$p_i = p_{out} + \Delta p_t \left(\frac{F}{F_0}\right)^2 \left(1 - \frac{i + 0.5}{8}\right) \tag{12.1}$$

where reaction 1 is $CH_4 + H_2O \rightleftharpoons CO + 3H_2$, reaction 2 is $CO + H_2O \rightleftharpoons CO_2 + H_2$, y are mole fractions, $p_{out} = 25\,\text{bar}\,a$, $\Delta p_t = 2\,\text{bar}$ at the design total flow F_0, and $\Delta T_{app} = 10\,\text{K}$ stands for the catalyst's activity. The two extents are found by Newton's method on the logarithms of the residuals with a fixed number of iterations (12 in the first cell, 8 after it, warm-started from the cell before), each step cut so that no species falls below a tenth of its value, so the result depends only on its inputs and the derivative function stays pure. Ethane reforms completely where the feed enters, $C_2H_6 + 2H_2O \rightarrow 2CO + 5H_2$. At design the outlet carries 171.32, 1638.00, 427.70, 268.66 and 2316.01 kmol/h of methane, water, carbon monoxide, carbon dioxide and hydrogen, and a hand calculation from the same constants reproduces every figure to the last digit.

12.3.2 Equilibrium Constants From the Tables

Each species' enthalpy is a straight line $h_k(T) = a_k + c_k T$ fitted to the NIST-JANAF tables [18] from 600 to 1200 K, its largest error 1.15 kJ/mol for methane. The reaction enthalpies follow from those lines, and van 't Hoff integrated with them fixes $\ln K$ up to one constant, as shown in Equation (12.2).

$$\ln K = -\frac{\Delta A}{RT} + \frac{\Delta c}{R} \ln T + I, \qquad \Delta H(T) = \Delta A + \Delta c\, T \tag{12.2}$$

where ΔA and Δc are the stoichiometric sums of a_k and c_k and only I is fitted to the tables' $\log K_f$ from 700 to 1200 K ($\Delta A = 209.84$ and −44.97 kJ/mol, $\Delta c = 0.01531$ and 0.010 148 kJ/(mol K), $I = 15.776$ and -13.475, K_1 in bar^2). The constants and the heats of reaction are therefore one consistent set. Against JANAF interpolated in $1/T$, K_1 is −1.44 % low at 850 °C and −1.11 % at 900 °C, and within 1.6 % from 700 to 900 °C (−1.51 % at 800 °C is the worst). The NIST WebBook's Shomate fit of the same tables gives −1.34 % and −1.12 %. K_2 is within 0.25 % of JANAF at 850 and 900 °C and within 0.19 % from 350 to 450 °C, the shift bed's range. At 900 °C $\Delta H_1 = 227.80\,\mathrm{kJ/mol}$ against JANAF's 227.1 and $\Delta H_2 = -33.06\,\mathrm{kJ/mol}$ against −33.1. With JANAF's own constants the design slip would be 5.321 % instead of 5.367 %. Ethane's formation enthalpy, −84.0 kJ/mol, is the WebBook's. Equilibrium makes the slip a function of temperature, pressure and ratio alone: at the design ratio it is 6.79, 5.37 and 4.17 % at 840, 860 and 880 °C, and 4.10 and 6.56 % at 20 and 30 bar a.

12.3.3 Tubes and a Radiant Firebox

Each cell's catalyst is heated by its wall through an inside conductance, each wall by radiation from a well-stirred firebox at the bridge-wall temperature, and the firebox by the burners, as shown in Equation (12.3).

$$\begin{aligned} C_{cat}\frac{\mathrm{d}T_i}{\mathrm{d}t} &= \dot{H}_{i-1} - \dot{H}_i + UA_i(T_{w,i} - T_i) \\ C_w\frac{\mathrm{d}T_{w,i}}{\mathrm{d}t} &= \sigma\,\varepsilon A\,\phi_i\left(T_{bw}^4 - T_{w,i}^4\right) - UA_i(T_{w,i} - T_i) \\ C_{fb}\frac{\mathrm{d}T_{bw}}{\mathrm{d}t} &= Q_{fired} + Q_{air} - \textstyle\sum_i Q_{rad,i} - Q_{flue} \end{aligned} \tag{12.3}$$

where C_{cat} and C_w are each an eighth of the totals, $UA_i = UA/8$, $\dot{H}$ is the enthalpy flow with formation enthalpies included (so the reaction heats need no separate term), $\dot{H}_{-1}$ is the mixed feed's, temperatures in the radiation term are in kelvin, and the flux shares ϕ_i of a top-fired box are 0.15, 0.15, 0.14, 0.13, 0.12, 0.11, 0.10 and 0.10 from the top. The flue gas is computed species by species from the fuels and the air, and $Q_{flue} = \sum_k n_k c_k (T_{bw} - 298.15)/3600$ and Q_{air} likewise at the 250 °C air preheat. At design the tubes absorb 52.14 MW of the 103.43 MW fired, 50.41 %, the bridge wall is at 1040.9 °C and the hottest wall, in the bottom cell, at 899.6 °C. A test of the application confirms that the heat stored in the firebox, walls and catalyst equals the integrated balance over a disturbed hour.

12.3.4 The Fuel Balance

The burners release the lower heating value of the tail gas and the makeup, as shown in Equation (12.4).

$$Q_{fired} = F_{TG}\sum_k x_k\,\mathrm{LHV}_k + F_{MU}\,\mathrm{LHV}_{NG} \tag{12.4}$$

where x_k is the tail-gas drum's composition, k runs over hydrogen, methane and carbon monoxide at 241.826, 802.3 and 282.995 kJ/mol (from the same formation enthalpies), and natural gas carries 802.3 and 1428.5 kJ/mol for methane and ethane. At design $Q_{fired} = 73.917 + 29.515 = 103.432$ MW, closing to 4×10^{-15}, so the tail gas, 9.279 MJ/Nm^3 at 28 678 Nm^3/h, supplies 71.46 % of the fuel and the makeup is 2885 Nm^3/h. The instantaneous firebox residual at $t = 0$, $103.432 + 9.463 - 60.735 - 52.137 = 0.023$ MW, comes from the air controller's initial estimate. Two simplifications of the furnace are declared and kept. The flue and air sensible heats use the slopes of the fitted lines from 298.15 K rather than the tables' own enthalpy rise. With JANAF enthalpies at the design bridge wall the flue would carry 58.37 MW instead of 60.65 and the air 8.48 instead of 9.45, and a hand balance needs 102.03 MW fired against 103.33 on the model's basis, so the model over-fires by about 1.3 MW and buys about 4 % more makeup (2855 against 2728 Nm^3/h by hand). The same linear enthalpies put the shift outlet at 416.8 °C, 3.5 K below the 420.3 that JANAF enthalpies give, with 3.43 % carbon monoxide dry against 3.52, and ethane's mean heat capacity overstates the feed's sensible heat by 0.134 MW.

12.3.5 Boiler, Shift and Cooler

A bypass holds the boiler's gas outlet at 350 °C, and the duty raises saturated steam at 40 bar a from feedwater at 105 °C, $\dot{m}_{WHB} = Q_{WHB}/(h'' - h_{fw})$. The properties come from the IAPWS-IF97 formulation [19] in the framework's shared module, $h'' = 2800.90\,\mathrm{kJ/kg}$ at $T_{sat} = 250.358\,^{\circ}\mathrm{C}$ and $h_{fw} = 443.45\,\mathrm{kJ/kg}$ at 45 bar a, so the 24.752 MW duty raises 37.80 t/h. That is less than the 46.9 t/h of process steam, so the boiler is not an export. The module's table stops at 25 bar, and the boiler calls its direct saturation functions, which match Region 2 at saturation to the printed digit. The shift bed is one adiabatic vessel with a heat capacity C_{HTS}, its gas at the shift equilibrium evaluated 10 K hotter than the bed, $C_{HTS}\,\mathrm{d}T_{HTS}/\mathrm{d}t = \dot{H}_{in}(350\,^{\circ}\mathrm{C}) - \dot{H}_{out}(T_{HTS})$. The cooler condenses water down to its IF97 vapor pressure at 40 °C, 0.073 84 bar, which leaves 3077 ppm in the adsorber's feed at 24 bar a.

12.3.6 The Adsorber

Four beds stand on a ring of four positions, adsorption, cocurrent depressurization, blowdown and purge, and repressurization, and every step of $t_s = t_c/4$ each bed moves one position, so exactly one bed adsorbs at every instant. Within a step each bed's pressure follows a prescribed shape: an equalization from $p_{hi} = 24\,\mathrm{bar}\,a$ toward $p_{eq} = (p_{hi} + p_{lo})/2$ in the first fifth, a provide-purge ramp to 5 bar a, a blowdown to $p_{lo} = 1.6\,\mathrm{bar}\,a$ in the first quarter, and product repressurization, each segment an exponential that ends exactly on its target. The flows these imply are exact derivatives, $C_b\,\mathrm{d}p/\mathrm{d}t$ with $C_b = 0.6\,\mathrm{kmol/bar}$, so the product taken by the repressurizing bed and the tail gas sent by the blowdown and purge both pulse once a step. The cycle-average performance follows a declared teaching correlation, as shown in Equation (12.5).

$$\begin{aligned} R &= 1 - \ell_N \frac{t_{s,0}}{t_s}\frac{F_{\mathrm{H}_2,0}}{F_{\mathrm{H}_2}}, \qquad \varphi = \varphi_0 \frac{F_{imp}}{F_{imp,0}}\frac{t_s}{t_{s,0}} \\ y_{imp} &= y_0\,\mathrm{e}^{(\varphi-\varphi_0)/w}\,(0.4 + 1.2x) \end{aligned} \tag{12.5}$$

where F_{H_2} and F_{imp} are the hydrogen and the impurities fed, the subscript 0 marks design, $\ell_4 = 0.15$ with four beds and $\ell_3 = 0.20$ with three, φ is the bed's loading per step ($\varphi_0 = 0.80$), $y_0 = 400\,\mathrm{ppm}$, $w = 0.12$, and x is the fraction of the step elapsed, so the impurity rises as the front climbs the bed. A fixed hydrogen loss per step makes recovery rise and purity fall with a longer step: 81.25, 85.00, 86.36 and 87.50 % recovery and 99.989, 99.960, 99.922 and 99.848 % purity at steps of 2.0, 2.5, 2.75 and 3.0 min, and the same with 20 % more feed. These are the directions of the usual adsorber trade [12], in which a longer adsorption step wastes less void gas per mole fed but runs the front nearer breakthrough. The purge pressures shape the pulses only, and setting them to 3 and 2.5 bar a leaves recovery and purity unchanged. With a bed out the ring closes to three positions at the next step boundary, the adsorption step is kept, and the cycle shortens to three steps. The product sent to the drum is the cycle-average product less the repressurization pulse about its mean, and the tail gas is the remainder of the feed plus the blowdown pulse, carrying its average composition.

Pause. Try Exercise 12.2 before reading on.

12.3.7 Drums, Header and Valves

The surge drum, the header and the tail-gas drum are isothermal ideal-gas volumes, as shown in Equation (12.6).

$$\begin{aligned} \frac{\mathrm{d}n_d}{\mathrm{d}t} &= F_{prod} - k_{dh}(p_d - p_h) \\ \frac{\mathrm{d}n_h}{\mathrm{d}t} &= k_{dh}(p_d - p_h) - F_{users} - F_{vent} \\ \frac{\mathrm{d}g_k}{\mathrm{d}t} &= F_{tail}\,z_k - F_{TG}\,x_k, \qquad p = \frac{nRT}{V} \end{aligned} \tag{12.6}$$

where k_{dh} passes the design product across a 2.5 bar drop, $V_d = 200\,\mathrm{m}^3$ and $V_h = 1000\,\mathrm{m}^3$ at 40 °C, $V_{tg} = 1000\,\mathrm{m}^3$ at 35 °C, z_k is the tail gas's composition and F_{users} is the demand. An imbalance of 5000 Nm^3/h moves the header 0.0968 bar a minute. The drum's time constant $V_d/(RTk_{dh})$ is 0.13, 0.52 and 1.03 min at 50, 200 and 400 m^3. The feed follows its setpoint through a 0.1 min lag up to 145 % of design, the makeup and tail gas through 0.1 min lags, and the process steam passes a valve whose stem follows its command through a 0.5 min lag and at most 4 % of travel a minute, as shown in Equation (12.7).

$$\begin{aligned} \frac{\mathrm{d}x_s}{\mathrm{d}t} &= \mathrm{clip}\left(\frac{x_s^* - x_s}{\tau_{sv}},\ \pm r_{sv}\right) \\ F_s &= F_{s,\max}\, x_s \sqrt{\frac{p_{sup} - p_{mix}}{p_{sup,0} - p_{mix}}} \end{aligned} \tag{12.7}$$

where $F_{s,\max}$ passes the design steam at 60 % open and $p_{mix} = 28\,\mathrm{bar}\,a$ at the tee. The ramp stands for the steam system behind the valve, which can raise the reformer's draw only so fast.

🕮 **Pause.** Try Exercise 12.3 before reading on.

12.3.8 Instruments

The controllers see only a packet of instrument readings, and a test of the application whitelists its keys. The feed, steam, makeup, tail-gas, product and demand meters have a 2 s lag and 0.3 % noise. A feed chromatograph reports ethane and nitrogen, and so the carbon number, from a sample every 5 min analyzed in 3 min behind 30 s of line. A calorimeter reports the tail gas's heating value every 30 s, 1 min after sampling. The outlet chromatograph reports the dry slip every 6 min, 5 min late, and a purity analyzer the header's impurities every minute. The outlet, tube-metal (an infrared scan), bridge-wall and shift temperatures have a 12 s lag and 0.3, 1.5, 1.0 and 0.3 K noise, the flue oxygen a 20 s lag, and the pressures a 2 s lag. Every execution consumes a fixed count of random numbers regardless of the settings, which keeps two runs that differ in a single setting aligned draw for draw. In the exported file every true value carries the prefix `truth_`, so a controller tested on the file can be restricted to the measured columns.

12.3.9 Initialization, Protection and Numerics

Table 12.2 gives every parameter with its default. Each case begins lined out on whatever settings it ships with. A bisection on the bridge-wall temperature brings the tube's outlet to its setpoint, the shift and the adsorber follow from the outlet gas, the makeup closes the firebox balance at the oxygen setpoint of 2 % wet, and the drums are filled to the header setpoint. The status strip raises CARBON RISK below the floor on the true ratio, TUBE METAL LIMIT above 925 °C, PURITY LOW below 99.9 % and HEADER LOW below 19.0 bar g, and five minutes below the floor stops the run. The integrator is fourth-order Runge-Kutta with a step of 0.5 s, the controllers run every second and the historian writes a row every 5 s, on the clocks of Section 1.6.1. Halving the step changes no verdict of the challenge and moves its reported figures by at most 0.006 bar on the header, and case 2's lowest ratios not at all.

12.3.10 Assumptions

The reformer is one representative tube at equilibrium with a declared approach, so no kinetics [9] enter and the result does not depend on the catalyst's activity profile, and the model is least reliable where a real tube is far from equilibrium near its inlet. The firebox is one well-stirred volume with a fixed exchange area and flux profile, and the convection section that preheats the feed and the air and raises more steam is not modeled. The flue and air enthalpies use the fitted slopes, which over-fires the furnace by about 1.3 MW. The adsorber's recovery and purity follow a teaching correlation, not an isotherm, and its pulses come from prescribed pressures, not from bed dynamics. Gases are ideal and the drums isothermal. Every limit is a teaching value: real steam-to-carbon floors come from the

Table 12.2 Parameters and Defaults. A, Adjustable in the Parameters Panel. C, a Control on a Card. F, Fixed in the Source

Symbol	Meaning	Value	Unit	
F_0	design natural gas feed	18,700	Nm^3/h	A
T_{in}	mixed feed into the tubes	600	°C	A
τ_F, —	feed loop lag, feed valve capacity	0.1, 145	min, %	A
τ_{sv}, r_{sv}	steam valve lag, ramp limit	0.5, 4	min, %/min	A
$p_{sup,0}$	steam supply pressure	40	bar a	A
p_{out}, Δp_t	outlet pressure, tube drop	25, 2	bar	A
ΔT_{app}	approach to reforming equilibrium	10	K	A
UA, εA	wall conductance, radiant area	0.75, 600	MW/K, m^2	A
C_{cat}, C_w, C_{fb}	catalyst, metal, firebox capacity	30, 70, 30	MJ/K	A
—	burner capacity, air preheat	170, 250	MW, °C	A
$T_{w,\max}$, —	tube-metal, bridge-wall limits	925, 1120	°C	A
—	shift inlet, approach, capacity	350, 10, 40	°C, K, MJ/K	A
—	boiler steam pressure	40	bar a	A
p_{hi}, p_{lo}, p_{pp}	adsorber pressures	24, 1.6, 5	bar a	A
C_b	bed gas holdup	0.6	kmol/bar	A
ℓ_4, ℓ_3	hydrogen loss share at design	0.15, 0.20	—	A
φ_0, y_0, w	loading, impurity, sharpness	0.80, 400, 0.12	—, ppm, —	A
—	drum-to-header drop at design	2.5	bar	A
V_d	surge drum (50, 100, 200, 400)	200	m^3	C
V_h, V_{tg}	header, tail-gas drum	1000, 1000	m^3	A
—	vent capacity	15,000	Nm^3/h	A
p_{tg}	tail-gas drum setpoint	1.25	bar a	A
$(S/C)_{\min}$, —	floor, time below it to stop	2.8, 5	mol/mol, min	A
—	purity spec, HEADER LOW	99.9, 19.0	%, bar g	A
ϕ_i	flux shares, top to bottom	0.15 to 0.10	—	F
—	makeup valve capacity	12,000	Nm^3/h	F
h_k, K_1, K_2	enthalpy lines, constants	Sec. 12.3.2		F

catalyst supplier's carbon-limit data for the feed [11], real tube-metal limits from the tube alloy's creep-rupture life under API Standard 530, and real purity specifications and header alarms from the users.

12.4 The Control Layers

Each physics step runs the controllers, then the integrator, then the adsorber's ring and the protection, and finally the outputs, in the order of Figure 1.11. The controllers execute every $\Delta t_c = 1$ s. The header, steam flow, oxygen trim and tail-gas drum loops are the velocity-form PID of Section 1.6.2, Equations (1.16) to (1.19), with $a = \pm 1$ setting the direction, the running sum never clamped, a limit correcting the integral alone through a tracking time equal to the integral time, and no derivative action. Setting the integral time to its 60 min maximum leaves a proportional loop. In Manual each loop solves for the integral that reproduces the station's value, so every transfer is bumpless.

12.4.1 Header Pressure and the Plant Rate

The header loop raises the plant rate r when the measured header falls ($a = +1$), with output limits of 40 and 140 % and the shipped tuning 15 % per bar and 8 min. The rate then passes a rate-of-change limit of 20 % a minute that binds operator and controller alike, and the feed demand is $D = rF_0/100$. A proportional vent opens 50 % per bar above a setpoint 1 bar over the header's. The header Manual/Auto switch belongs to this loop alone, and the ratio and outlet loops have their own switches on their cards.

12.4.2 The Ratio and Its Lead-Lag

The ratio station multiplies a steam-to-carbon setpoint by the carbon fed, the feed meter times the chromatograph's carbon number. With the lead-lag on it works from the demand D instead of the

meter, and the two streams get opposite dynamics, as shown in Equation (12.8).

$$\begin{aligned} f_k &= \begin{cases} f_{k-1} + \alpha_{ld}(D_k - f_{k-1}) & D_k > f_{k-1} \\ D_k & \text{otherwise} \end{cases} \\ c_k &= \begin{cases} c_{k-1} + \alpha_{lg}(D_k - c_{k-1}) & D_k < c_{k-1} \\ D_k & \text{otherwise} \end{cases} \\ F_{feed,SP} &= f_k, \qquad F_{s,SP} = (S/C)_{SP}\, n_{C,m}\, c_k \end{aligned} \tag{12.8}$$

where $\alpha_{ld} = 1 - \mathrm{e}^{-\Delta t_c/T_{ld}}$ with the lead time $T_{ld} = 1.5\,\text{min}$ and $\alpha_{lg} = 1 - \mathrm{e}^{-\Delta t_c/T_{lg}}$ with $T_{lg} = 1\,\text{min}$. Going up the steam setpoint takes the new demand at once while the feed follows through a lag, and going down the feed falls at once while the steam follows. With the lead-lag off $F_{feed,SP} = D_k$ and $F_{s,SP} = (S/C)_{SP}\, n_{C,m} F_{feed,m}$. A steam flow PI ($a = +1$ on percent of design steam, 0.5 % per percent and 1 min) moves the valve. The lead-lag acts on the demand and not on the feed meter, because no filter on a feed that has already moved can make the steam lead it. The single floor of 2.8 serves the status flag, the run stop and the challenge. It sits 0.2 below the design so that the feed steps of cases 1 and 2 reach it, while case 1's held steam bottoms at 2.609 and every plain-ratio dip measured falls between 2.63 and 2.76, so a floor of 2.5 would never be crossed.

Pause. Try Exercise 12.1 before reading on.

12.4.3 Outlet Temperature, Override and Feedforward

The outlet loop and the tube-metal override are a pair of PI laws written with external reset feedback under a low select, the same form as the carbon monoxide override of Volume 3, Chapter 4 with a high select, as shown in Equation (12.9).

$$\begin{aligned} u_T &= K_T\, e_T + \rho_T, \qquad u_O = K_O\, e_O + \rho_O \\ q &= \min(u_T, u_O) \\ \rho_j &\leftarrow \rho_j + \left(1 - \mathrm{e}^{-\Delta t_c/\tau_j}\right)(q - \rho_j), \quad j \in \{T, O\} \end{aligned} \tag{12.9}$$

where $e_T = T_{out,SP} - T_{out,m}$, $e_O = (T_{w,\max} - 10) - T_{w,m}$, the shipped tuning is $K_T = 0.5\,\text{MW/K}$ with $\tau_T = 5\,\text{min}$ and $K_O = 1\,\text{MW/K}$ with $\tau_O = 2\,\text{min}$, and both resets follow the selected output, so the law not selected stays higher by the size of its own proportional term and cannot wind up. The build record reports that a low select with ordinary PID tracking, tried first, ratcheted the makeup to zero on the tube-metal signal's 1.5 K of noise, which external reset feedback cannot do. On *Loops* q is the makeup's heat and the tail gas is a disturbance the loop meets only through the outlet. With the tail-gas feedforward q is the total heat at the design feed, scaled by the measured load, and the makeup is what the measured tail gas does not bring, the heat-balance feedforward of Equation (1.2) with the tail gas as the measured load, as shown in Equation (12.10).

$$F_{MU,SP} = \frac{3600}{\text{LHV}_{NG}} \left(q\, \frac{F_{feed,m}}{F_0} - F_{TG,m}\, \text{LHV}_{TG,m} \right) \tag{12.10}$$

where the tail gas's heating value is the calorimeter's. The load factor keeps the feedforward from over-firing when the header loop cuts the feed. The burners cap the total at 170 MW.

12.4.4 Air, Adsorber Step, Tail-Gas Drum and Slip

The combustion air is the oxygen the measured fuels need, from the calorimeter and the design tail gas's oxygen per megajoule, times one plus an excess that holds 2 % wet oxygen, trimmed by a PI on the flue oxygen. The adsorber's step is $t_c/4$, or with *cycle follows the feed* $(t_c/4)\,\text{clip}\big(C_0/(F_{feed,m} n_{C,m}), 0.5, 1.5\big)$ with C_0 the design carbon, so each bed takes the same load. An averaging PI on the tail-gas drum, 1200 kmol/h per bar and 4 min, sets the tail gas to the burners

through a valve capped at three times the design flow scaled by the drum's pressure over the burners' 1.05 bar a. A softer loop let the drum rise past the purge pressure or let Loops fall into a limit cycle. The slip is inferred by Equation (12.1) at the measured outlet, the measured ratio and 25 bar a with the commissioned 10 K approach, and a bias updater with a factor of 0.5 trims it each time the outlet chromatograph reports, the update of Equation (1.14).

12.4.5 Model Predictive Control as a Comparison

The predictive controller plans three setpoints every 30 s over 15 min: the plant rate, the ratio setpoint and the outlet setpoint, which the loops below then follow. Its model is linear and was identified by steps on the plant at design: the header integrator $\dot{p} = K_h(Gr - F_{users,m} - b)$ with K_h from the declared volume and a bias b updated from the measured pressure, a 1.2 min feed lag, the ratio as the steam lags the feed by 1.5 min, first-order outlet (4 min), tube-metal (5 min) and bridge-wall (3 min) responses. The quadratic program, discretized exactly with a zero-order hold, is Equation (12.11).

$$\begin{aligned} \min\ & \textstyle\sum_{k=1}^{30} \left[w_p e_{p,k}^2 + w_{S/C} e_{S/C,k}^2 + w_T e_{T,k}^2 \right] \\ & + w_{\Delta u} \textstyle\sum_j \| S \Delta u_j \|^2 + \rho \| s \|^2 \\ \text{s.t. } & 40 \le r \le 140, \quad 2.8 \le (S/C)_{SP} \le 3.6 \\ & 840 \le T_{SP} \le 880 \\ & |p - p_{SP}| \le 0.4 + s, \quad S/C \ge 2.85 - s \\ & T_w \le T_{w,\max} - 5 + s, \quad T_{bw} \le 1120 + s, \quad s \ge 0 \end{aligned} \tag{12.11}$$

where $w_p = 100$ per bar^2, $w_{S/C} = 50$, $w_T = 1$ per K^2, $w_{\Delta u} = 30$ with move scales $S = (0.05, 5, 0.05)$, a terminal header weight of $5w_p$, $\rho = 10^4$, 30 steps of 30 s in 10 graduated move blocks, and moves capped at 5 % of rate, 0.2 and 5 K per decision. Without the cycle following the feed a rate ceiling keeps the inferred bed loading, judged against the feed at which that loading was adsorbed, under that of a 99.92 % product. The framework's solver applies Goldfarb and Idnani's dual method [20] to the program, only the first planned move is sent, which is the receding-horizon practice of industry [21] described in Section 1.6.3, and a solve that fails or leaves its box hands the setpoints to the loops with the feedforward. No solve was rejected in any challenge run of this chapter. The controller is offered as a comparison and not as the challenge's answer.

The model starts each decision at the measured ratio, and its gains are sized once from the declared defaults, so a learner's change to the plant is model mismatch. The ceiling's feed is the measured feed delayed by the purity analyzer's 1.5 min and mixed through the product drum. Before a correction in November 2026 the predicted ratio started mirrored about its setpoint, and the challenge passed at two seeds of five.

12.4.6 Three Models

Section 12.3 is the plant. What a learner identifies is the header's integrating gain and the ratio's dip on a step. The controllers' models are the ratio station's carbon number, the feedforward's design heat and meters, the slip inference's equilibrium with the commissioned approach, and the predictive controller's linear plant at design. None of them knows a catalyst that has aged, a makeup carrying inert gas, or a bed that has failed until a measurement shows it.

12.5 Nine Cases, Measured

The cases appear in Table 12.3. Each starts from the defaults above on Loops with the lead-lag off and the cycle fixed, and a row lists what differs. Every number below was measured headless on the engine at seed 20261005, stepped as the browser steps it (Section 1.6.1), with a row's settings entered before Run. Four more seeds, 0 to 3, were run for the challenge, and seeds 1 and 2 for cases 1, 2, 6 and 8. Every case runs identically loaded first, chosen after another case, and after Reset.

Table 12.3 The Nine Cases

#	Case and events	Teaching point
1	header Manual, ratio on Man, cycle follows, demand and rate +15 % at 2 min	the steam held
2	header Manual, ratio Auto, cycle follows, lead-lag off, same step	plain ratio, lead-lag
3	approach aged to 20 K	outlet, slip inference
4	surge drum 200 (or 50, 400) m^3	the sawtooth
5	header Manual, cycle 10 to 11 min at 5 min	recovery is firing
6	demand +10 % at 5 min, cycle follows	every loop moves
7	demand +40 % at 5, firm header, lead-lag, follows	infeasible
8	a bed valve fails at 5 min	three beds
9	challenge, 8 h: demand +20 % at 1, ethane 10 % at 4, demand back at 5, a bed fails at 6 h	the whole plant

12.5.1 The Steam Held, Then a Plain Ratio

In case 1 the operator raises the plant rate to 115 % at 2 min with the ratio loop on Man, and the ratio falls from 3.00 to 2.609 by 5.80 min, the same 15 % the carbon rose. CARBON RISK appears at 2.45 min and five minutes below the floor stop the run at 7.45 min. Switching the ratio loop to Auto at 3 min has the ratio back above 2.95 at 7.11 min, and the *Set S/C to design* helper, which commands the valve to 69.0 % from the meters and the supply pressure transmitter, does it at 5.03 min. Because the valve ramps, a late remedy cannot beat the stop. In quarter-minute steps Auto saves the run up to 5.75 min and not at 6.00, and the helper up to 6.25 and not at 6.50 min. The stop was left at 5 min on purpose. A 10 min stop moves the deadlines to 10.75 and 11.25 min, which is 4.2 s of wall time at the default 2 min/s rather than 1.65 s, so the Pause the briefing asks for is needed either way, it pushes the stop outside the case's 10 min trend window, and a real low-ratio trip acts faster, not slower.

> **TRY IT IN THE APP** **1 · More feed, the same steam (Manual)**
>
> Load the case and press Run. At 2 min the plant rate goes to 115 % and nobody touches the steam. Watch the Steam-to-carbon trend: the ratio falls from 3.00 to 2.609 by 5.80 min, the same 15 % the carbon rose, and CARBON RISK appears at 2.45 min. Left alone, five minutes below the floor stop the run at 7.45 min.

In case 2 the ratio loop is on Auto with the lead-lag off, and the same step drops the true ratio to 2.729 at 2.84 min, 0.78 min below the floor, as shown in Figure 12.2a. With the lead-lag on it never falls below 2.994. The cause is the steam's ramp, not the loop's tuning. A steam loop of 3 % per percent and 0.2 min still reaches 2.753, a ramp of 20 % a minute with the shipped loop 2.758, and only both together hold 2.957. Over 64 tunings from 0.05 to 5 % per percent and 0.05 to 60 min the plain ratio's low lies between 2.629 and 2.763. A bound by hand, a valve ramping at 4 % a minute from the instant the rate starts to rise, before any meter sees the feed, with the feed following the 20 %-a-minute rate through its 0.1 min lag, reaches at best 2.774, 0.85 min after the step. The claim is therefore proved for this case's rise and not for every rise. At a rate limit of 10 or 5 % a minute the best of nine steam tunings holds 2.870 or 2.984. Figure 12.2b shows how the lead-lag succeeds. The steam valve moves the same way in both runs, ramping from 60 to 69 %, and what changes is the feed, which the lead-lag holds back through its 1.5 min lag until the steam has caught up. Seeds 1 and 2 give 2.732 and 2.729 plain and 2.998 and 2.996 with the lead-lag, and halving the step changes neither.

12.5.2 A Slow Furnace and an Inferred Slip

Case 3 ages the catalyst to a 20 K approach that the inference still assumes to be 10 K. At 0.5 min the true slip is 6.08 % and the inference reads 5.43. The first chromatograph report, 6.08 % at 4.01 min, moves the bias to 0.39, and later reports take it to 0.54 and 0.64. Raising the outlet setpoint to 875 °C at 15 min brings the outlet past 874 °C only at 24.14 min, the furnace's slowness from burner to firebox to wall to catalyst, while the makeup climbs from 2375 to a peak of 3447 Nm^3/h and settles at

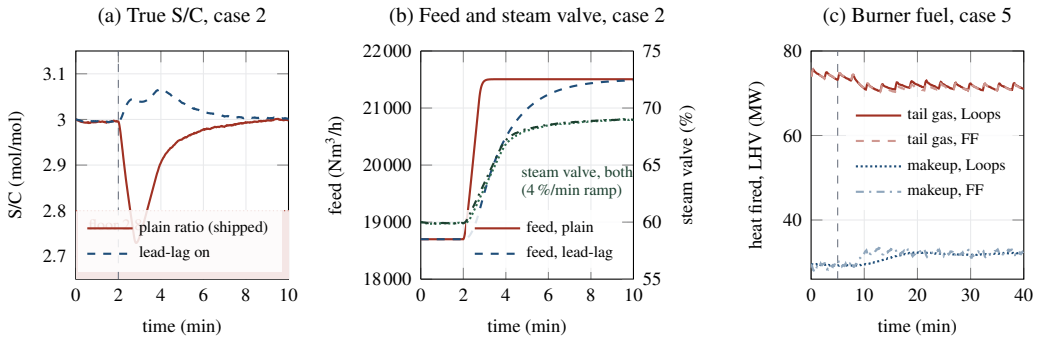


Figure 12.2 Three measured responses. (a) The true steam-to-carbon ratio in case 2 after the 15 % feed step at 2 min, plain ratio as shipped (solid) against the lead-lag (dashed), with the 2.8 floor (dotted) and the region below it shaded. (b) The feed (solid plain, dashed lead-lag) and the steam valve (dotted plain, dash-dotted lead-lag) in the same two runs. The valve ramps alike in both, and the lead-lag works by holding the feed back. (c) The heat the tail gas and the makeup bring to the burners in case 5, whose adsorber cycle lengthens from 10 to 11 min at 5 min: the tail gas solid on Loops and dashed with the tail-gas feedforward (FF), the makeup dotted on Loops and dash-dotted with FF. Data written from the engine by a measurement script.

3136, the hottest tube rises from 898.5 to 914.2 °C, and the true slip falls from 6.09 to 5.07 %, which the chromatograph confirms as 5.18 at 30 min and 4.95 at 35 min.

12.5.3 A Cyclic Unit on a Continuous Header

In case 4 the adsorber's product swings from 43,213 to 53 693 Nm^3/h each step as the repressurizing bed draws on it, and its tail gas leaves the beds at 25,780 to 42 997 Nm^3/h. Over 20 to 30 min the 200 m^3 surge drum swings 0.221 bar and the header 0.030 bar, against 0.390 and 0.052 with 50 m^3 and 0.132 and 0.018 with 400 m^3. The swing falls less than in proportion to the volume because the drum's time constant, 0.13 to 1.03 min, is short against a repressurization that draws product for most of a 2.5 min step. The tail-gas drum holds 1.240 to 1.267 bar a and sends 28,392 to 29 049 Nm^3/h to the burners, so the furnace barely sees the pulses.

12.5.4 Recovery Is Firing

Case 5 lengthens the cycle from 10 to 11 min at 5 min with the header on Manual. On Loops the recovery rises from 85.01 to 86.37 %, the tail gas's heat falls from 73.19 to 72.31 MW and its heating value from 9.274 to 9.229 MJ/Nm^3, the makeup rises from 2862 to 3102 Nm^3/h, 8.4 %, and the purity falls from 99.959 to 99.922 % by 30 min, as shown in Figure 12.2c. The outlet dips to 857.37 °C at 16.7 min before the loop catches up. With the feedforward the makeup moves as the tail gas's meters change and the outlet stays within 0.23 K of its setpoint.

> **TRY IT IN THE APP** **5 · A longer cycle: recovery is a firing change**
>
> Load the case and press Run on Loops. At 5 min the adsorber's cycle lengthens from 10 to 11 min. The recovery rises from 85.01 to 86.37 %, the makeup fuel from 2862 to 3102 Nm^3/h, and the outlet dips to 857.37 °C. Choose Loops + tail-gas feedforward as the Control architecture and press Reset: the outlet now stays within 0.23 K of its setpoint.

Case 8 fails a bed at 5 min, and the adsorber closes to three beds at the next step. On Loops the recovery falls from 85.03 to 80.98 %, the tail gas's heat rises from 73.24 to 82.12 MW and its heating value from 9.277 to 9.367 MJ/Nm^3, and the outlet loop cuts the makeup from 2864 to 2382 Nm^3/h, 16.8 % less (16.6 and 17.0 % at seeds 1 and 2). The plan for the application expected the makeup to

Table 12.4 The Eight-Hour Challenge at Five Seeds (20261005, 0 to 3)

Settings	$\lvert p_h - p_{SP} \rvert$ (bar)	Purity (%)	Lowest S/C	Hottest tube (°C)	Verdict
shipped	0.829–0.831	99.760–99.761	2.858–2.862	914.8–915.0	fails
suggested	0.392–0.395	99.927–99.929	2.884–2.886	916.3–916.7	passes
lead-lag off	0.312–0.317	99.930–99.932	2.711–2.725	914.4–914.7	fails
cycle fixed	0.371–0.373	99.671–99.673	2.888–2.889	912.9–913.0	fails
15 %/bar, 8 min	0.952–0.957	99.944–99.945	2.868–2.870	916.7–917.2	fails
on Loops	0.393–0.400	99.920–99.923	2.910–2.916	926.9–927.1	fails
MPC[a]	0.432–0.580	99.938–99.941	2.862–2.868	921.6–922.4	3 of 5

[a]With the lead-lag on and the cycle following the feed. Passes at the default seed (0.463 bar), seed 0 (0.437) and seed 3 (0.432).

climb, and the balance does the opposite: the lost hydrogen is burned. The header loop raises the feed from 18,730 to 19 640 Nm^3/h to deliver the same product, so the natural gas fed and fired per Nm^3 of hydrogen rises from 0.4323 in the first 5 min to 0.4403 from 30 to 45 min, and that is the number the case teaches. The richer tail gas first over-fires the furnace on Loops, the outlet reaching 865.2 °C at 16.3 min and the hottest tube 907.4 °C, while the feedforward holds the outlet within 1.66 K with the tube at 903.9 °C.

12.5.5 A Demand Step and a Demand Too Large

In case 6 the users draw 10 % more at 5 min. Under the shipped header tuning the header falls to 19.54 bar g at 16.2 min, the rate peaks at 115.7 % at 23.3 min, the header is back within 0.15 bar from 26.0 min, the worst error is 0.463 bar, the outlet dips to 854.4 °C, the makeup reaches 3537 Nm^3/h, the step shortens to 2.167 min, and the ratio stays at or above 2.956. At 40 % per bar the worst error is 0.245 bar (0.245 and 0.242 at seeds 1 and 2) and the ratio's low 2.919, and at 60 per bar and 3 min 0.152 bar and 2.874, so a firmer header loop buys pressure with ratio margin.

Case 7 asks for 40 % more and is infeasible by design. The rate reaches its 140 % maximum at 7.28 min, the firing peaks at 168.7 MW, the outlet sags to 834.0 °C at 11.0 min with the slip at 7.53 %, and the tube override takes the firing from 16.96 min, the hottest tube peaking at 919.0 °C. From 60 to 360 min the drum sends a mean 69 021 Nm^3/h to a demand of 70,006, and the header falls from 17.73 to 12.01 bar g, about 0.019 bar a minute. Across header gains of 30, 60 and 120 % per bar and outlet setpoints of 840, 860 and 880 °C the header spends 50.3 to 51.8 min of the first hour below 19.0 bar g and the rate is at its maximum in all nine. Shedding demand to 125 % at 12 min brings the header above 19.5 bar g by 20.5 min.

12.5.6 Eight Hours on the Header

The challenge lasts 8 h, with the demand 20 % up in half a minute at 1 h, the feed ethane-rich at 4 h, the demand back at 5 h and a bed failure at 6 h. Its target, checked against the simulator's truth at each physics step after the first minute, is the header within ±0.5 bar, the purity at least 99.9 %, the ratio never below 2.8 and the hottest tube never above 925 °C. The outcomes are in Table 12.4, and Figure 12.3 draws its four columns against their limits. As shipped the gentle header loop lets the demand step take the header 0.831 bar off at 69.6 min, and the fixed cycle loads each bed past breakthrough so the purity falls to 99.760 %. The suggested settings make four changes, the lead-lag, 60 % per bar and 3 min, the cycle following the feed and the tail-gas feedforward, and pass at every seed with the header within 0.395 bar. Each change is needed. Without the lead-lag the firm header loop moves the rate fast enough to drop the ratio to 2.711, with the cycle fixed the purity falls to 99.671 %, and with the shipped header tuning the header strays 0.957 bar.

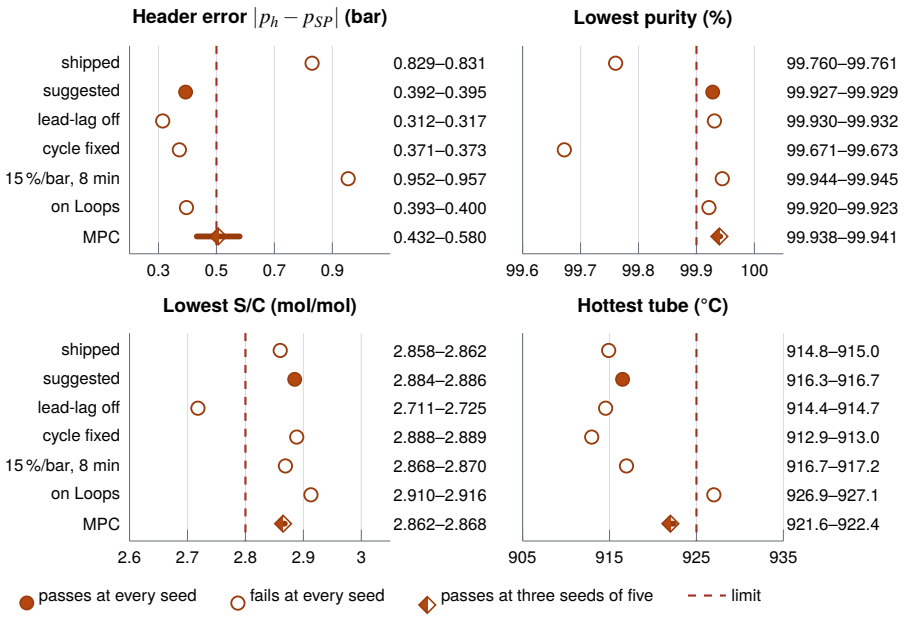


Figure 12.3 Table 12.4 drawn against the challenge's limits (dashed). Each bar runs from the lowest to the highest value over the five seeds, and rows three to six each undo one of the suggested settings. Notice that each of those four rows crosses a limit in a different panel: the lead-lag off on the ratio, the fixed cycle on the purity, the gentle header loop on the header and Loops on the tube. The shipped settings cross two, and the predictive controller straddles the header's limit.

> ✦ **DID YOU KNOW?** **The Steadiest Header Still Fails**
>
> From the simulator (Table 12.4). With the lead-lag switched off, the suggested settings hold the header closer than any other configuration, within 0.312 to 0.317 bar at five seeds. They still fail the challenge, because the firm header loop moves the rate fast enough to drop the steam-to-carbon ratio to 2.711.

On Loops the hottest tube passes 925 °C 11 to 13 times between 74.9 and 296.9 min across the five seeds, in a sustained cycle of 18.3 to 20.2 min in which, at the default seed, the rate swings from 98.8 to 140 % and the outlet from 831.0 to 884.2 °C. The firm header loop moves the rate, the outlet loop hears of it only when the outlet has cooled, and the two chase each other through the 120 % hours. A first account of the failure blamed the bed failure's richer tail gas, but after the failure the tube reaches only 911.8 °C, so the app's texts were corrected during its technical review. With the feedforward there is no cycle (rate 113.5 to 124.6 %, outlet 855.1 to 865.3 °C), nor with the shipped gentle header loop on Loops (rate 110.8 to 118.3 %). The cycle is a true interaction of a firm rate loop with a slow outlet loop and is the measured reason the feedforward is needed, so it was explained rather than detuned away.

The predictive controller, given the lead-lag and the following cycle, keeps the ratio at or above 2.862, the purity at or above 99.938 % and the tube at or below 922.4 °C, but its linear model hunts at 120 % load and the header strays 0.432 to 0.580 bar, so it passes at three seeds of five. Its weights were tuned by search to a header weight of 100 and a move weight of 30. Played at 120 min/s, the fastest speed, the eight hours take 240 frames and 2.3 s of computation headless. Case 2 at 10 min is shown in Figure 12.4.

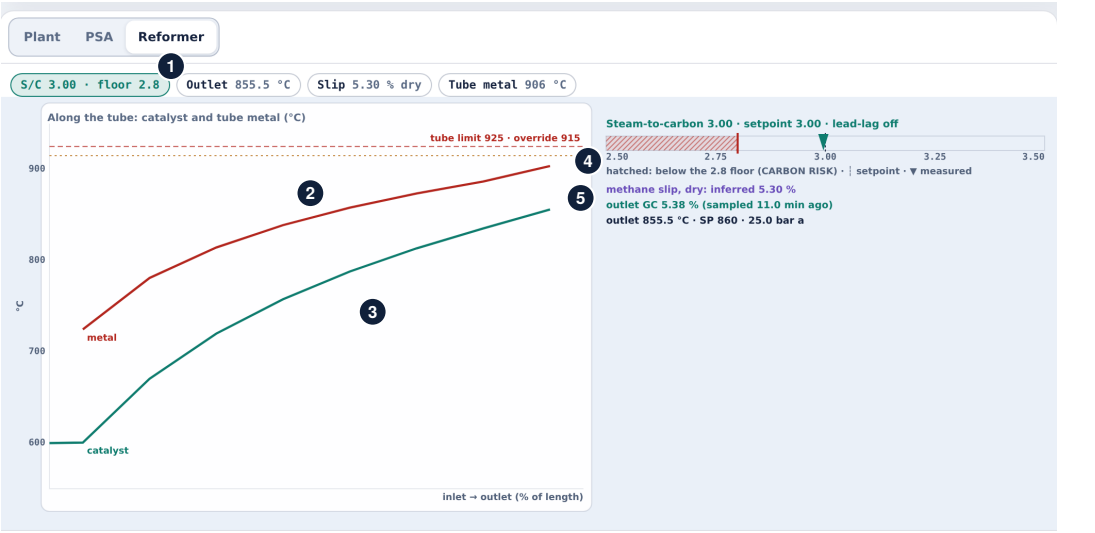


Figure 12.4 Case 2 as shipped (lead-lag off) at $t = 10$ min on the Reformer view, light theme, cropped to the plant view with numbered callouts. (1) The status chips. (2) The tube walls, up to 906 °C against the 925 °C limit (dashed). (3) The catalyst, from 601 °C at the inlet to 855.5 °C measured at the outlet. (4) The S/C scale, which shows 3.00 against its hatched floor with the lead-lag off. (5) The slip strip, with the inference at 5.30 % and the chromatograph's last report of 5.38, the true slip of 5.64 hidden while the teaching overlay is off.

12.6 Technology in Practice

Figure 12.5 draws the ratio station, the outlet loop with its override and the header loop of Section 12.4 with the hardware a refinery installs.

What Measures It

The ratio needs three readings: the feed flow, the steam flow and the carbon number of the feed. The carbon number comes from a chromatograph, which separates the species of a sample in time, so its cycle takes minutes and some dead time is inevitable [22]. The chapter's chromatograph samples every 5 min and reports 3 min later, and between reports the ratio station uses its last value. The override reads the tube walls through an infrared scan, whose 12 s lag and 1.5 K of noise in Section 12.3 are the noise that once made an ordinary low select ratchet the makeup to zero.

> **TECHNOLOGY** **Sensor: Flow Meters Behind the Ratio**
>
> The ratio station multiplies a measured flow, so the steam-to-carbon ratio is only as good as its meters. An orifice plate infers flow from the pressure drop across a restriction, which grows with the square of the flow, and it is commonly understood not to keep good accuracy much below a third of its full range, a turndown of 3 to 1. Modern differential pressure sensors stretch that to about 10 to 1. A vortex meter counts the vortices that a bluff body sheds, at a frequency in proportion to the flow, and so reads a wider range, but below a low-flow cutoff its signal falls to zero [22]. The chapter's feed, steam, makeup and tail-gas meters have a 2 s lag and 0.3 % of noise.

What Moves It

Four handles move the plant: the feed valve under its flow loop, the steam valve, the makeup fuel valve and the adsorber's switching valves, which a sequence opens and closes on the step schedule.

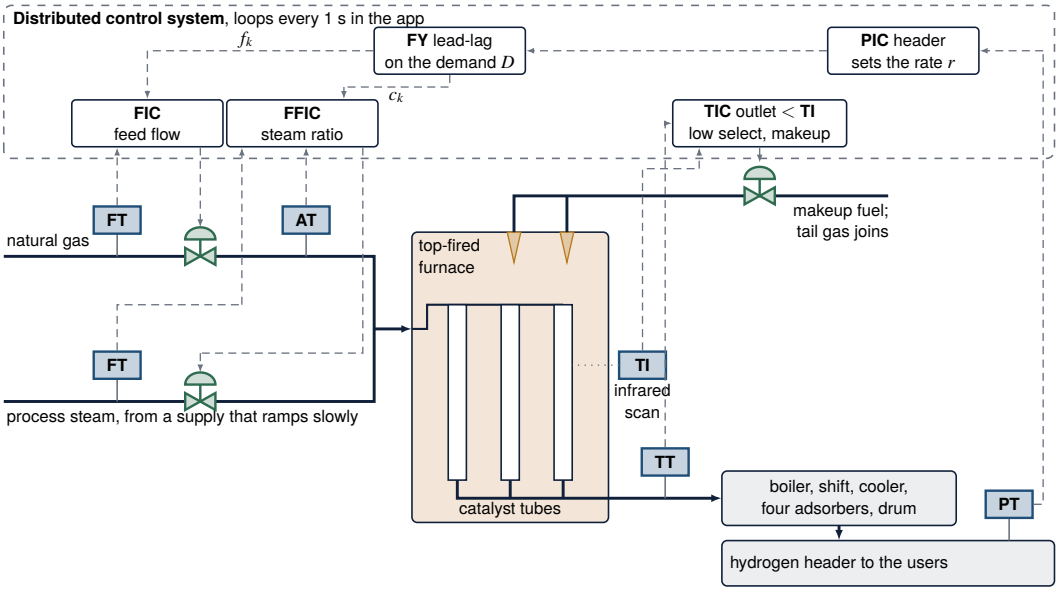


Figure 12.5 The ratio station, the outlet loop with its tube-metal override and the header loop with their hardware. Lettered boxes are sensing devices: the header pressure transmitter (PT), the feed and steam flow meters (FT), the feed chromatograph (AT), the outlet thermocouple (TT) and the infrared scan of the tube walls (TI). Valve symbols are final elements, and dashed arrows are signals. Notice that the lead-lag acts on the rate demand before it reaches either flow loop, and that the steam valve sits behind a slow steam supply. The adsorber, the tail-gas drum, the oxygen trim and the feedforward are left out. Not to scale, and not a piping diagram.

The steam valve is the slowest of them in this model, and the bed valve that fails in case 8 is one of the switching valves.

> **TECHNOLOGY** **Final Element: The Steam Control Valve**
>
> The process steam reaches the mixing tee through a control valve. A pneumatic valve with a spring takes a set position when its air fails, fixed by the actions of its actuator and its body: an air-to-open valve fails closed. Good practice chooses that position from the risks of the process, not from convention or habit. Friction in the stem makes a valve move in steps, a behavior called stiction, and leaves it deaf to small reversals of its signal, worst in valves that are poorly maintained or lack a positioner [22]. The chapter's valve follows its command through a 0.5 min lag and at most 4 % of travel a minute, and the ramp stands for the steam system behind the valve.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level oversees local controllers, loops are sometimes nested, and cycle times range from milliseconds to minutes [23]. The app executes its loops every second and its predictive controller every 30 s, inside that range. The header, steam flow, oxygen trim and tail-gas drum loops are PI blocks. The lead-lag is a dynamic function block, which modern control systems provide as digital function blocks [22], and the outlet loop and the override meet in a low select. Burners get the same order of moves from selectors instead: in a cross-limited air and fuel ratio the air leads when the firing rate rises and the fuel leads when it falls [22]. The lead-lag of Equation (12.8) obtains that order from dynamics.

> **IN PRACTICE** **Reading the Tubes Through the Peep Holes**
>
> At one industrial reformer studied in a 2008 thesis, the plant's data were logged every second, but the tube walls were read by hand. Each night the operators measured them at an upper and a lower peep hole with a hand-held infrared pyrometer and logged the readings, and a consultant surveyed the tubes every quarter [24]. The care has a reason. Reformer tubes are designed for a life of typically 100,000 hours, and by a rule of thumb a wall 20 °C hotter halves that life for a given alloy at its design pressure [24]. The reading itself needs judgment, because a surface's emittance is uncertain and it also reflects radiation from other hot bodies [22].

What Is Modernizing It

Three developments touch this plant. The first is the tube wall. The thesis behind the nightly readings built a model of the reformer for the on-line monitoring of tube-wall temperatures [24], and thermal imaging, a dense array of infrared sensors, shows hot spots as a picture where a pyrometer reads one spot [22]. The second is the inferred slip. Predicting a quality that is analyzed slowly is the original and still dominant use of soft sensors, and their predictions deteriorate as the process changes [25], as the aged catalyst of case 3 shows before its bias catches up. The third is predictive control, with more than 4600 applications reported and refining the largest share [21]. The chapter's controller plans the rate, the ratio and the outlet together, and it passes the challenge at three seeds of five.

12.7 Limitations

Nothing in the simulation was fitted to an operating unit. Its sizes, compositions, correlation and limits are teaching values that make one consistent design, a 50 000 Nm^3/h plant firing about 103 MW. The reformer is one tube at equilibrium with a declared approach, so maldistribution between tubes, kinetics near the inlet and the catalyst's own activity profile are absent. The firebox is one stirred volume, the convection section is not modeled, and the linear enthalpies over-fire the furnace by about 1.3 MW and put the shift outlet 3.5 K low, which no lesson depends on. The adsorber's recovery and purity come from a correlation whose directions match practice but whose numbers were not fitted, and its purge pressures do not enter it. The boiler's steam is not exported and its drum is not modeled. Carbon formation is represented by a floor, not by a kinetic criterion.

The controllers carry the plant's declared design values, and they meet a mismatch only where a case creates one. The predictive controller's model is linear at design, it hunts at 120 % load, and it passes the challenge at three seeds of five, which is why it is offered as a comparison. The lead-lag's necessity is proved by hand for case 2's 20 %-a-minute rate rise and measured for the challenge's, and a slower rise does not need it. The figures come from a single model run at the application's seed, and extra seeds and a halved integration step were tried only where noise or timing could change them.

> **OPEN CHALLENGE** **Capturing the Carbon of a Fired Reformer**
>
> In this chapter the tail gas is burned, so its carbon leaves in the flue gas. A federal study of 2022 found only a few hydrogen plants capturing carbon, at less than 90 %, and several in development above it. To capture more than 90 % of its feed's carbon, a steam methane reformer must treat both its process gas at pressure and the low-pressure flue gas of its externally heated furnace, a step that a reformer heated from inside does not need. The study also found that the absence of rules on such emissions limits capture [26]. Progress would be a plant that captures nearly all of its carbon while its furnace still burns the adsorber's tail gas.

12.8 Implementation History

The boiler's water and steam properties first lived in the flash-steam application of Volume 1, Chapter 6. The hydrogen plant's build moved that module unchanged in arithmetic into the shared framework, kept the flash-steam simulator's fifteen cases logging byte-identical runs before and after the move, and added a shared test that reproduces every verification value of the IAPWS release the

module computes, to 0.37 of the release's tolerance at worst. The hydrogen plant is the module's second user. After the build, three reviews by separate agents changed texts and the display but no behavior: every case's logged run kept the same hash through the usability, graphics and technical stages.

12.9 Conclusion

A steam methane reforming hydrogen plant is a ratio with a floor, a furnace that burns the adsorber's waste and a cyclic purifier on a continuous header, and this simulator logs all three side by side from an equilibrium reformer whose constants are within 1.6 % of the JANAF tables.

With the steam supply ramping at 4 % a minute, a plain ratio drops to 2.729 on a 15 % feed rise and no steam tuning of 64 lifts it past 2.763, while a lead-lag on the demand holds 2.994 by holding the feed back. The tail gas supplies 71.46 % of the fuel, so a longer cycle needs 8.4 % more makeup and a failed bed 16.8 % less while natural gas per hydrogen rises from 0.4323 to 0.4403.

The challenge needs the lead-lag, a firmer header loop, a cycle that follows the feed and the tail-gas feedforward at every seed, and without the feedforward a firm header loop and a slow outlet loop cycle the tubes past their limit. The model is lumped, its adsorber a correlation and its limits teaching values. A fuller model would add reforming kinetics near the inlet, the convection section and its steam, an adsorber model with breakthrough fronts, and a carbon criterion in place of the floor. The simulator shows that in this plant the order in which two streams move matters as much as the ratio between them.

KEY POINTS

- A hydrogen plant must raise its steam before its feed, because steam short of the carbon lets carbon form on the catalyst.
- With the steam ramping at 4 % a minute, a plain ratio drops to 2.729 on a 15 % feed rise, while a lead-lag on the demand holds 2.994 by holding the feed back.
- The tail gas supplies 71.46 % of the fuel, so a longer cycle needs 8.4 % more makeup and a failed bed 16.8 % less.
- Without the tail-gas feedforward, a firm header loop and a slow outlet loop cycle the tubes past their limit.

Where else this applies. Volume 3, Chapter 4 cross-limits fuel and air so that the air leads on the way up and follows on the way down, and Volume 6, Chapter 7 studies the recycle whose coupling the tail gas shares by returning as fuel.

References

[1] U.S. Energy Information Administration, "Hydrogen for refineries is increasingly provided by industrial suppliers," Today in Energy, January 20, 2016, eia.gov/todayinenergy/detail.php?id=24612, read November 2026.

[2] J. J. Alves and G. P. Towler, "Analysis of refinery hydrogen distribution systems," *Industrial & Engineering Chemistry Research*, vol. 41, no. 23, pp. 5759–5769, 2002.

[3] D. A. Latham, K. B. McAuley, B. A. Peppley, and T. M. Raybold, "Mathematical modeling of an industrial steam-methane reformer for on-line deployment," *Fuel Processing Technology*, vol. 92, no. 8, pp. 1574–1586, 2011.

[4] L. Lao, A. Aguirre, A. Tran, Z. Wu, H. Durand, and P. D. Christofides, "CFD modeling and control of a steam methane reforming reactor," *Chemical Engineering Science*, vol. 148, pp. 78–92, 2016.

[5] H. Khajuria and E. N. Pistikopoulos, "Optimization and control of pressure swing adsorption processes under uncertainty," *AIChE Journal*, vol. 59, no. 1, pp. 120–131, 2013.

[6] F. G. Shinskey and B. G. Lipták, "Ratio control," in *Process Control*. Elsevier, 1995, pp. 102–105.

[7] J. R. Rostrup-Nielsen and T. Rostrup-Nielsen, "Large-scale hydrogen production," *CATTECH*, vol. 6, no. 4, pp. 150–159, 2002.

[8] J. Rostrup-Nielsen and L. J. Christiansen, *Concepts in Syngas Manufacture*, ser. Catalytic Science Series. London: Imperial College Press, 2011.

[9] J. Xu and G. F. Froment, "Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics," *AIChE Journal*, vol. 35, no. 1, pp. 88–96, 1989.

[10] Occupational Safety and Health Administration, "OSHA technical manual, section IV, chapter 2: Petroleum refining process," U.S. Department of Labor, Washington, DC, osha.gov/otm/section-4-safety-hazards/chapter-2, read November 2026.

[11] D. L. Trimm, "Coke formation and minimisation during steam reforming reactions," *Catalysis Today*, vol. 37, no. 3, pp. 233–238, 1997.

[12] D. M. Ruthven, S. Farooq, and K. S. Knaebel, *Pressure Swing Adsorption*. New York: VCH Publishers, 1994.

[13] S. Sircar, "Pressure swing adsorption," *Industrial & Engineering Chemistry Research*, vol. 41, no. 6, pp. 1389–1392, 2002.

[14] S. Sircar and T. C. Golden, "Purification of hydrogen by pressure swing adsorption," *Separation Science and Technology*, vol. 35, no. 5, pp. 667–687, 2000.

[15] C. W. Skarstrom, "Method and apparatus for fractionating gaseous mixtures by adsorption," U.S. Patent 2 944 627, Jul. 12, 1960.

[16] J. L. Wagner, "Selective adsorption process," U.S. Patent 3 430 418, Mar. 4, 1969.

[17] R. Soltani, M. A. Rosen, and I. Dincer, "Assessment of CO_2 capture options from various points in steam methane reforming for hydrogen production," *International Journal of Hydrogen Energy*, vol. 39, no. 35, pp. 20 266–20 275, 2014.

[18] M. W. Chase, Jr., *NIST-JANAF Thermochemical Tables*, 4th ed., ser. Journal of Physical and Chemical Reference Data, Monograph 9. Washington, DC and Woodbury, NY: American Chemical Society and American Institute of Physics for the National Institute of Standards and Technology, 1998.

[19] W. Wagner, J. R. Cooper, A. Dittmann, J. Kijima, H.-J. Kretzschmar, A. Kruse, R. Mareš, K. Oguchi, H. Sato, I. Stöcker, O. Šifner, Y. Takaishi, I. Tanishita, J. Trübenbach, and T. Willkommen, "The IAPWS industrial formulation 1997 for the thermodynamic properties of water and steam," *Journal of Engineering for Gas Turbines and Power*, vol. 122, no. 1, pp. 150–184, 2000.

[20] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[21] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[22] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[23] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[24] D. A. Latham, "Mathematical modelling of an industrial steam methane reformer," M.Sc.(Eng.) thesis, Queen's University, Kingston, Ontario, Dec. 2008, hdl.handle.net/1974/1650.

[25] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[26] E. Lewis, S. McNaul, M. Jamieson, M. Henriksen, H. Matthews, L. Walsh, J. Grove, T. Shultz, T. Skone, and R. Stevens, "Comparison of commercial, state-of-the-art, fossil-based hydrogen production technologies," National Energy Technology Laboratory, Pittsburgh, PA, Report DOE/NETL-2022/3241, 2022.

12.10 Exercises

12.1 ○ Concept check · 2 min

In case 2 the same 15 % feed step drops the true steam-to-carbon ratio to 2.729 under the plain ratio, while with the lead-lag on it never falls below 2.994. What does the lead-lag change?

(a) It drives the steam valve faster than its ramp allows while the rate is rising

(b) It holds the feed back through a lag on the rate demand while the steam setpoint moves at once

(c) It raises the gain of the steam flow loop whenever the measured feed is rising

(d) It filters the feed meter so that the steam setpoint no longer follows its noise

12.2 Concept check · 1 min

In case 5 the adsorber's cycle is lengthened from 10 to 11 min while the plant rate is held. What happens to the natural gas makeup to the burners?

(a) It falls, because the plant recovers more hydrogen from the same feed

(b) It is unchanged, because the outlet loop holds the outlet temperature

(c) It rises, because the tail gas carries less hydrogen to the burners

(d) It rises, because the purity falls and the header loop raises the feed

12.3 Concept check · 2 min

In case 4 the 200 m^3 surge drum is replaced by one of 50 m^3. How does the drum's pressure swing from 20 to 30 min change?

(a) It grows four times, because the swing is inversely proportional to the volume

(b) It is unchanged, because the adsorber's product pulse does not depend on the drum

(c) It shrinks, because a smaller drum passes the pulse to the header more quickly

(d) It grows less than four times, because the pulse is long against the drum's time constant

12.4 Theory and calculation · 5 min

The design feed is 18 700 Nm^3/h of natural gas with a carbon number $n_C = 1.040$, at a steam-to-carbon ratio of 3.00. Take 22.414 Nm^3/kmol and 18.015 kg/kmol for water. The steam valve passes the design steam at 60 % open, its flow is proportional to its opening at a fixed supply pressure, and its stem moves at most 4 % a minute.

(a) Compute the carbon fed and the process steam at design, in t/h. *Checkpoint:* 867.7 kmol/h of carbon and 46.9 t/h of steam.

(b) The feed rises 15 % with the steam held, as in case 1. Compute the new ratio. *Checkpoint:* 2.609.

(c) Compute the valve opening that restores 3.00 at the new feed, and the shortest time the ramp allows to reach it. *Checkpoint:* 69.0 %, at least 2.25 min.

12.5 Theory and calculation · 6 min

The header holds $V_h = 1000\,m^3$ of hydrogen at 40 °C, and the surge drum $V_d = 200\,m^3$ at the same temperature. The drum passes the design product of 50 004 Nm^3/h to the header across a 2.5 bar drop, so k_{dh} is that flow divided by 2.5 bar. Take $R = 0.083\,14\,\text{bar}\,m^3/(\text{kmol}\,\text{K})$ and 22.414 Nm^3/kmol.

(a) Compute how fast the header pressure moves when the users draw 5000 Nm^3/h more than the drum supplies. *Checkpoint:* 0.0968 bar a minute.

(b) With the plant rate unmoved, how long would the case 6 step take to pull the header 0.5 bar down, and how fast does the challenge's 20 % step pull it? *Checkpoint:* 5.2 min, and 0.194 bar a minute.

(c) Compute the drum's time constant $V_d/(RTk_{dh})$ at 200 m^3 and at 50 m^3, and compare them with the adsorber's 2.5 min step. *Checkpoint:* 0.52 min and 0.13 min.

12.6 Theory and calculation · 7 min

Use the adsorber correlation of Equation (12.5) with four beds ($\ell_4 = 0.15$), a design step $t_{s,0} = 2.5$ min, $\varphi_0 = 0.80$, $y_0 = 400$ ppm and $w = 0.12$. Take the product's impurity as the step average of y_{imp}, and the hydrogen and the impurities fed in proportion to the feed.

(a) Compute the recovery and the purity for an 11 min cycle at the design feed, as in case 5. *Checkpoint:* 86.36 % and 99.922 %.

(b) The feed rises 20 % with the cycle fixed at 10 min, as in the challenge. Compute the recovery, the purity and the impurity at the end of a step, where $x = 1$. *Checkpoint:* 87.50 %, 99.848 %, 2428 ppm.

(c) With *cycle follows the feed* the step shrinks in proportion to the carbon fed. Compute the step, the loading and the purity at the same 20 % rise, with the carbon number unchanged. *Checkpoint:* 2.083 min, 0.80, 99.960 %.

12.7 App challenge · 8 min

Load scenario 2, Ratio control, then the lead-lag, and change no setting. The users draw 15 % more and the plant rate steps to 115 % at 2 min. Use the default seed, run to 12 min at the default speed of 2 min/s, and export the CSV file.

(a) Read the Lowest S/C tile, and find in the file the lowest true ratio, column `truth_sc`, and its time, column `time_min`. *Checkpoint:* 2.73 on the tile. 2.729 at 2.83 min, within 0.02 min.

(b) Set Lead-lag On on the S/C ratio card, press Reset, run to 12 min again and export the file. Read the Lowest S/C tile. *Checkpoint:* 2.99.

(c) In both files read the true feed, column `truth_feed_nm3h`, in the row at 3.0 min, and explain the difference. *Checkpoint:* 21 477 Nm^3/h plain and 19 527 Nm^3/h with the lead-lag, within 10 Nm^3/h.

12.8 App challenge · 8 min

Load scenario 8, A PSA bed fails: three beds, and change no setting. A bed valve fails at 5 min. Use the default seed, run to 45 min at 10 min/s, and export the CSV file.

(a) In the rows at 5.0 and 45.0 min read the recovery, `psa_recovery_pct`, and the true makeup, `truth_makeup_nm3h`. Compute the change in makeup. *Checkpoint:* 85.03 % and 2866 Nm^3/h, then 80.98 % and 2382 Nm^3/h, a change of −16.9 %.

(b) From the ledger columns `truth_feed_total_nm3`, `truth_makeup_total_nm3` and `truth_product_total_nm3`, compute the natural gas fed and fired per Nm^3 of hydrogen from 0 to 5 min and from 30 to 45 min. *Checkpoint:* 0.4323 and 0.4403, within 0.0005.

(c) Read the NG per H_2 tile at 45 min and say why it differs from the second value of (b). *Checkpoint:* 0.440.

12.9 App challenge · 10 min

Load scenario 9, Challenge, 8 h on the header. Before Run, set the Pressure gain to 60 % per bar and the Pressure integral time to 3.0 min on the PIC card, Lead-lag On on the S/C ratio card, Cycle follows the feed On on the PSA card, and Control architecture to Loops + tail-gas feedforward. Use the default seed and the top speed, 120 min/s, so that each 8 h run takes four seconds.

(a) Run 8 h and read the verdict at the end. *Checkpoint:* Target met: header within 0.394 bar, purity at least 99.929 %, S/C at least 2.886, hottest tube 916.7 °C.

(b) Set Lead-lag Off, press Reset and run 8 h again. Read the verdict, and find in the exported file the time of the lowest `truth_sc`. *Checkpoint:* Missed on S/C: 2.720, at 63.8 min, within 0.2 min.

Explain why the plain ratio fails with this header tuning, although the challenge as shipped, with the same plain ratio, keeps the ratio at 2.861 or above at the default seed.

12.10 Application area · 4 min

The chapter calls its steam-to-carbon floor of 2.8, its tube limit of 925 °C, its 99.9 % purity specification and its header alarm teaching values.

(a) Say where each value comes from on a real plant. *Checkpoint:* The catalyst supplier, the tube alloy's creep-rupture life, the users.

(b) Explain why a feed change on a reformer is a sequencing problem, and give the order of the two streams going up and going down. *Checkpoint:* Steam first up, steam last down.

12.11 Application area · 4 min

In case 8 a failed bed cuts the makeup natural gas by 16.8 %, and the natural gas fed and fired per Nm^3 of hydrogen rises from 0.4323 to 0.4403. The plant delivers its design product of 50 004 Nm^3/h.

(a) Compute the relative rise in natural gas per unit of hydrogen and the extra natural gas the plant uses each hour. *Checkpoint:* 1.85 %, and about 400 Nm^3/h.

(b) If all the carbon of that natural gas, $n_C = 1.040$, leaves in the flue gas, how much more carbon dioxide does the plant emit? Use 22.414 Nm^3/kmol and 44.01 kg/kmol. *Checkpoint:* About 0.82 t/h.

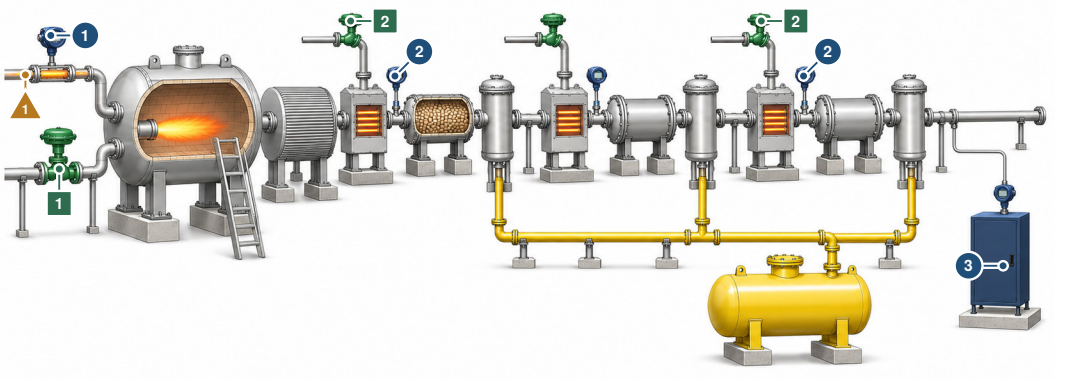
Say why a fuel report would describe this failure wrongly.

12.12 Review the agent · 5 min

An AI assistant is asked why the plain ratio of case 2 falls to 2.73 and replies as follows. “The steam flow loop is far too gentle at 0.5 % per percent and 1 min. Raise its gain to 3.00 % per percent and cut its integral time to 0.20 min. The steam will then track the ratio setpoint within seconds, and the plain ratio will stay above the 2.8 floor without any lead-lag.” Find the flaw in the claim and say how it could be checked.

CHAPTER 13

A Claus Sulfur Recovery Simulator for Feedforward, Dead Time and an Optimum Turned Into a Zero



The controller moves 1 combustion air valve, set by the feedforward and the trim
2 reheater valves that set the three bed inlet temperatures

The controller reads 1 acid gas flow meter 2 bed inlet temperatures 3 tail gas analyzer that reads the air demand

What disturbs it 1 the acid gas: its flow, its hydrogen sulfide and a hydrocarbon nobody measures

Notice the long sample line at the far right: the analyzer that sets the air reads gas that left the furnace minutes earlier, after three condensers. Illustration, not to scale.

IN PLAIN TERMS **Just Enough Air to Make Sulfur**

The main idea. A sulfur plant burns a third of a poisonous gas, hydrogen sulfide. The rest reacts with what burned to make liquid sulfur. An instrument at the end of the plant reads a value that is zero when the air is just right and shows which way it is wrong.

The challenge. That reading arrives minutes after the air has changed, so no tuning can shrink the first error when the gas changes. Measuring the gas coming in and setting the air from it ahead of time does most of the work.

Why it matters. One percent too little air can push a plant over its pollution permit.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the Did you know? box in the Introduction, the History boxes in Section 13.2, the Try it boxes and the second Did you know? box in Section 13.5, and Key points in Section 13.9. No equations.
Study	45 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	20 min	This page, Section 13.5 with Figure 13.2, Section 13.6 and Section 13.7 with its Open challenge box.

The control question Sulfur recovery has a maximum at one air rate, and a maximum has no sign, yet an ordinary PI loop holds the air of every Claus unit. What measurement makes that possible, and with the analyzer minutes downstream of the furnace, what do tuning, feedforward and a Smith predictor each do to the error a feed change leaves?

Learning objectives

- Explain how the air demand $y_{H_2S} - 2y_{SO_2}$ turns the recovery maximum into a zero a PI loop can hold, and price a percent of air from the stoichiometry.
- Compute the stoichiometric air of a feed, and the ratio the trim must find after an unmeasured change or cannot reach after a measured one.
- Show by a run that the first peak after a composition step is the same for every tuning, and say what shortens it.
- Compare no feedforward, flow feedforward and flow and analysis on the same disturbances, and read the result from the scorecard and tiles.
- Judge what a Smith predictor on the trim can and cannot remove, and why a drifting analyzer must be checked from outside the loop.

Prerequisites Sections 1.4.6, 1.4.7, 1.6.2 and 1.6.4, with Section 1.4.4 for the beds. Chapter 8 for a late analyzer in this volume.

The app https://apmonitor.com/apps/sulfur-recovery/

See also Volume 3, Chapter 4 for an oxygen trim on a fuel-air ratio and Volume 3, Chapter 6 for a feedforward with a slow ratio trim around an optimum. Volume 5, Chapter 3 for a loop on a slow analyzer and Volume 5, Chapter 5 for a Smith predictor built from a process model.

13.1 Introduction

A sulfur plant is judged on the second decimal place of its recovery, and the recovery has a maximum. One third of the hydrogen sulfide is burned to sulfur dioxide in a reaction furnace and the rest reacts with it, two to one, to make sulfur. Any other air rate leaves one of the two in excess and that excess cannot react with anything. Measured on the unit modeled here, recovery is 97.619 % at an air ratio of 1.000, 96.640 % at 0.990 and 96.640 % again at 1.020: one percent short of air costs exactly what two percent of surplus costs, because too little air strands hydrogen sulfide one for one while too much strands only half as much sulfur dioxide.

> **+ DID YOU KNOW? Sulfur From 109 Plants in 2005**
>
> In 2005 the United States produced 8.8 million metric tons of elemental sulfur, recovered at 109 petroleum refineries, natural gas processing plants and coking plants. About 90 % of the sulfur consumed was used as sulfuric acid [1]. About 90 to 95 % of recovered sulfur is made by the Claus process [2].

A maximum has no sign, and a fixed-sign controller cannot sit on one. That is the argument Volume 12, Chapter 6 makes at length, where a chloride optimum defeats an otherwise competent PID. Yet the air on every Claus unit in the world is on a PI loop and it works. The reason is not a better controller. It is a better measurement. The tail gas analyzer reports the air demand

$$D = y_{H_2S} - 2y_{SO_2}$$

which is zero at the recovery maximum, positive when the furnace is short of air and negative when it has too much. The optimization has been turned into a regulation by changing what is measured, and that move is worth more than any amount of search machinery. Section 1.4.6 sets this measurement beside the volume's other analyzers, and Section 1.4.7 among the few refinery optima that lie inside the range rather than on a limit.

The misconception the simulator corrects is that an optimum is an optimization problem. Faced with an objective that has a maximum, the first question is not what optimizer to build but whether

some measurable quantity is zero there: a difference, a ratio against a stoichiometric number, a derivative somebody can instrument. Where one exists the problem collapses to a loop.

Finding it does not make the problem easy. The analyzer sits at the far end of the plant behind three condensers, 3.2 min of transport and cycle away from the furnace it is controlling. Measured, a step in acid gas composition produces a peak air demand error of 2.67 mol% whether the loop is sluggish, correctly tuned or oscillating, because the error arrives with the step and for the next 3.2 min the controller has no information that anything has changed. On a dead-time-dominant loop, tuning changes how long an error lasts and not how big it gets [3]. The things that shorten the peak are feedforward and a faster measurement, which is engineering rather than tuning.

This chapter works with an interactive Claus unit whose stoichiometry, analyzer delay and dew point correlation are hidden, so that the reader discovers the air demand relationship, measures the division of labor between feedforward and feedback rather than taking it on trust, and watches what a drifting analyzer does. The Claus chemistry, feedforward and air demand control are all established practice. Section 13.2 gives the history of the process and of its control. Section 13.3 writes out the three stage Claus train, with its furnace, its equilibrium limited converters, its sulfur dew point fouling and its analyzer faults, in the detail needed to rebuild it. Section 13.4 describes the controllers, from the feedforward and its trim to the dew point override and the optional Smith predictor. Section 13.5 reports the scenario sequence, in which the same disturbance is answered about 168 times better by a ratio station than by any tuning, and the graded challenge. Section 13.6 sets the hardware of a real plant beside the model. Sections 13.7 and 13.8 state the limitations and the three results that carry past sulfur plants, and Section 13.9 sums up.

13.2 Background

Acid gas from an amine regenerator is mostly hydrogen sulfide and carbon dioxide, and something must be done with it. The Claus process burns part of the hydrogen sulfide and reacts the remainder with the sulfur dioxide produced, first thermally in a reaction furnace and then over alumina in a series of catalytic converters with a condenser after each. Modern plants recover between 95 and 98 percent of the sulfur before tail gas treating, and the balance leaves as sulfur dioxide from a stack under a permit.

> ⧖ HISTORY **1883 and 1937: One Catalytic Step Becomes Two**
>
> Carl Friedrich Claus patented the original process in 1883. It partially oxidized hydrogen sulfide over a catalyst at 200 to 350 °C, making sulfur and water in a single step. In 1937 the process was divided into two parts. A flame burned one third of the hydrogen sulfide completely to sulfur dioxide, and that sulfur dioxide then reacted over a catalyst with the remaining hydrogen sulfide at 170 to 370 °C [4]. Most modern plants that treat acid gas are built on this modified process [4], and the furnace and three converters of this chapter follow its two parts.

Three technical literatures bear on the control problem. The reaction furnace itself is the least understood part of the unit and has been reviewed as such, because the chemistry at 1200 °C in under a second involves far more than the two nominal reactions [5]. The catalytic section is an equilibrium problem, and process modeling of it has been used to evaluate modifications aimed at raising recovery and reducing emissions [6]. Alternative configurations, including high temperature air combustion and routes that recover hydrogen as well as sulfur, continue to be proposed [7, 8]. The environmental driver behind all of it is direct: sulfur dioxide from refinery sulfur plants is a measured contributor to local air quality and the subject of its own assessments [9].

Control developed around the analyzer. Before continuous tail gas analysis the air was set from the acid gas flow with a ratio a supervisor adjusted after a laboratory result, and the unit ran deliberately rich because the penalty for being lean was larger. The air demand analyzer changed that. Once the difference between the hydrogen sulfide and twice the sulfur dioxide could be measured continuously,

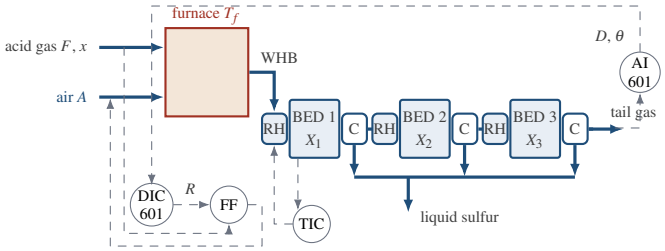


Figure 13.1 Claus unit U-601. Acid gas F of hydrogen sulfide fraction x burns with air A in a reaction furnace at T_f, followed by a waste heat boiler WHB. Three reheaters RH and catalytic beds of conversion X_i each precede a condenser C that removes liquid sulfur. AI-601 measures the tail gas air demand D after a dead time θ. The feedforward FF computes the air from the acid gas flow and composition and multiplies it by the ratio R that the trim controller DIC-601 sets. A temperature loop TIC, drawn for bed 1, holds each bed inlet on its reheater.

the whole problem became a trim loop on a stoichiometric feedforward, and the structure has not changed since. The oxygen trim on the air-to-fuel ratio of the fired heater in Volume 3, Chapter 4 has the same structure. What has changed is the feedforward. Plants with an acid gas chromatograph compute the air demand from the measured composition rather than from a commissioning number, and plants without one do not.

> ⧖ HISTORY **1957 and 1977: The Analyzer Takes Over the Air**
>
> A patent filed in 1957 found the ratios of a sulfur plant hard to hold, largely because the composition of its feed gas varied. Its remedy was to analyze a sample of the process gas for the ratio of hydrogen sulfide to sulfur dioxide and let the result reset the flow controller of the furnace air. The analyzer was a chromatograph: a timer injected the sample into a packed column at intervals, and the two gases left the column at different times [10]. A patent filed in 1977 had the structure of this chapter. A PI controller on the tail gas ratio multiplied the nominal air for a known volume of acid gas, and that product was multiplied by the measured acid gas flow to give the total air [11].

The dead time was and remains the difficulty. A sample point closer to the furnace shortens it and costs a hot, wet, corrosive sample line that plugs, so the trade is an instrumentation decision rather than a control one. Dead-time compensators exist and are surveyed [3], and on a plant whose disturbances are mostly composition rather than setpoint changes the feedforward is usually the better investment. Above the regulatory layer, the constraint and economic arbitration that a large unit needs is provided by multivariable control [12], as in the layers of Section 1.4.7, and the plantwide question of which unit owns which specification is the same one the fractionation train of Volume 10, Chapter 10 poses [13].

13.3 Process Model

Unit U-601 takes 8.0 kmol/min of acid gas at 85 mol% hydrogen sulfide with one percent hydrocarbon, and 17.33 kmol/min of air for a ratio of exactly one. The flame sits at 1226 °C, the thermal stage makes 63.5 % of the sulfur, and three converters with inlets at 245, 210 and 200 °C take the recovery to 97.619 %, leaving 7.46 t/d of sulfur to the stack against a permit of 9.0 t/d. Figure 13.1 shows the arrangement, and Table 13.1 classifies its variables.

13.3.1 Stoichiometry, and Where It Leaves the Plant

With R the ratio of air fed to air demanded, the furnace burns $R/3$ of the hydrogen sulfide and leaves $1 - R/3$. The Claus reaction wants those two in a ratio of two to one, which happens at $R = 1$ and

Table 13.1 Classification of Variables at the Design Operating Point

Symbol	Role	Nominal	Range
D (AI-601)	CV, analyzer	0.00 mol%	−1–1.5 SP
$T_{in,1..3}$	CV, continuous	245, 210, 200 °C	180–275
R	MV	1.000	0.90–1.10
A	MV	17.33 kmol/min	0–40
reheaters	MV	245, 210, 200 °C	150–285
F	DV, measured	8.0 kmol/min	3–13
x	DV, analyzed	85.0 mol%	45–95
hydrocarbon	DV, unmeasured	1.00 mol%	0–4
ammonia	DV, unmeasured	0 mol%	0–6
analyzer bias	DV, unmeasured	0	±1.5
T_f, $T_{out,i}$, activity	states	—	—

nowhere else, so whichever is in excess is sulfur that cannot react, Equation (13.1).

$$\text{loss} = \begin{cases} 1 - R, & R < 1 \\ \frac{1}{2}(R-1), & R > 1 \end{cases} \tag{13.1}$$

The asymmetry in Equation (13.1) is worth noticing, because every operator's instinct to run rich follows from it, before adding that excess oxygen sulfates the catalyst.

The air the furnace actually demands is a stoichiometric calculation, Equation (13.2).

$$A_{st} = \frac{F}{0.21}\left(\tfrac{1}{2}y_{\mathrm{H_2S}} + \tfrac{3}{4}y_{\mathrm{NH_3}} + 3\,y_{HC}\right) \tag{13.2}$$

Only the first two terms of that are measured on most plants, and the third never is.

13.3.2 Equilibrium Limited Beds

Each converter is adiabatic and equilibrium limited, which makes it the opposite of the hydrocracker beds of Chapter 7. More conversion is more heat and more heat is a worse equilibrium, so the bed limits itself, Equations (13.3) and (13.4).

$$T_{out} = T_{in} + K\,\dot{n}\,X, \qquad X = X_{eq}\left(\tfrac{T_{in}+T_{out}}{2}\right)a \tag{13.3}$$

$$X_{eq}(T) = X_0 - k_X\,(T - 200\,^\circ\mathrm{C}) \tag{13.4}$$

with $\dot{n}$ the reactable sulfur reaching the bed in kmol/min and a the bed activity. Because $\mathrm{d}X_{eq}/\mathrm{d}T < 0$ the residual has exactly one root and it is stable, so a Claus converter cannot run away. In the terms of Section 1.4.4, its runaway index is negative and its steady gain from inlet to outlet is less than one.

13.3.3 The Sulfur Dew Point

What a converter can do is drown. Below the sulfur dew point liquid sulfur condenses in the pores, Equation (13.5), and the activity decays, Equation (13.6).

$$T_{dew} = T_0 + k\,\log_{10}\left(\frac{y_S}{0.01}\right) \tag{13.5}$$

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \begin{cases} -k_f\,(T_{dew} - T_{in})\,a, & T_{in} < T_{dew} \\ +k_r\,(1-a), & \text{otherwise} \end{cases} \tag{13.6}$$

Recovering the activity takes a hot gas soak measured in days, so the damage is delayed and large while the benefit of running colder is immediate and small.

Table 13.2 Model Parameters

Symbol	Meaning	Value	Class
thermal	furnace conversion, net of COS	65.5 %	adjustable
X_0	equilibrium at 200 °C	0.800	adjustable
k_X	equilibrium lost per °C	0.0030	adjustable
K	rise per kmol/min converted	45.0 °C	adjustable
T_0, k	dew point correlation	182, 30.0	adjustable
k_f	activity lost below the dew point	0.055/(°C h)	adjustable
COS made	sulfur as COS and CS_2	3.0 %	adjustable
T_h	half hydrolysis, bed 1 outlet	305 °C	adjustable
θ	analyzer dead time	3.2 min	adjustable
T_{NH_3}	ammonia destruction	1250 °C	adjustable
permit	sulfur to the stack	9.0 t/d	visible

13.3.4 Carbonyl Sulfide and the First Bed

Carbon dioxide ballast in the acid gas makes carbonyl sulfide and carbon disulfide in the furnace, and only the first bed hydrolyzes them back, Equation (13.7).

$$h = \frac{1}{1 + \exp[-s\,(T_{out,1} - T_h)]} \tag{13.7}$$

That single term is why the first bed wants to run hot while the other two want to run cold, and it is the subject of scenario 10.

13.3.5 Assumptions

The thermal stage converts a fixed fraction of whatever the stoichiometry allows. Real furnace chemistry depends on residence time, burner mixing and configuration, and is the least settled part of the unit [5].

All sulfur species are lumped as sulfur atoms. The real vapor is an equilibrium between several allotropes and the dew point depends on which.

Each converter is one lumped adiabatic equilibrium stage, so there is no axial profile and no mass transfer limit.

Condensers are perfect and there is no tail gas treating unit, which is what takes a 97 percent Claus unit to 99.9 percent overall [6].

Carbonyl sulfide and carbon disulfide are one lumped species hydrolyzed only in the first bed. Real hydrolysis also occurs in the second bed on the right catalyst.

The analyzer dead time is a pure transport delay followed by a first-order average of 0.8 min, with no dispersion in the sample line, and the instrument drifts but never fails outright. Real sample systems plug, and a plugged sample system reports the last thing it saw.

13.4 Controller Design

13.4.1 Velocity Form PID

Every loop uses the shared velocity-form PID of Section 1.6.2, with the derivative on the measurement, executed every $\Delta t = 0.05$ min. Equations (13.8) to (13.11) restate it in the symbols used below.

$$e_k = a\,(\mathrm{SP}_k - \mathrm{PV}_k) \tag{13.8}$$

$$P_k = K_c e_k, \qquad D_k = -K_c \tau_D\, a\, \frac{\mathrm{PV}_k^f - \mathrm{PV}_{k-1}^f}{\Delta t} \tag{13.9}$$

$$s_k = s_{k-1} + (P_k - P_{k-1}) + (D_k - D_{k-1})$$

$$c_k = s_k + I_{k-1} \tag{13.10}$$

$$
\begin{aligned}
u_k &= \mathrm{clamp}(c_k, u_{\min}, u_{\max}) \\
I_k &= I_{k-1} + \frac{K_c}{\tau_I} e_k \Delta t + \frac{\Delta t}{\tau_t} (u_k - c_k)
\end{aligned}
\tag{13.11}
$$

where a is the controller direction, s_k accumulates the proportional and derivative increments without limit, c_k is the output the algorithm demands, and the integral used at scan k was accumulated through scan $k-1$. These are Equations (1.16) to (1.19) with $s_k = S_k$ and $c_k = v_k$. The integral state I also carries the starting output, so the bias term u_b is zero. Anti-windup is the back-calculation term of Equation (13.11), with tracking time $\tau_t = \tau_I$ [14]. Every loop here runs with $\tau_D = 0$, so in practice all four are PI. The trim DIC-601 has $a = -1$ and limits of 0.90 and 1.10 on the ratio, and the reheater loops have $a = +1$ and limits of 150 and 285 °C on the reheater outlets. A loop in manual sets $I_k = u_{\text{man}} - s_k$, so the switch to automatic is bumpless. The 0 to 40 kmol/min limit on the air flow is applied after the trim's ratio has been multiplied into the feedforward and is not seen by the trim's anti-windup, but it is reached only near the top of the flow and composition ranges. Integration is fixed-step fourth-order Runge-Kutta at 0.01 min, on the engine clock of Section 1.6.1.

13.4.2 Feedforward, and What It Is Allowed to Know

The air flow is computed from Equation (13.2) and multiplied by the ratio the trim controller sets. Three settings differ in what the calculation uses. *None* holds the selected base air flow, which the feedback trim still multiplies. The base does not follow the trimmed output, avoiding a repeated multiplication on each scan. *Flow* uses the measured acid gas rate with the composition the unit was commissioned with. *Flow and analysis* uses the composition an acid gas chromatograph reports through a first-order lag of 2.0 min. In none of the three does the calculation measure the ammonia or the hydrocarbon. It uses commissioning values for both, set in the app's advanced feedforward settings, because no plant has an on-line analyzer for amine carryover. Everything the feedforward is wrong about is what the trim loop is for.

13.4.3 The Trim Loop and Its Authority

DIC-601 has $a = -1$, direct acting in the convention of Section 1.6.2, because more air is more sulfur dioxide and therefore a lower air demand. The process gain is about −36 mol% of air demand per unit of ratio, so the suggested gain of 0.022 is a loop gain of 0.8, which is about as far as a loop with 3.2 min of dead time and integral action can be pushed [3, 15].

The trim range is 10 % either side on purpose, because an analyzer that fails high should not be allowed to swing the air by a quarter. The same limit that stops the analyzer doing harm stops the trim doing good, which is the subject of scenario 8.

🕮 **Pause.** Try Exercise 13.1 before reading on.

13.4.4 The Dew Point Override

A one-directional override raises a bed inlet setpoint to the sulfur dew point plus a guard of 10 °C whenever the setpoint is closer to the dew point than that, up to the 275 °C top of the setpoint range. It can only push up, because there is no amount of conversion worth a wet catalyst bed, and because the damage is delayed and the benefit is immediate.

13.4.5 Smith Compensation of the Air Demand Trim

The optional Smith predictor compensates the modeled response to commanded air while retaining the stoichiometric feedforward and the three reheater loops. Its feedback variable remains signed air demand D in mol%. A Smith correction separates the command response from the measured residual [3, 16]. The nominal command-to-analyzer model is Equation (13.12).

$$
\widehat{G}_0(s) = \frac{K_S}{(0.25s+1)(0.8s+1)}, \qquad \widehat{G}_\theta(s) = e^{-\widehat{\theta} s} \widehat{G}_0(s)
\tag{13.12}
$$

Here s has units of inverse minutes, $K_S = -36.3728042448$ mol% per unit ratio, and $\widehat{\theta} = 3.2$ min. The 0.25 min air lag precedes transport, and the 0.8 min analyzer lag follows it. The delay-free branch retains both lags.

The gain is a centered finite secant at ratios 1 ± 10^{-5}, using declared chemistry at 8 kmol/min, 85 mol% hydrogen sulfide, 1 mol% hydrocarbon, zero ammonia, unit activities and bed inlets 245, 210 and 200 °C. Its limiting derivative is −36.3729014, a difference of 2.67 parts per million. Applied to in-run ratio steps of −0.005 and +0.005, the executed frozen model has root mean square analyzer errors of 0.000489 and 0.000 937 mol% over 40 min. A separate optimized second-order-plus-dead-time fit gives $K = -36.3251$, lags 0.800793 and 0.247 454 min, effective delay 3.227 694 min, and RMS 0.000 009 76 mol%. The effective fitted delay includes sample-grid phase. That smaller residual describes the fitted model, not the implemented one.

The model advances on the original 0.05 min scan using the command actually delivered over the preceding interval. Its input u is applied air divided by the controller's own feedforward base, or the held base when feedforward is off. It is distinct from the physical ratio $R = A/A_{st}$. For model air state x_k, initial ratio u_0, and incremental demand v_k, the normal-scan recursions are Equation (13.13).

$$\begin{aligned} x_k &= u_{k-1} + (x_{k-1} - u_{k-1})e^{-\Delta t/0.25} \\ v_k &= K_S(x_k - u_0), \qquad b = e^{-\Delta t/0.8} \\ \widehat{D}_{0,k} &= b\widehat{D}_{0,k-1} + (1-b)v_k \\ \widehat{D}_{\theta,k} &= b\widehat{D}_{\theta,k-1} + (1-b)v(t_k - \widehat{\theta}) \\ D_{c,k} &= D_{m,k} + \widehat{D}_{0,k} - \widehat{D}_{\theta,k} \end{aligned} \tag{13.13}$$

The timestamped history interpolates v and supplies zero incremental demand before its first sample. The two analyzer states initialize to zero, while x starts at u_0. Here D_m is the actual analyzer reading and D_c is the corrected feedback. The residual $D_m - \widehat{D}_\theta$ carries unknown loads into the PI after detection. Neither the hidden feed nor the actual stoichiometric requirement enters the predictor.

A separate PI with the same action $a = -1$ uses D_c in the velocity recursion already given, with the same 0.90–1.10 trim bounds. An aggregate-lag internal model control (IMC) tuning approximation [17] sets $\tau_I = 0.25 + 0.8 = 1.05$ min and $K_c = 1.05/(|K_S|\lambda)$. The fixed $\lambda = 2$ min gives $K_c = 0.014433861$ ratio per mol%. This approximates the two lags for tuning and does not exactly invert Equation (13.12). Integral action removes measured offset and cannot identify sensor bias.

While Smith acts, ordinary PI tracks delivered trim using actual analyzer history. Smith entry and recovery reseed its own corrected-measurement history and hold the first move. The independent 0 kmol/min to 40 kmol/min air bound remains authoritative. A nonfinite packet, state or prediction first holds delivered trim within its bounds, then transfers to tracked ordinary PI with a recorded fault. Numerically valid model mismatch does not trigger this transfer. If previous delivered trim is below 0.90 under severe air clipping, its bounded hold cannot preserve the same command. All histories are local to a run and are rebuilt on restart.

13.5 Results

Thirteen scenarios run from a manual ratio sweep to a graded challenge. Table 13.3 lists them, and Section 1.6.4 describes their completion rules, the scorecard and the exported log.

13.5.1 What a Percent of Air Costs

Table 13.4 reports the air ratio swept either side of stoichiometric.

One percent short of air costs exactly what two percent of surplus costs, and the air demand measurement passes through zero at the maximum with a gain of about −36 mol% per unit of ratio. The permit is 9.0 t/d, so one percent of air is the difference between a plant inside its license and a plant outside it. Run as scenario 1, the steps to 0.990 at 20 min and to 1.020 at 60 min settle on

Table 13.3 Scenario Sequence

#	Scenario	Teaching point
1	Two to one	what a percent of air costs
2	A measurement with a sign	optimum into a zero
3	Three minutes of nothing	dead time and peak error
4	Tuning against dead time	the limit cycle
5	Feedforward on the flow	measured disturbances
6	Feedforward on the composition	analyzed disturbances
7	The disturbance nobody measures	what feedback is for
8	When the trim runs out of room	authority
9	The analyzer that lies	sensor drift
10	Three beds, three windows	middle against end
11	A wet bed	slow constraints
12	A trim that remembers what it sent	Smith compensation
13	Challenge	graded run

Table 13.4 Recovery and Stack Against Air Ratio

R	Recovery (%)	Stack (t/d)	D (mol%)
0.980	95.661	13.60	+0.736
0.990	96.640	10.53	+0.366
1.000	97.619	7.46	0.000
1.010	97.130	8.99	−0.360
1.020	96.640	10.53	−0.712

Table 13.5 Three Tunings on One Composition Step

K_c	τ_I (min)	Peak $\lvert D\rvert$	IAE	Ratio band
0.010	20	2.67	153	settles slowly
0.022	12	2.67	42	settles
0.040	8	2.67	335	cycles, 0.064 wide

the values of Table 13.4, and the stack permit alarm comes up 0.19 min after each of them. It clears only when the stack falls below 7.65 t/d, 85 % of the permit, which happens for 0.18 min as the ratio passes 1.000 at 60 min and then not until 100.71 min, after the return to stoichiometric.

13.5.2 Tuning Does Not Reduce the Peak

Table 13.5 reports three tunings against a step in acid gas composition from 85 to 78 mol% at 20 min with the flow feedforward running. The runs are scenario 3 with the gains set before the run, and the last row is also scenario 4 as loaded.

The peak is 2.67 mol% for all three, and it arrives with the step, at 20.01 min, because the composition moves the true air demand at once. It then holds to within 0.01 mol% until 23.3 min, because the analyzer reading first moves at 23.26 min and until then no gain and no integral time has anything to act on. A faster measurement or a composition feedforward does not lower that first value either, but both shorten it. One minute after the step the error is 1.70 mol% with the dead time set to zero and 1.87 mol% with the composition analyzer in the feedforward, against 2.67 mol% for every tuning at 3.2 min. The integrated absolute error in mol% min, taken over the 380 min following the step at the 0.01 min model step, separates the tunings: correct tuning is 3.6 times better than sluggish tuning and eight times better than tuning that cycles. The sluggish loop is still 0.010 mol% from setpoint at 400 min. The gain of 0.040 never settles at all, swinging the air ratio between 0.900, the trim's lower stop, and 0.964, a band of 0.064 that is about seven percent of the air.

Table 13.6 Integrated Absolute Air Demand Error Over 380 min Following Each Disturbance

Disturbance	None	Flow	Flow and analysis
Flow 8.0 kmol/min to 9.6 kmol/min	1227	7.3	7.3
Composition swings, 4 h	—	163	69
Hydrocarbon 1.0 mol% to 1.6 mol%	—	21.6	21.6

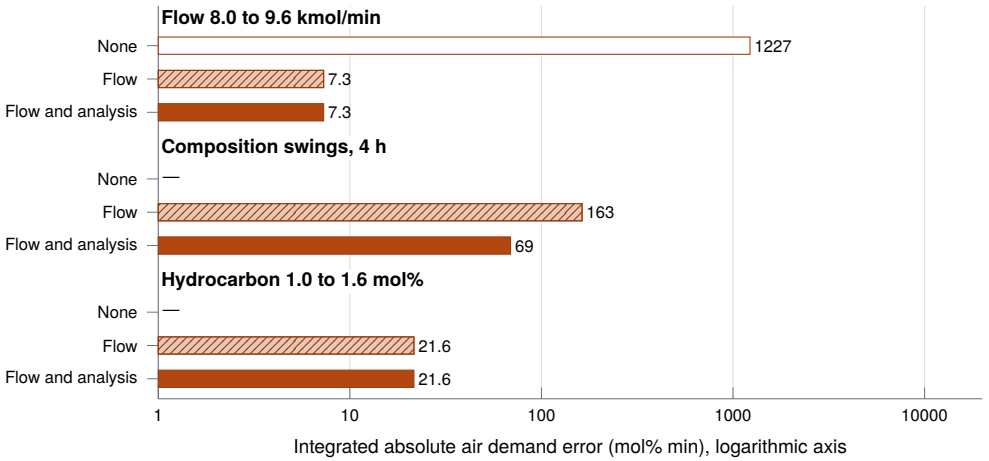


Figure 13.2 The integrated absolute air demand errors of Table 13.6 on a logarithmic axis, by disturbance and feedforward setting: none (white bar), flow (hatched) and flow and analysis (solid). A dash marks a case the table does not report. Notice that any feedforward takes the flow step from 1227 to 7.3, that the composition analyzer helps only on the composition swings, and that no setting changes the hydrocarbon slug, which only the feedback trim answers.

🕮 **Pause.** Try Exercise 13.2 before reading on.

13.5.3 Feedforward Against Feedback

Table 13.6 reports the integrated absolute error in mol% min over the same window for three disturbances and three feedforward settings, from scenarios 5, 6 and 7 with the feedforward chosen before the run, and Figure 13.2 draws it on a logarithmic axis. Each window starts at the first event of its scenario, and the error is integrated at the model step. The application scorecard, which integrates its 0.05 min log, agrees except on the brief flow transient, where it reads 7.2.

About 168 times less integrated error on the flow step. With no feedforward the base air remains at 17.33 kmol/min and the trim multiplies it up to 19.06 kmol/min at its high stop of 1.10. Its ten percent authority cannot supply the twenty percent increase needed by the load. The true air ratio settles at 0.9165 and recovery at 89.8 %, with 9.71 t emitted over the six hours after the step. The flow ratio brings the air up within about a minute, but the analyzer reports that minute 3.2 min later, and the trim answers an error the feedforward has already removed by taking the ratio to 1.032. The stack, over the permit from the step to 21.1 min, goes back over it four more times before 38.4 min. The permit alarm raised at 20.02 min therefore latches, because the stack settles at 7.73 t/d, inside the permit but above the 7.65 t/d at which the alarm resets.

The composition swings separate the second and third settings. Over the 380 min from the first of six moves, four hours of an amine unit taking the acid gas between 79 and 88 mol% hydrogen sulfide and two of settling, the flow feedforward alone leaves an integrated error of 163 and 2.67 t of sulfur to the stack, and adding the composition analyzer gives 69 and 2.26 t. Four hundred kilograms, about one and a half tonnes a day at that rate, bought with one chromatograph.

The hydrocarbon slug is invisible to both, because nothing on the plant measures it. Six tenths of a percent more hydrocarbon raises the true air demand by four percent, and the trim finds every bit of it, walking the ratio to 1.040 and recovering 97.62 %. That is the honest division of labor: feedforward handles what is measured and feedback handles what is real.

> **TRY IT IN THE APP** **7 · The disturbance nobody measures**
>
> Load the case and press Run. At 20 min the hydrocarbon in the acid gas rises from 1.0 to 1.6 mol%, and the feedforward cannot see it. Watch the trend of the air ratio the trim asks for: it walks to 1.040, and the recovery comes back to 97.62 %. Feedback finds the four percent of extra air demand that nothing on the plant measures.

13.5.4 The Analyzer That Lies

A tail gas analyzer drifting 0.5 mol% high leads the trim to add air until the analyzer reads its setpoint, which puts the true demand at −0.500 mol%. The plant runs 1.4 % rich with the air demand indication exactly on target, recovery falls to 96.94 % and the stack goes to 9.60 t/d, over the 9.0 t/d permit. The air demand alarm stays quiet, because the measured error is near zero and even the true error of 0.5 mol% is inside its 1.2 mol% band. The one warning is the stack permit alarm, raised at 21.77 min from the stack figure, which the analyzer does not influence. It latches: the stack drops back under the permit twice while the trim settles, from 25.0 to 30.4 min and from 34.4 to 38.1 min, but the alarm clears only below 7.65 t/d and so stays up for the rest of the run, with the air demand reading sitting on its false target throughout. That failure mode is why air demand analyzers are the most carefully maintained instruments on a sulfur plant, and none of the maintenance is control engineering [18].

Pause. Try Exercise 13.3 before reading on.

> **TRY IT IN THE APP** **9 · The analyzer that lies**
>
> Load the case and press Run. At 20 min the tail gas analyzer starts to read 0.5 mol% high, and every loop stays in Auto. Watch the air demand bar under the plant: the open marker, the analyzer, stays on two to one while the filled marker, the true demand, settles near −0.5 mol%. The stack permit alarm comes up at 21.77 min and stays.

13.5.5 Three Beds, Three Windows

Measured at steady state on bed 3, recovery is 97.79 % with the inlet at 185 °C and 97.27 % at 230, with the dew point at 180. Colder is better, exactly as the equilibrium says, with a phase change as the floor. Measured on bed 1, recovery is 96.94 % at 210 °C of inlet and 98.24 % at 275, which the equilibrium flatly denies. Hydrolysis of carbonyl sulfide and carbon disulfide runs from 19 to 76 % across that range, and everything bed 1 gives up in conversion is recovered by beds 2 and 3, while nothing is downstream of bed 3 to recover its losses. Scenario 10 makes the same moves on a smaller scale, bed 1 to 210 and then 265 °C and bed 3 to 225 °C, and settles at 96.94, 98.07 and 97.74 %. A stage in the middle of a train has different economics from the stage at the end.

13.5.6 The Graded Challenge

The challenge runs fifteen hours with two flow changes, three composition swings and a hydrocarbon slug, starting with the feedforward off and the trim tuned at half its correct gain. It is met when the air demand is held within 0.25 mol% of setpoint through the last third, every bed stays above its dew point, and the shift emits less than five tonnes of sulfur. The run as loaded makes 199.5 t of sulfur and emits 13.29 t. Configured as suggested it makes 208 t and emits 4.5 t, and the stack permit alarm still comes up at every one of the six events and clears between them.

The verdict is judged at 900 min on the true air demand at every logged sample from 594 min, the last third of the run, with the model parameters and disturbances as given. Changing either voids

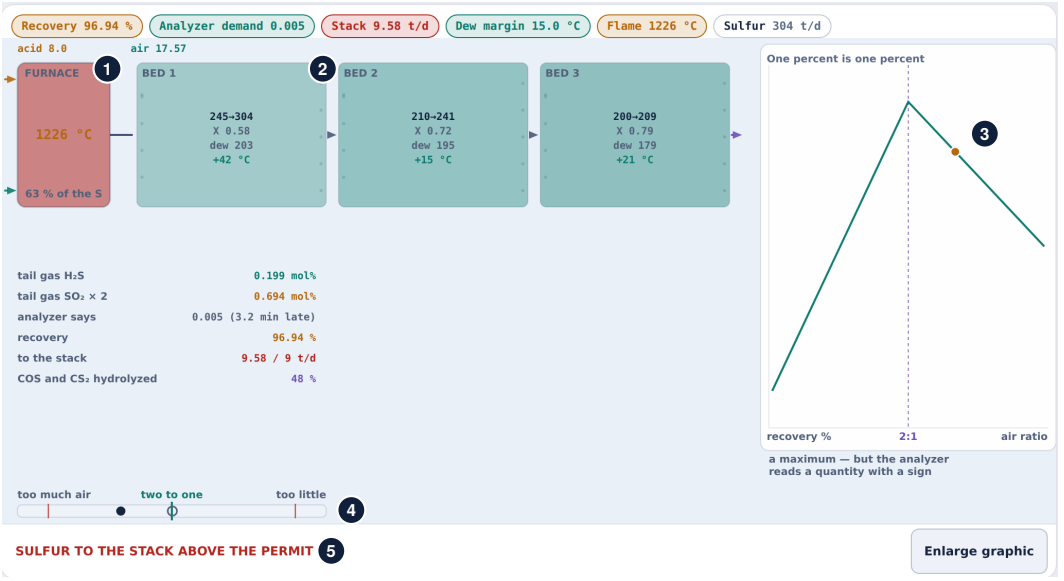


Figure 13.3 The Claus unit application running scenario 9, the analyzer bias case, at $t = 120$ min, 100 min after the +0.5 mol% drift, with every loop in automatic, cropped to the plant view with numbered callouts. (1) The furnace and (2) the three catalytic beds are drawn on one line, each bed carrying its inlet and outlet temperatures, its conversion, its sulfur dew point and its margin above it. (3) Beside them a sketch of recovery against air ratio marks the operating point at 1.014 and 96.94 %, on a curve recalculated from the current feed, bed temperatures and catalyst activities using the same model as the operating point. (4) The bar under the scene is the air demand, with two to one marked and the alarm limits either side. The true value of –0.495 mol% is the filled marker and the analyzer reading of 0.005 mol% the open one, on the optimum. (5) The latched stack permit alarm is on the status line below.

the attempt. As loaded, with no feedforward, $K_c = 0.010$, $\tau_I = 20$ min, bed inlets at 225, 205 and 205 °C and the override off, the worst error is 0.332 mol%, above the required band, and 13.29 t is emitted, although the dew point margin never falls under 11.1 °C. With the suggested settings, flow and analysis feedforward, $K_c = 0.022$, $\tau_I = 12$ min, inlets at 258, 208 and 194 °C and the override on, the worst error is 0.144 mol%, with 0.106 mol% to spare, the margin never falls under 11.6 °C, and 4.48 t is emitted. The same settings chosen before Run or carried through Reset give identical trajectories. Over seven other analyzer noise seeds the worst error ranges from 0.134 to 0.180 mol% and the emission from 4.45 to 4.49 t, all passing.

13.5.7 Measured Comparison of PI and Smith Trim

Scenario 12, the command-memory case, runs 300 min on flow feedforward. Hydrogen sulfide changes from 85 to 82 mol% at 20 min and returns at 150 min. Unmeasured hydrocarbon changes from 1.0 to 1.3 mol% at 90 min and returns at 210 min. The model contains none of these future events. Each paired experiment uses seed 20260922, identical instruments, noise, feedforward and event times. PI retains each scenario's ordinary tuning, including the unchanged suggested challenge settings. Smith uses $\lambda = 2$ and $\widehat{\theta} = 3.2$ throughout, except for explicitly labeled model-delay experiments. No tuning is adjusted by case.

Table 13.7 integrates true signed-demand error from $t = 0$ through each stated duration at the 0.05 min logging interval. These windows include startup, unlike the earlier event-relative tables. Travel sums absolute requested-ratio increments. Rail time uses the scorecard tolerance of 2×10^{-7}

Table 13.7 Fixed-Setting Trim Comparisons Including Startup

Case	Trim	Duration (min)	Peak $\lvert D \rvert$ (mol%)	IAE (mol% min)	Travel (ratio)	Rail (min)	Permit (min)	Crossings	Emitted (t)
Command-memory case	PI	300	1.209	52.239	0.281	0.00	60.65	10	1.778
	Smith	300	1.177	19.953	0.105	0.00	22.70	4	1.635
	Stronger PI	300	1.177	24.125	0.204	0.00	25.55	6	1.654
	Same-fast PI	300	3.849	541.683	7.333	32.65	262.70	44	3.942
	Smith, $\hat{\theta} = 1.6$	300	1.177	25.497	0.202	0.00	29.10	7	1.660
Composition step	PI	400	2.666	41.958	0.196	0.00	29.85	2	2.141
	Smith	400	2.666	14.549	0.077	0.00	7.45	1	2.063
Flow step, FF off	PI	400	6.710	1226.928	0.100	375.95	380.00	1	10.346
	Smith	400	6.708	1227.458	0.100	375.55	380.00	1	10.349
Composition swings	PI	420	3.581	162.604	1.023	0.00	171.70	17	2.882
	Smith	420	3.750	70.667	0.374	0.00	43.35	6	2.479
Hydrocarbon step	PI	400	1.366	21.578	0.097	0.00	37.90	2	2.202
	Smith	400	1.366	7.482	0.040	0.00	7.90	1	2.117
Stale-composition FF	PI	400	6.052	1198.411	0.100	375.75	380.00	1	5.391
	Smith	400	6.052	1198.979	0.100	375.30	380.00	1	5.393
Analyzer bias	PI	400	0.500	184.011	0.038	0.00	369.10	3	2.619
	Smith	400	0.500	188.985	0.014	0.00	377.55	1	2.634
Actual delay 6.4 min	PI	300	1.229	60.633	0.423	0.00	68.80	10	1.815
	Smith	300	1.174	65.042	0.367	0.00	76.05	12	1.833
	Smith, $\hat{\theta} = 1.6$	300	4.026	505.368	3.590	29.75	260.10	25	3.803
	Smith, $\hat{\theta} = 6.4$	300	1.177	32.232	0.105	0.00	35.55	4	1.690
Challenge, loaded	PI	900	5.915	1159.652	0.286	458.50	842.65	29	13.294
	Smith	900	5.890	1094.027	0.367	499.25	842.60	73	12.997
Challenge, suggested	PI	900	5.487	93.929	1.414	0.00	63.30	16	4.481
	Smith	900	5.440	66.411	0.815	0.00	37.20	11	4.312
Flow step, flow FF	PI	400	6.183	7.188	0.132	0.00	10.45	5	2.183
	Smith	400	6.183	3.881	0.049	0.00	3.75	2	2.165
Swings, analysis FF	PI	420	3.760	68.618	0.897	0.00	71.05	24	2.469
	Smith	420	3.732	49.312	0.327	0.00	52.75	11	2.385

at either trim bound. Permit time counts samples above the current stack limit, and crossings count upward transitions rather than hysteretic alarm resets. Stack rates and integrated mass are sulfur-equivalent quantities, not species-specific sulfur dioxide measurements. None of these comparisons reaches the independent air cap.

On this case, IAE is 52.239 for ordinary PI and 19.953 for Smith, with emitted sulfur 1.778 and 1.635 t. Permit time is 60.65 against 22.70 min, with 10 against 4 upward crossings. A fixed stronger PI ($K_c = 0.018$, $\tau_I = 4$ min) gives IAE 24.125, so part of the shipped-PI difference is tuning. Ordinary PI at the exact Smith tuning ($K_c = 0.014433861$, $\tau_I = 1.05$ min) gives 541.683 and 32.65 min on a trim rail. Compensation permits that faster integral setting in this experiment, without establishing global optimality.

At 32 min, ordinary PI has true demand -0.611391 and analyzer -0.310293 mol%. Smith has true demand -0.016843, analyzer -0.134111 and corrected feedback -0.029178 mol%, as shown in Figure 13.4. The correction reflects previously commanded air. On the isolated unknown-composition step, PI and Smith have exactly the same sampled peak of 2.665 970 mol%. Different whole-sequence peaks in Table 13.7 include different residuals from earlier events and do not demonstrate anticipation of unknown feed.

Smith loses when the feedforward leaves insufficient authority: IAE is 1227.458 against 1226.928 with feedforward off and 1198.979 against 1198.411 with stale composition. Analyzer bias gives 188.985 against 184.011. With actual transport 6.4 min, the nominal predictor gives 65.042 against PI's 60.633. An explicit matched model delay of 6.4 min lowers that IAE to 32.232, while a wrong 1.6 min model gives 505.368 and 29.75 min on a ratio rail. These losses remain numerically valid predictions, without fallback.

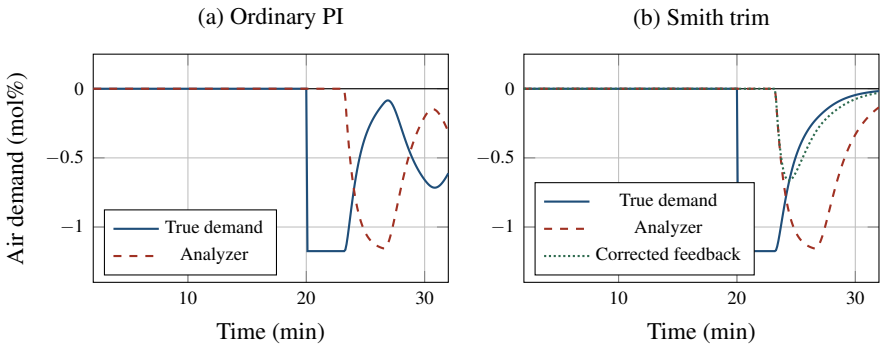


Figure 13.4 First unknown composition change under ordinary PI and Smith trim. Both panels show the same 30 min window ending at 32 min. True demand is the solid line, the analyzer the dashed line and, in panel (b), the corrected feedback the dotted line. Corrected feedback combines the analyzer with two incremental model branches and is distinct from true demand.

> ✦ DID YOU KNOW? **The Predictor Does Not Help Against a Lying Analyzer**
>
> From the simulator (Table 13.7). With the analyzer biased, the Smith trim integrates an error of 188.985 against 184.011 mol% min for ordinary PI, and spends 377.55 of the 400 min over the permit against 369.10. The predictor corrects for the air it sent, but the analyzer's bias stays in the residual it feeds back.

Across the 61 measured variants, active Smith runs execute 6000–18000 model/PI calls and reject zero samples. On the Node host of the measurements, the case's mean module time is 0.908 µs per call, including both model branches and PI. This timing depends on host load, JIT compilation and clock resolution, and does not measure plant integration or rendering.

13.6 Technology in Practice

Figure 13.5 draws the air demand trim, its feedforward and one of the bed inlet loops of Section 13.4 with the hardware a refinery installs.

What Measures It

The trim reads one instrument at the far end of the plant, and the feedforward reads the acid gas flow and, in the flow and analysis setting, a chromatograph on the acid gas. Chromatographs are relatively slow, and a 1989 patent records why a feed analysis corrected by the tail gas still leaves the plant off ratio: it does not see the hydrocarbons until the gas they leave reaches the tail gas analyzer, and the heavier hydrocarbons are usually the ones that upset the air [19]. Scenario 7 is that case.

> TECHNOLOGY **Sensor: The Tail Gas Analyzer**
>
> A chromatograph separates the species of a sample in time, so its result arrives a cycle of minutes after the sample and some dead time is inevitable [20]. A non-dispersive photometer measures a gas continuously by the infrared or ultraviolet light it absorbs, and each compound absorbs in its own pattern [20]. Either kind drifts: analytical instruments drift more than those for pressure, level, temperature or flow, so they are checked with span gases of known concentration, often by a self-calibration system on a schedule [20]. A drift looks calm on a trend. The chapter's analyzer has 3.2 min of transport delay and a 0.8 min average, and it drifts but never fails outright.

What Moves It

The trim moves the air, and the reheaters move the bed inlets. Each catalytic stage of a Claus plant is a reheater, a catalyst chamber and a condenser, and the beds must run above the sulfur dew point so

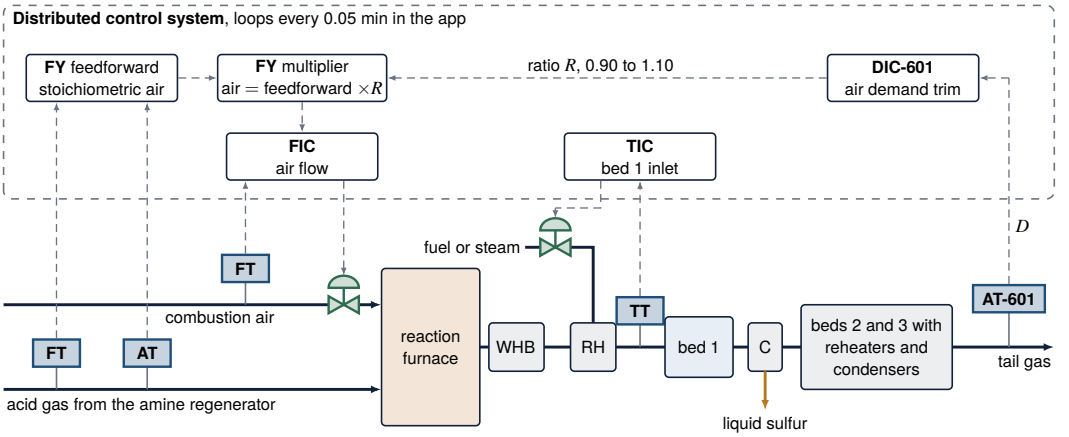


Figure 13.5 The air demand trim, its feedforward and the bed 1 inlet loop with their hardware. Lettered boxes are sensing devices: the acid gas flow meter (FT), the acid gas chromatograph used by the flow and analysis setting (AT), the bed inlet thermocouple (TT) and the tail gas analyzer (AT-601). Valve symbols are final elements, and dashed arrows are signals. Notice the length of the path from the air valve through the furnace and three condensers to the analyzer, which is the 3.2 min of the chapter, and that the trim multiplies the feedforward instead of adding to it. Beds 2 and 3, their loops, the dew point override and the Smith option are left out. Not to scale, and not a piping diagram.

that liquid sulfur does not deactivate the catalyst [2]. Plants reheat by mixing in hot gas bypassed from the waste heat boiler, by steam exchangers, by exchangers against hot gas from a converter, or by fired heaters that burn fuel or acid gas short of air, so that no oxygen reaches the catalyst [7]. The model sets each reheater outlet between 150 and 285 °C and leaves the method open.

> **TECHNOLOGY** **Final Element: Main and Trim Air Valves**
>
> The 1977 patent split the furnace air into a main line and a trim line, each with its own flow controller and control valve. The trim valve took the faster and finer adjustment over the short term, and a slow PI controller moved the main air so that the trim air returned to a chosen steady flow over the long term [11]. A pneumatic valve takes a set position when its air fails, chosen from the risks of the process, and friction in its stem makes it move in steps and deaf to small reversals [20]. The chapter has one air flow, which follows the computed demand through a 0.25 min lag between 0 and 40 kmol/min.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level oversees local controllers, loops are sometimes nested, and cycle times range from milliseconds to minutes [21]. The app executes every loop each 0.05 min, inside that range. The trim and the three reheater loops are PI blocks, the feedforward is a calculation block whose output the trim's ratio multiplies, and the dew point override raises a setpoint through a high select. The 1977 patent built the same calculation on an analog computer and noted that a digital computer could easily be programmed to do it [11].

> **IN PRACTICE** **The Stack Monitor Sits Outside the Loop**
>
> In the United States, a refinery sulfur plant whose construction or modification began after May 14, 2007, with a capacity above 20 long tons a day, burning ambient air and incinerating its tail gas, must keep its sulfur dioxide below 250 ppm by volume, dry and at zero excess air. It must monitor and record that concentration

continuously, with an oxygen monitor to correct for excess air, and each rolling 12-hour average above the limit is reported as excess emissions [22]. Emissions analyzers are often given daily calibration checks [20]. The stack monitor reads what leaves, not the air demand, so it is the check from outside the loop that the drifting analyzer of scenario 9 calls for.

What Is Modernizing It

Three developments touch this loop. The first is the inferred analyzer. A soft sensor trained on other plant measurements can act as a back-up for a failed hardware sensor, and it needs maintenance of its own because the data drift [18]. One such sensor estimated the hydrogen sulfide in the tail gas of a refinery sulfur unit, designed to replace the analyzer during its maintenance [23]. The second is dead-time compensation. The 1977 patent already compared a delayed analyzer with an analog model of the process run without its dead time [11], the idea of the Smith option of Section 13.4. The third is predictive control, with more than 4600 applications reported and refining the largest share [12], which can arbitrate the constraints a large unit shares with its neighbors.

13.7 Limitations

The thermal stage is a fixed conversion. The reaction furnace is the part of a Claus unit that a review describes as least understood [5], and this model does not attempt it.

There is no tail gas treating unit, which is what takes a modern plant from 97 to 99.9 percent overall [6, 8], so the application teaches the Claus section and not the plant.

Sulfur species are lumped and the dew point correlation is fitted, so the absolute dew point values carry less meaning than their changes.

Ammonia is represented only through its air demand and a furnace temperature alarm. Real ammonia chemistry in a Claus furnace is complicated and dominates the sour water debottlenecking literature.

The analyzer drifts but never plugs. A plugged sample system reports the last good value and looks calm, which is a different and harder fault than the bias implemented here.

The Smith model freezes local gain and both lags at the declared design point. Ratio normalization does not capture changes in the feedforward base, chemistry, catalyst condition or instrument dynamics. Unknown composition is learned through the delayed measured residual, and analyzer bias is preserved in that residual. The authority and bias rows in Table 13.7 show the resulting losses. Agreement of individual model recursions does not establish closed-loop robustness over the full adjustable plant range. The 6.4 min transport experiment illustrates the sensitivity to dead-time error studied in compensator design [16].

Finally, the constraint and economic arbitration a large unit needs [12] is absent, as is the plantwide question of which upstream unit owns the acid gas composition [13].

OPEN CHALLENGE **A Furnace Model Good Enough to Steer By**

The air a furnace needs depends on what burns in it, and no model predicts that reliably. A 2001 thesis found no comprehensive kinetic model of the reaction furnace, and found the equilibrium models in use failing worst when carbon dioxide and hydrocarbons are present [4]. A 2018 review describes detailed schemes of 130 species and more than 1500 reactions, too large and stiff for flow simulation, and a reduced scheme with mean absolute errors of 7.62 and 7.88 % against plant data [5]. The difficulty is a flame zone above 2000 °C lasting milliseconds, followed by up to two seconds without oxygen [5]. Progress would be a model fast and accurate enough to compute the air for a feed whose hydrocarbons change.

13.8 Discussion

Three results generalize past sulfur plants.

The first is the move that makes this optimum controllable. Recovery has a maximum and cannot be held by feedback. The difference between the hydrogen sulfide and twice the sulfur dioxide is zero

there and can. Where an objective has a maximum, the first question is whether some measurable quantity changes sign across it, and finding one converts a search into a loop.

The second is that on a dead-time-dominant loop the peak error is set by the disturbance, and how long it lasts by the delay, rather than by the tuning. Measured here, the peak is 2.67 mol% for every tuning from sluggish through correct to oscillating, and it holds for the 3.2 min of dead time. Effort on such a loop belongs in the measurement and the feedforward, and the useful question about a slow loop is where the delay is and whether it can be moved.

The third is the division of labor. Feedforward answered a flow disturbance about 168 times better than feedback, cut the error on a composition swing from 163 to 69 when it was given an analyzer, and answered a hydrocarbon slug not at all. Every feedforward is a model, every model is incomplete, and the feedback loop underneath stands in for the rest.

13.9 Conclusion

A Claus unit is an optimum that behaves like a setpoint, and the reason is a measurement rather than a controller. In the simulated unit of this chapter recovery is 97.619 % at an air ratio of one, 96.640 % at 0.990 and 96.640 % at 1.020, so one percent short of air costs exactly what two percent of surplus costs, and the stack figure moves from 7.46 t/d to 10.53 t/d against a 9.0 t/d permit for the same one percent. The air demand difference is zero at that maximum and carries a gain of −36 mol% per unit of ratio, which is what makes an ordinary PI loop sufficient. The analyzer is 3.2 min downstream, so the peak error after a composition step is 2.67 mol% for every tuning tried and lasts as long as the delay, and the remedy is feedforward rather than tuning: measured over the 380 min after a twenty percent flow step, the integrated air demand error is 7.3 with a flow ratio station and 1227 without one. A tail gas analyzer drifting half a mole percent takes the plant to 9.60 t/d of stack sulfur with the air demand reading on target, and only the latched stack permit alarm says so. A fuller model would add a tail gas treating unit, a reaction furnace with kinetics rather than a fixed conversion, ammonia chemistry, and a plugged sample system as a second analyzer fault. Even without them, the sulfur plant is a compact case for learning what feedforward is for.

The optional Smith trim reduces IAE from 52.239 to 19.953 on the fixed 300 min comparison, while the stronger ordinary PI reaches 24.125 and model-delay error, sensor bias and exhausted authority retain their measured limitations.

KEY POINTS

- The air demand reading is zero at the best air rate and has a sign on either side, so an ordinary PI loop holds an optimum like a setpoint.
- One percent short of air costs exactly what two percent of surplus costs: recovery is 96.640 % at 0.990 and at 1.020 against 97.619 % at 1.000.
- With the analyzer 3.2 min downstream, the first peak after a composition step is 2.67 mol% for every tuning, and the remedy is feedforward rather than tuning.
- An analyzer drifting half a mole percent takes the stack to 9.60 t/d with the reading on target, and only the stack permit alarm says so.

Where else this applies. Volume 3, Chapter 4 holds an oxygen trim on its air-to-fuel ratio with the same structure, and Volume 12, Chapter 6 shows a fixed-sign controller failing on an optimum that has no such measurement.

References

[1] U.S. Geological Survey, “Sulfur,” Mineral Commodity Summaries 2006, pp. 164–165, Reston, VA, 2006, usgs.gov.

[2] U.S. Environmental Protection Agency, “AP-42, fifth edition, volume I, chapter 8, section 8.13: Sulfur recovery,” Compilation of Air Pollutant Emissions Factors, Research Triangle Park, NC, April 2015, 2015, epa.gov.

[3] J. E. Normey-Rico and E. F. Camacho, “Dead-time compensators: A survey,” *Control Engineering Practice*, vol. 16, no. 4, pp. 407–428, 2008.

[4] A. E. Pollock, "Kinetic modeling of the modified Claus reaction furnace," Ph.D. dissertation, University of Calgary, Calgary, Alberta, Jun. 2001.

[5] S. Zarei, "A review of investigations on claus reaction furnace of sulfur recovery unit," *COJ Technical & Scientific Research*, vol. 1, no. 3, 2018.

[6] H. Ghahraloud, M. Farsi, and M. Rahimpour, "Modification of claus sulfur recovery unit by isothermal reactors to decrease sulfur contaminant emission: Process modeling and optimization," *Chemical Product and Process Modeling*, vol. 14, no. 2, 2018.

[7] M. Sassi and A. K. Gupta, "Sulfur recovery from acid gas using the claus process and high temperature air combustion (hitac) technology," *American Journal of Environmental Sciences*, vol. 4, no. 5, pp. 502–511, 2008.

[8] A. Taghizadeh Damanabi and F. Bahadori, "A new approach for hydrogen production in claus sulfur recovery process," *Journal of Sulfur Chemistry*, vol. 40, no. 2, pp. 137–148, 2018.

[9] S. Delfi, M. Mosaferi, A. Khalafi, and K. Zoroufchi Benis, "Sulfur dioxide emissions in iran and environmental impacts of sulfur recovery plant in tabriz oil refinery," *Environmental Health Engineering and Management*, vol. 5, no. 3, pp. 159–166, 2018.

[10] F. W. Karasek, "Sulfur production," U.S. Patent 3 026 184, Mar. 20, 1962.

[11] D. E. Smith, "Control system for a sulfur process," U.S. Patent 4 100 266, Jul. 11, 1978.

[12] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[13] M. L. Luyben, B. D. Tyreus, and W. L. Luyben, "Plantwide control design procedure," *AIChE Journal*, vol. 43, no. 12, pp. 3161–3174, 1997.

[14] K. J. Åström and T. Hägglund, *Advanced PID Control*. Research Triangle Park, NC: ISA – The Instrumentation, Systems, and Automation Society, 2006.

[15] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *Journal of Process Control*, vol. 13, no. 4, pp. 291–309, 2003.

[16] J. E. Normey-Rico and E. F. Camacho, "Unified approach for robust dead-time compensator design," *Journal of Process Control*, vol. 19, no. 1, pp. 38–47, 2009.

[17] D. E. Rivera, M. Morari, and S. Skogestad, "Internal model control: PID controller design," *Industrial & Engineering Chemistry Process Design and Development*, vol. 25, no. 1, pp. 252–265, 1986.

[18] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[19] R. L. Reed and D. A. Holdeman, "Method for controlling Claus furnace with variable hydrocarbon feed composition," U.S. Patent 4 836 999, Jun. 6, 1989.

[20] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[21] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[22] U.S. Environmental Protection Agency, "Standards of performance for petroleum refineries for which construction, reconstruction, or modification commenced after may 14, 2007," Code of Federal Regulations, Title 40, Part 60, Subpart Ja, sections 60.102a and 60.106a, edition of July 1, 2025, 2025, govinfo.gov.

[23] A. di Bella, L. Fortuna, S. Graziani, G. Napoli, and M. G. Xibilia, "Soft sensor design for a sulfur recovery unit using genetic algorithms," in *2007 IEEE International Symposium on Intelligent Signal Processing*, 2007, pp. 1–6.

13.10 Exercises

13.1 Concept check · 2 min

Recovery on the Claus unit has a maximum at an air ratio of 1.000, yet the trim DIC-601 is an ordinary PI loop with a fixed direction. Why can it hold the plant at the maximum?

(a) The recovery curve is flat at its top, so a small air error costs too little for the loop to need a direction

(b) The analyzer reads the air demand, which is zero at the maximum and changes sign across it, so the error has a direction

(c) The trim's limits of 0.90 and 1.10 keep the ratio close enough to the maximum that the sign of the error never matters

(d) The stoichiometric feedforward computes the optimum air exactly, so the PI loop has only analyzer noise left to remove

13.2 Concept check · 2 min

In scenario 3 the acid gas steps from 85 to 78 mol% hydrogen sulfide at 20 min, on flow feedforward. Trim gains of 0.010, 0.022 and 0.040 all leave the same peak air demand error, 2.67 mol%. Why?

(a) The trim sits on its 0.90 stop through the excursion, so none of the three gains is acting

(b) The integral time is too long in all three tunings for the integral to act before the peak

(c) The error arrives with the step and the analyzer reports nothing for about 3.2 min, so no tuning has anything to act on

(d) The flow feedforward cancels whatever part of the error depends on the tuning of the trim

13.3 Concept check · 2 min

In scenario 9 the tail gas analyzer drifts 0.5 mol% high at 20 min, with every loop in Auto and the feedforward on flow and analysis. Which indication on the panel tells the operator that something is wrong?

(a) The air demand alarm, because the true air demand settles 0.5 mol% away from its setpoint

(b) The trim ratio, which runs to its 1.10 stop as the trim adds air to correct the reading

(c) The Analyzer air demand tile, which settles 0.5 mol% away from the setpoint

(d) The stack permit alarm, which is raised from the stack figure that the analyzer does not influence

13.4 Theory and calculation · 7 min

Use the stoichiometric loss of Equation (13.1) and the measured recoveries of Table 13.4.

(a) Compute the loss at air ratios of 0.98, 0.99, 1.01 and 1.02. How many times more does a unit of air shortfall cost than a unit of surplus? *Checkpoint:* 0.02, 0.01, 0.005 and 0.01; twice as much.

(b) From the rows at 1.000 and 0.990, find the drop in recovery per unit of loss. Use it to predict the recoveries at 0.980 and 1.010 and compare them with the table. *Checkpoint:* 97.9 points per unit of loss; 95.661 and 97.130 %, as measured.

(c) The stack carries 7.46 t/d at 97.619 % recovery. Find the sulfur fed, the recovery the 9.0 t/d permit allows at least, and the band of air ratio that keeps the stack inside the permit. *Checkpoint:* About 313 t/d of sulfur, a floor of 97.13 %, and a band from about 0.995 to 1.010.

13.5 Theory and calculation · 7 min

Use the air demand of Equation (13.2). The design feed is 8.0 kmol/min of acid gas at 85 mol% hydrogen sulfide, 1.00 mol% hydrocarbon and no ammonia.

(a) Compute the stoichiometric air at the design point. *Checkpoint:* 17.33 kmol/min.

(b) The hydrocarbon rises to 1.60 mol%, as in scenario 7. The feedforward on flow and analysis still assumes 1.00 mol%. By how much does the demand rise, and what ratio must the trim find? *Checkpoint:* 18.02 kmol/min, 3.96 % more; a ratio of 1.040.

(c) In scenario 5 the flow rises to 9.6 kmol/min with the feedforward off, so the trim multiplies the held base of 17.33 kmol/min and stops at 1.10. Find the true air ratio and the ratio the trim

would need. A shortfall costs 0.979 points of recovery per percent of air, and at the new flow the unit recovers 97.94 % at a ratio of one. Predict the recovery. *Checkpoint:* 0.9165 against 1.200 needed; about 89.76 %.

13.6 Theory and calculation · 7 min

Take the trim's process gain as $K_S = -36.37$ mol% of air demand per unit of ratio.

(a) Compute the loop gain $|K_S|K_c$ for the three tunings of Table 13.5. Which exceeds one, and what does the table show for it? *Checkpoint:* 0.36, 0.80 and 1.45; the gain of 0.040 cycles.

(b) The challenge holds the air demand within 0.25 mol% of setpoint. What band of air ratio is that, and how many points of recovery can sit at its lean and its rich edge? Use 97.9 points per unit of loss. *Checkpoint:* ± 0.0069 of ratio; 0.67 points lean and 0.34 rich.

(c) The Smith trim sets $K_c = 1.05/(|K_S|\lambda)$ and $\tau_I = 1.05$ min. Compute K_c for $\lambda = 2$ and 4 min, and the loop gain at $\lambda = 2$. Why can the Smith PI run this short integral time when ordinary PI at the same setting rails? *Checkpoint:* 0.01443 and 0.00722; loop gain 0.525.

13.7 App challenge · 8 min

Load scenario 1, "1 · Two to one (MANUAL)", as shipped (default seed) and run 150 min at the default speed of 5 min per second, about 30 s. The air ratio steps to 0.990 at 20 min, to 1.020 at 60 min and back to 1.000 at 100 min.

(a) Pause once between 30 and 59 min, once between 75 and 99 min and once after 130 min. Each time read the Recovery, To the stack and Analyzer air demand tiles. *Checkpoint:* 96.64 %, 10.53 t/d, 0.366 mol%; then 96.64, 10.53, −0.712; then 97.62, 7.46, 0.000.

(b) From the two off-design readings, estimate the gain of air demand on air ratio. Is it the same on both sides? *Checkpoint:* About −36 mol% per unit of ratio: −36.6 lean, −35.6 rich.

(c) Watch the alarm banner. After which steps does the stack permit alarm come up, and when does it clear for good? *Checkpoint:* Just after 20 and 60 min; it clears for good about 0.7 min after the return to 1.000.

13.8 App challenge · 8 min

Load scenario 3, "3 · Three minutes of nothing", as shipped (Feedforward on Flow only, Air demand trim gain 0.022, Air demand integral time 12.0 min, default seed). Run 60 min at 5 min per second and export the CSV.

(a) Find the largest |`demand_pct`|, the true air demand, after the step at 20 min, and the logged time at which `demand_meas`, the analyzer, first differs from zero by more than 0.01 mol%. Read the scorecard IAE at 60 min. *Checkpoint:* 2.666 mol% at 20.05 min; the analyzer at 23.30 min; IAE 34.43 (± 0.1).

(b) Pause, press Reset, set the gain to 0.040 and the integral time to 8.0 min, and run 60 min again. Repeat with 0.010 and 20.0 min. Record the peak and the IAE at 60 min each time. *Checkpoint:* Peak 2.666 mol% both times; IAE 48.01 and 66.45 (± 0.1).

(c) Table 13.5 ranks the three tunings 42, 153 and 335 over 380 min. Why does the gain of 0.040 rank second, not last, at 60 min? *Checkpoint:* Its cycle has run only half an hour; the sluggish loop's error is large early.

13.9 App challenge · 9 min

Load scenario 5, "5 · Feedforward on the flow", as shipped (Feedforward None, default seed). The acid gas rises from 8.0 to 9.6 kmol/min at 20 min. Run 400 min at 20 min per second, about 20 s, and export the CSV.

(a) When does `ratio_cmd` reach the 1.10 stop? At 400 min, read the true ratio on the status line of the Air demand card, the Recovery tile and the scorecard IAE. *Checkpoint:* About 24.1 min; 0.9165, 89.76 %, IAE 1226.9 (±1).

(b) Pause, press Reset, switch Feedforward to Flow only and run 400 min again. Read the scorecard IAE, the largest `ratio_cmd` and its time, the To the stack tile and whether the stack permit alarm is still on. *Checkpoint:* IAE 7.19 (±0.1); 1.032 at 23.70 min; 7.73 t/d; the alarm is still on.

(c) Why does the trim move the ratio to 1.032 at all, when the flow feedforward has already answered the step? *Checkpoint:* It acts on the transient the analyzer reports 3.2 min late.

13.10 Application area · 3 min

A plant engineer proposes moving the tail gas sample point closer to the furnace to shorten the analyzer delay, at the cost of a hot, wet, corrosive sample line.

(a) What would a shorter delay do to the first peak after a composition step, and what to the error one minute after it? *Checkpoint:* Nothing to the peak; it cuts the error at one minute.

(b) What does the chapter suggest as the better investment on a plant whose disturbances are mostly composition, and what fault of such a sample line does the model leave out? *Checkpoint:* Composition feedforward; a plugged sample line.

13.11 Application area · 4 min

Measured at steady state, bed 3 gives 97.79 % recovery at 185 °C of inlet and 97.27 % at 230, with its dew point at 180. Bed 1 gives 96.94 % at 210 °C and 98.24 % at 275.

(a) Why does bed 3 want to run cold and bed 1 hot? *Checkpoint:* Equilibrium favors cold; bed 1 alone hydrolyzes COS and CS_2.

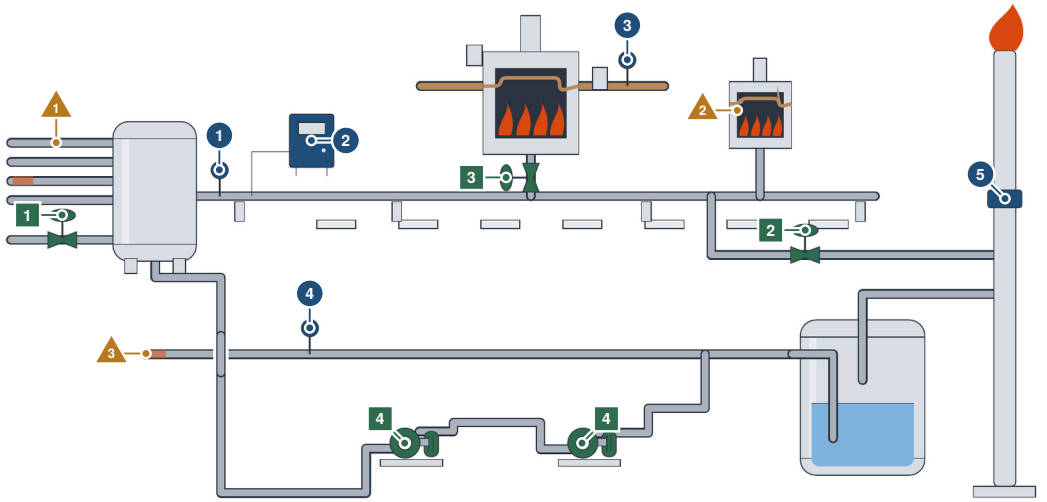
(b) Why is the dew point protected by a one-directional override rather than an alarm, and in which direction do the challenge's suggested inlets of 258, 208 and 194 °C move from the loaded 225, 205 and 205 °C? *Checkpoint:* The damage is delayed and the gain immediate; bed 1 hotter, bed 3 colder.

13.12 Review the agent · 5 min

An AI assistant recommends: "Switch the Air demand trim to Smith predictor. It takes the 3.2 min dead time out of the loop, so the trim sees a composition change at once. The first peak after an unknown feed change falls, a drifting analyzer is corrected as well, and the comparison table shows the IAE falling from 52.239 to 19.953." Find the flaw and say how you would check it in the app.

CHAPTER 14

A Refinery Fuel Gas Header That Should Never Import and Flare at Once



The controller moves 1 natural gas makeup valve 2 flare valve on the same controller output 3 fuel valve of heater H-1 4 two recovery compressor stages

The controller reads 1 fuel gas header pressure 2 Wobbe meter on the mixed gas 3 outlet temperature of H-1 4 flare header pressure 5 flow meter at the flare, for the record

What disturbs it 1 the four source gases: a coker drum switch, an FCC trip, a purge rise 2 heater H-20 lights 3 relief load into the flare header

One controller output opens either the makeup valve or the flare valve, and the recovery stages take gas from the flare header ahead of the water seal and return it to the mix drum. Illustration, not to scale.

IN PLAIN TERMS Buying Gas and Burning It at Once

The main idea. A refinery burns the gas its own units make and buys natural gas to make up the rest. One controller holds the pressure in the shared pipe. It opens a valve that buys gas or a valve that sends spare gas to the flare, an open flame that burns it.

The challenge. If the two valves overlap, the controller can keep both open. The plant then buys gas and burns gas at the same time, while the pressure looks normal. The pipe holds less than a minute of gas, so any upset moves the pressure fast.

Why it matters. A refinery should never pay for gas while it burns its own as waste.

THREE WAYS TO READ THIS CHAPTER

Explore	11 min	This page, the Did you know? and History boxes in Section 14.2, the Try it boxes and the second Did you know? box in Section 14.5, and Key points in Section 14.8. No equations.
Study	45 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 14.5 with Figure 14.3, Section 14.6 and Section 14.7 with its Open challenge box.

The control question A refinery fuel gas header holds well under a minute of its throughput, and one pressure controller splits its output between a natural gas makeup valve and a flare valve that pass money in opposite directions. How should that split range, the flare gas recovery and the heaters' firing be set so that the refinery never buys gas while it flares its own?

Learning objectives

- Calculate the integrating gain of a gas header from its volume and temperature, and the time an unbalanced flow takes to move its pressure.
- Explain why an overlap in a split range imports and flares while the pressure stays on setpoint, and why a dead band costs almost nothing on an integrating header.
- Compute the heating value and Wobbe index of a fuel gas mixture, and predict how much heat a burner at fixed pressure loses when hydrogen is added.
- Set the load and unload thresholds of a staged recovery compressor from where the loaded stages hold the flare header.
- Compare feedback, flow feedforward and MPC on a measured consumer surge, and say why an MPC on two valves still needs a positioner interlock.

Prerequisites Sections 1.4.2 (heaters, fuel heating value and the Wobbe index), 1.4.7 (split ranges and constraint control), 1.6.2 (velocity form) and 1.6.4 (scenarios). Section 1.6.3 is needed only for the MPC. Chapter 11 treats the drum switch that this header's coker gas follows.

The app https://apmonitor.com/apps/fuel-gas-header/

See also Volume 4, Chapter 5 for split ranges on heated and cooled barrel zones and Volume 4, Chapter 9 for valve-position control on a site with an import limit. Volume 3, Chapter 4 for the air side of a heater burning fuel gas and Volume 3, Chapter 6 for feedforward from a measured load. Volume 6, Chapter 7 for the choice of which flow of an inventory loop is held fixed.

14.1 Introduction

A refinery burns its own waste. The catalytic cracker's gas plant, the coker, a hydrotreater's recycle purge and the reformer's surplus hydrogen leave the units as gases nobody can sell as liquid, and a mix drum and header deliver them to every fired heater on the site. In the generic system simulated here those four sources send 18 500 Nm^3/h (normal cubic meters at 0 °C and 1.013 25 bar absolute), a recovery compressor returns 900 Nm^3/h of relief gas, and 1009 Nm^3/h of natural gas makes up the 250 MW the heaters burn. The 60 m^3 header at 3.5 bar g holds 10.4 kmol, which is 41 s of that throughput. When the coker switches drums and sends 3000 Nm^3/h less, the header falls 0.97 bar a minute and passes 3.0 bar g 45 s after the switch began.

The simulator is built around two beliefs it is meant to correct. The first is that a split range is a wiring detail. One pressure controller output opens the makeup valve below the middle of its range and the flare valve above it, and the two valves pass money in opposite directions. Configured with the two valve ranges overlapping by 4 % of output, the controller holds the header on setpoint while buying natural gas and flaring fuel gas at the same time, for 14.4 min of a 30 min case. The pressure trend shows nothing wrong. The second belief is that the heat a burner releases follows the heating value of its gas. At a fixed burner pressure it follows the Wobbe index, the higher heating value divided by the square root of the specific gravity, and adding 10 % hydrogen to the design gas lowers the heating value 7.1 % but the Wobbe index only 2.7 %. The class of problem is wider than refining. A single output split across two actuators with opposite costs appears in heating and cooling of one zone (the barrel zones of Volume 4, Chapter 5), in acid and base dosing, and in pressure held between a vent and a supply, and split-range control is a common structure in the process industry whose systematic design is still being worked out [1–3]. A header whose quality changes the response of every consumer at once has the same shape wherever gases of different origin are blended into one supply, and gas quality criteria for such interchange were argued for engines in 1986 [4].

However, the literature a student meets keeps the pieces apart. Fuel gas systems are scheduled by mixed-integer optimization over hours [5–7], flare gas recovery is evaluated by design studies and economics [8, 9], and gas interchangeability is a combustion subject [10]. This chapter puts the split range, the recovery compressor and the heaters' response to gas quality into one simulated refinery fuel gas header running on one clock, and measures what an overlap, a dead band, a staging threshold, a hydrogen-rich purge and three control architectures do to it. The ideal-gas balance, the split range, the proportional-integral (PI) loop, the Wobbe index and the model predictive controller (MPC) are all standard. The findings are measurements on one generic system at one seed. With the PI loop every dead band of zero or more gave no import and flare in the case built to show it, and only a dead band of exactly zero left 4.5 s in the 8 h challenge, so the overlap rather than the narrowness of the band is the fault. An MPC whose plans never open both valves still imports and flares for 7.9 to 9.3 min in 8 h unless an interlock waits for the positioners. The model has 23 states, its pressure rate matches the integrating gain computed from the declared volume to a relative 10^{-9}, and it is given here down to its last default so that it can be rebuilt. Its controllers see only instruments and never the simulator's state. Nine cases, the last a challenge, carry the lessons, and their completion rules look at the learner's own moves and at the state the header reaches after them.

Section 14.2 follows the domain, from flaring and flare gas recovery to the Wobbe index. Section 14.3 sets out the header, the gas properties, the split range, the flare header and the heaters as modeled. Section 14.4 describes the pressure loop, the flow feedforward, the MPC, the compressor staging and the heater loop with its Wobbe feedforward. Section 14.5 measures the nine cases and the challenge, Section 14.6 sets the hardware of a real header beside the model, Section 14.7 states the limitations and Section 14.8 draws the conclusions. The exercises at the end of the chapter run the cases.

14.2 Fuel Gas, Flares and the Wobbe Index

Fuel gas is a byproduct of refining and the primary energy source of a refinery, which burns what its units make and purchases what they do not [7]. The flare is the other end of that balance. Flaring is widely used by the fossil fuel industry to dispose of gas it cannot use [11], and an industrial survey of the practice covers the principles of combustion, the types of flare and case studies of their operation [12]. Recovering that gas instead became a design subject in its own right. A 2012 study examined reforming the flare gas of a gas refinery into usable energy with no routine flaring [13], the technical and economic case for recovering flare gas in several gas-processing plants was evaluated in 2017 [8], and a flare gas recovery system built on liquid ring compressors was designed and simulated in 2020 [9]. What a flare emits when it does burn has also been measured rather than assumed. Airborne sampling across three basins that account for more than 80 % of United States flaring found flares destroying 91.1 % of their methane, against the 98 % generally assumed, once unlit flares are counted [11].

> **✦ DID YOU KNOW? Refineries Burn Their Own Gas**
>
> In 2024, United States refineries burned 213,576 thousand barrels of still gas, counted at 6 million British thermal units per fuel oil equivalent barrel, and 1,031,147 million cubic feet of natural gas [14]. Still gas is the mixture of methane, ethane and other light gases that distillation, cracking and reforming produce.

The plant-wide view of fuel gas came from optimization. A mixed-integer linear model for multi-period scheduling of a refinery's fuel gas system was published in 2008 [5], the energy configuration and operation of refinery fuel gas networks were optimized together in 2015 [6], and a 2018 study made such a schedule robust to the fluctuation of fuel gas generation and consumption that leaves a theoretical plan infeasible on the day [7]. Those fluctuations are what the regulatory layer absorbs, and the header pressure loop, its split range between makeup and flare and its relation to the recovery compressor sit below the schedule. Split-range control is used to extend one controlled variable's

operating range across more than one manipulated variable, and a systematic design procedure that weighs the different dynamic effect and cost of each input was proposed in 2019 [1]. A generalized form with one controller per input and a baton passed between them, and model predictive control as the alternative that needs a detailed model, followed in 2020 [2].

Gas quality has its own history, from a 1977 utility study of hydrogen in burners to the Wobbe index as a criterion in 1986, and the History box below tells it. Interchangeability indices for premixed burners are still measured experimentally [10]. A refinery header sees the same problem every time a hydrotreater purges harder, because hydrogen-rich purge gas reaches every heater at once.

> **HISTORY 1977 and 1986: How Much Hydrogen a Burner Takes**
>
> A 1977 utility study asked how much hydrogen could be blended into natural gas and burned in typical appliances with no adjustment. It burned blends of 8 to 31 % hydrogen in eleven pilot burners and twenty-three main burners [15]. A burner passed only if its heat input changed little and its flames neither lifted, flashed back, yellowed nor burned incompletely. The heat input index of the study was a ratio of Wobbe numbers, the heating value divided by the square root of the specific gravity. Main burners in the open failed above 20 to 22 % hydrogen, limited by flashback when turned down [15]. In 1986 the Wobbe index was argued as the gas quality criterion for interchanging fuels in engines, because a change of composition at the same index left the air-to-fuel ratio and the combustion velocity nearly unchanged [4].

14.3 The Fuel Gas System as Modeled

No refinery stands behind the simulated system, and every size, flow, composition, price and threshold is a teaching value. Four sources and a natural gas makeup valve feed a mix drum and header of volume V, which supply one fully simulated heater, H-1, a bank of eighteen heaters lumped behind one demand, and a twentieth heater, H-20, out for maintenance until case 8 lights it. A flare valve on the header sends surplus gas to the flare stack downstream of the water seal. A separate flare header gathers the relief load of the units behind a 60 mbar g water seal, and a recovery compressor with two on/off stages returns that gas to the mix drum before it can lift the seal. The layout, every symbol of the equations below and the values of the challenge at 2 h are shown in Figure 14.1.

The model integrates 23 states. Eleven are physical: the kilomoles n_i of hydrogen, methane, an ethane-and-heavier lump C_{2+} and nitrogen in the header, the kilomoles n_F in the flare header, the makeup and flare valve stems x_{MU} and x_{FL}, H-1's burner pressure p_b, its firebox heat Q_f and outlet T_{out}, and the lumped heaters' firing demand h_L. Twelve are ledgers that integrate what crosses the boundaries (each lump's and the flare header's net inflow, the gas into and out of the header, the volume flared and imported, the carbon dioxide from the flared carbon and the energy flared and imported), so that conservation and every scorecard total can be checked exactly. Time is in minutes, flows in Nm^3/h, header pressure in bar gauge, flare header pressure in mbar gauge, heating values in MJ/Nm^3 on the higher heating value, and heat in MW. Table 14.1 lists the variables by role.

14.3.1 The Header and Its Integrating Gain

The header is one well-mixed ideal-gas volume at a fixed temperature, and each lump balances what the sources, the recovered gas and the makeup bring in against what the heaters and the flare valve take out at the header's own composition, as shown in Equation (14.1).

$$\begin{aligned} c\,\frac{\mathrm{d}n_i}{\mathrm{d}t} &= \sum_k F_k y_{k,i} + F_{RC}\,y_{rel,i} + F_{MU}\,y_{NG,i} \\ &\quad - \left(F_h + F_L + F_{su} + F_{FV}\right)\frac{n_i}{n} \\ p &= \frac{nRT}{V} - p_a, \qquad n = \textstyle\sum_i n_i \end{aligned} \tag{14.1}$$

where $c = 22.414 \times 60$ converts Nm^3/h to kmol/min, k runs over the FCC dry gas, the coker gas, the hydrotreater purge and the reformer net gas with compositions y_k, F_{RC} is the gas the recovery

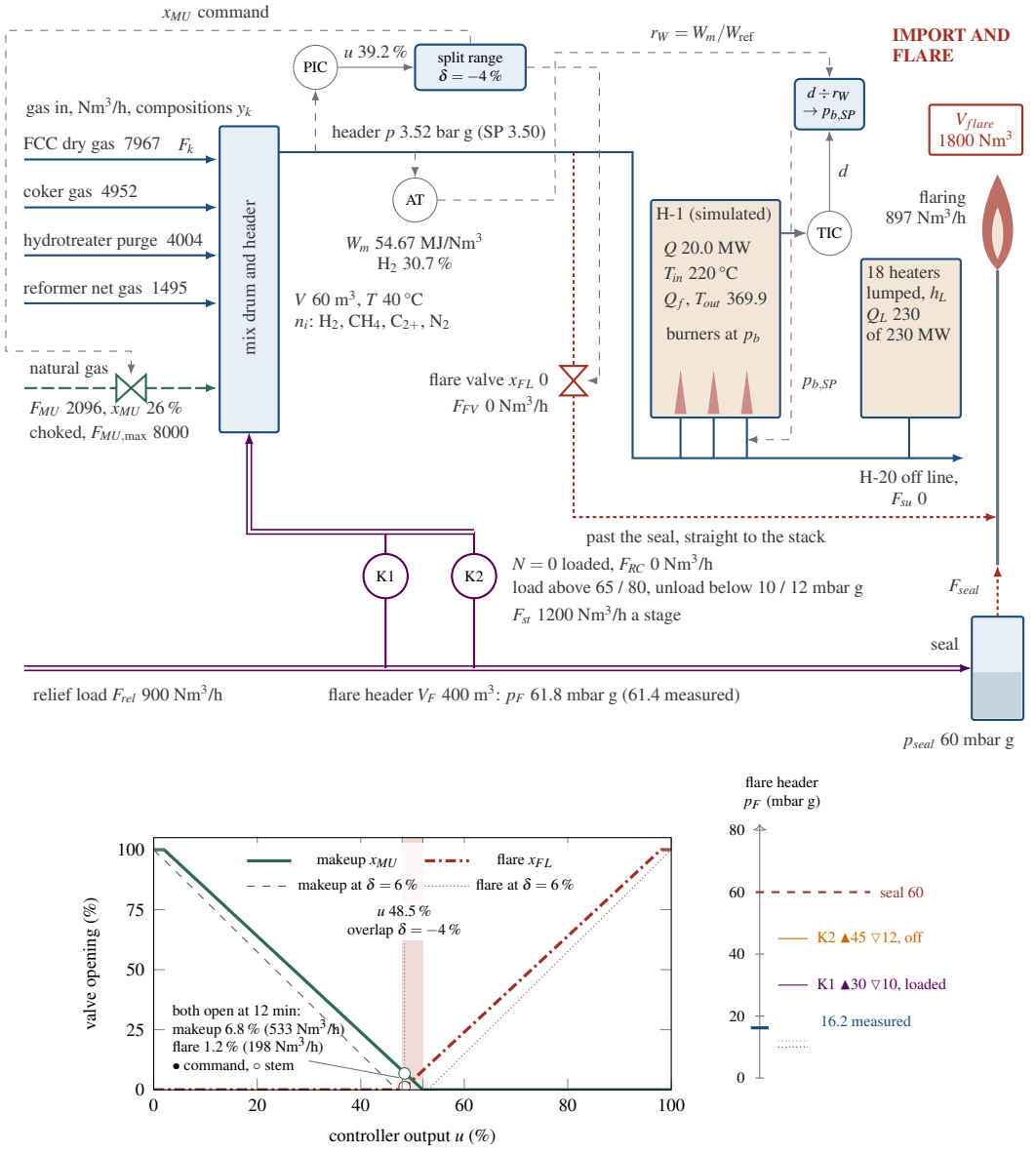


Figure 14.1 The fuel gas system as modeled. (a) Drawn after the System view of the challenge as shipped at $t = 2$ h: the four sources and the natural gas makeup into the mix drum, the PIC and its split range with a 4 % overlap, the Wobbe meter (AT) and the feedforward to the burner pressure of H-1 and the lumped bank, the header's flare valve discharging past the seal, and the flare header whose stages load above the seal, so the whole 900 Nm³/h relief load lifts the seal and burns (897 Nm³/h metered at the tip). Flows in Nm³/h at the meters unless marked. Fuel gas is drawn solid, natural gas long-dashed, gas to the flare dotted, relief gas as a double line and signals thin dashed. Process model only, not to scale. (b) The Split range view of case 2 at 12 min: one output, two valves, the overlap shaded, both valves open at $u = 48.5$ %, the makeup valve solid and the flare valve dash-dotted, with the suggested 6 % dead band thin (makeup dashed, flare dotted) for comparison, and the flare header's ladder with the seal and each stage's load (▲) and unload (▽) thresholds.

Table 14.1 Variables by Role at the Design Point

Symbol	Role	Design	Range
p	CV, header pressure (bar g)	3.50	SP 2.5–4.5
T_{out}	CV, H-1 outlet (°C)	370.0	SP 340–390
u	MV, split-range output (%)	41.07	0–100
x_{MU}, x_{FL}	MV, makeup and flare valves (%)	12.6, 0	0–100
d	MV, H-1 firing demand (%)	100	0–130
N	MV, recovery stages loaded	1	0–2
F_k	DV, sources (% of design)	100	0–200
F_{rel}	DV, relief load (Nm^3/h)	900	0–4000
d_L	DV, other heaters' firing (%)	100	70–120
s	DV, H-20 lit (%)	0	0–100
n	state, gas in the header (kmol)	10.40	—
y_{H_2}	output, hydrogen (mol %)	32.09	—
HHV, SG	output, heating value, specific gravity	44.097, 0.6501	—
W	output, Wobbe index (MJ/Nm^3)	54.69	alarm 52.5
p_F	output, flare header (mbar g)	16.25	seal 60
F_{MU}	output, makeup (Nm^3/h)	1009	0–8000
F_{flare}	output, flared (Nm^3/h)	0	—
Q, Q_L	output, heat released (MW)	20, 230	—
$t_{I\&F}$	output, import and flare (min)	0	—

compressor returns at the relief composition y_{rel}, F_{MU} the natural gas makeup at y_{NG}, F_h, F_L and F_{su} the fuel of H-1, the lumped bank and H-20, F_{FV} the flow through the header's flare valve, $T = 40\,°C$ and $p_a = 1.013\,25$ bar. Summing the four balances gives the integrator that every controller of this header fights, as shown in Equation (14.2).

$$\frac{\mathrm{d}p}{\mathrm{d}t} = K\,(F_{in} - F_{out}), \qquad K = \frac{RT}{V}\cdot\frac{1}{c}\cdot 10^{-2} \tag{14.2}$$

with R in J/(mol K), V in m^3, and the factor 10^{-2} turning kPa into bar per kmol. At $V = 60\,m^3$, $K = 3.2267 \times 10^{-4}$ bar/min per Nm^3/h, so 1000 Nm^3/h of imbalance moves the header 0.323 bar a minute, and at 240 m^3 it is 8.067×10^{-5}. The application's own test steps the FCC gas down by 800 Nm^3/h at both volumes and finds the engine's pressure rate equal to K times the step to a relative 10^{-9}. Nothing on the right side of Equation (14.2) answers the pressure. The sources send what their units make, the makeup valve is choked from a 20 bar absolute supply, and each heater's burner-pressure loop holds its own flow until its fuel valve is wide open. The header has no self-regulation until a heater starves, which for H-1 begins when the header falls to 2.216 bar g.

Pause. Try Exercise 14.1 before reading on.

14.3.2 Composition, Heating Value and the Wobbe Index

The gas properties follow from the four lumps by mole-fraction averages, as shown in Equation (14.3).

$$\begin{aligned} \mathrm{HHV} &= \textstyle\sum_i y_i H_i, \qquad SG = \sum_i y_i M_i / 28.96 \\ W &= \mathrm{HHV}/\sqrt{SG} \end{aligned} \tag{14.3}$$

where H_i is 12.75, 39.8, 92.0 and 0 MJ/Nm^3 and M_i is 2.016, 16.043, 38.5 and 28.013 kg/kmol for the four lumps, and the C_{2+} lump carries 2.6 carbon atoms a molecule. Table 14.2 gives each stream. The design gas, mixed from the sources, the recovered relief gas and the makeup, is 32.09 % hydrogen with $\mathrm{HHV} = 44.097\,MJ/Nm^3$, $SG = 0.6501$ and $W = 54.690\,MJ/Nm^3$. Pure hydrogen and methane come out at W of 48.32 and 53.47 MJ/Nm^3. The purge, at 42.48 MJ/Nm^3, lies below both, because a little heavy gas raises a light gas's specific gravity sharply.

Hydrogen lowers the heating value and the specific gravity together, so the Wobbe index moves much less than the heating value. Adding a fraction a of pure hydrogen to the design gas gives

Table 14.2 Stream Compositions (Mole Fractions of the Four Lumps) and Properties. Teaching Values

Stream	H_2	CH_4	C_{2+}	N_2	Nm^3/h	HHV	SG	W
FCC dry gas	0.14	0.32	0.44	0.10	8000	55.00	0.869	59.01
coker gas	0.10	0.48	0.38	0.04	5000	55.34	0.817	61.23
hydrotreater purge	0.85	0.09	0.04	0.02	4000	18.10	0.182	42.48
reformer net gas	0.90	0.04	0.05	0.01	1500	17.67	0.161	44.04
relief load	0.20	0.40	0.30	0.10	900	46.07	0.731	53.88
natural gas	0	0.93	0.06	0.01	1009	42.53	0.605	54.70
header (design)	0.321	0.327	0.293	0.059	20409	44.10	0.650	54.69

$\mathrm{d}\ln W/\mathrm{d}a = -0.711 + 0.447 = -0.264$ at $a = 0$, the first term from the heating value and the second from half the fall in specific gravity. Adding 5, 10 and 20 % hydrogen lowers the heating value 3.55, 7.11 and 14.22 % and the Wobbe index only 1.33, 2.66 and 5.35 %, as shown in Figure 14.2a. Natural gas, at 54.70 MJ/Nm^3, has the design gas's Wobbe index to within 0.01 MJ/Nm^3. That equality is a coincidence of the chosen teaching compositions and not a property of refinery gas. Three candidate designs during the build gave header indices of 54.69, 55.17 and 54.09 MJ/Nm^3, and nothing was tuned toward the natural gas value.

14.3.3 Makeup, Flare Valve and the Split Range

One output u drives both valves through the split range of Equation (14.4).

$$\begin{aligned} x^*_{MU} &= \mathrm{clip}\Big(\frac{50 - \delta/2 - u}{50 - \max(0, \delta/2)}\Big), \\ x^*_{FL} &= \mathrm{clip}\Big(\frac{u - 50 - \delta/2}{50 - \max(0, \delta/2)}\Big) \\ \frac{\mathrm{d}x}{\mathrm{d}t} &= \mathrm{clip}\Big(\frac{x^* - x}{\tau_x},\ \pm r_v\Big) \\ F_{MU} &= F_{MU,\max}\, x_{MU} \\ F_{FV} &= F_{FL,\max}\, x_{FL}\, (p + p_a)/(p_0 + p_a) \end{aligned} \tag{14.4}$$

where clip limits to $[0, 1]$, δ is the dead band in percent of output, x^* the command, τ_x the stem's lag (5 s for the makeup and 3 s for the flare valve), $r_v = 10\,\%/\mathrm{s}$ the stroking limit, $F_{MU,\max} = 8000\,\mathrm{Nm^3/h}$, $F_{FL,\max} = 15000\,\mathrm{Nm^3/h}$ at the design pressure p_0, and the stems are held at their stops. A positive δ leaves a band of output around 50 % in which neither valve opens. A negative δ is an overlap in which both do, and the case 2 and challenge settings as shipped carry $\delta = -4\,\%$, so the makeup shuts at 52 % and the flare valve opens at 48 %, as shown in Figure 14.1b. Both valves are choked, the makeup from its supply and the flare valve into the flare, so neither flow depends on the header except through the flare valve's absolute upstream pressure.

14.3.4 Flare Header, Water Seal and Recovery Compressor

The flare header holds its gas behind the water seal, and a loaded compressor stage recycles through its spillback whatever its suction cannot supply, as shown in Equation (14.5).

$$\begin{aligned} c\,\frac{\mathrm{d}n_F}{\mathrm{d}t} &= F_{rel} - F_{RC} - F_{seal}, \qquad p_F = \frac{n_F RT}{V_F} - p_a \\ F_{RC} &= N\,F_{st}\,\mathrm{clip}\big((p_F - p_{\min})/\Delta p_{sb}\big) \\ F_{seal} &= k_{seal}\,\max(0,\ p_F - p_{seal}) \end{aligned} \tag{14.5}$$

where p_F is in mbar gauge, $V_F = 400\,\mathrm{m^3}$, N is the number of loaded stages, $F_{st} = 1200\,\mathrm{Nm^3/h}$ per stage, $p_{\min} = 5\,\mathrm{mbar}\,g$, $\Delta p_{sb} = 15\,\mathrm{mbar}$, $k_{seal} = 500\,\mathrm{Nm^3/h}$ per mbar and $p_{seal} = 60\,\mathrm{mbar}\,g$. With

N stages loaded and enough capacity the flare header settles at $p_{\min} + \Delta p_{sb} F_{rel}/(N F_{st})$, which is 16.25 mbar g for one stage at 900 Nm^3/h and 15.0 mbar g for two at 1600 Nm^3/h. With no stage loaded the 900 Nm^3/h relief load lifts the seal and holds the flare header at $60 + 900/500 = 61.8$ mbar g. The flared flow is $F_{flare} = F_{seal} + F_{FV}$. The header's flare valve discharges past the seal, so the recovery compressor never takes back what the split range flares. Many plants route that valve into the flare header, where a compressor with spare capacity would recirculate it, and the simplification keeps the split range's flaring and the compressor's recovery separately measurable.

14.3.5 Burners, H-1 and the Lumped Heaters

Every burner manifold sits behind a fuel valve, and valve and burner tips both pass gas as the square root of the difference of squared pressures over the specific gravity. The flow at a burner pressure p_b and the heat it releases are shown in Equation (14.6).

$$\begin{aligned} F &= F_0\,\phi(p_b)\sqrt{SG_0/SG}, \qquad \phi(p_b) = \sqrt{\frac{p_b^2 - p_a^2}{p_{b0}^2 - p_a^2}} \\ Q &= \mathrm{HHV}\,F/3600 = Q_0\,(W/W_0)\,\phi(p_b) \\ p_{b,\max} &= \sqrt{\frac{r^2 P^2 + p_a^2}{r^2 + 1}}, \qquad r^2 = \frac{p_{b0}^2 - p_a^2}{u_0^2\,(P_0^2 - p_{b0}^2)} \end{aligned} \tag{14.6}$$

where pressures are absolute, $P = p + p_a$ is the header's absolute pressure and P_0 its design value, F_0, Q_0, SG_0 and W_0 are design values, $p_{b0} = 2.2$ bar absolute, $u_0 = 0.6$ is the fuel valve opening at design, and $p_{b,\max}$ is the burner pressure with the fuel valve wide open, 2.978 bar absolute at design ($r^2 = 0.682$). At a fixed burner pressure the heat is proportional to W, which is why the Wobbe index rather than the heating value is the property a burner answers to.

Pause. Try Exercise 14.2 before reading on.

H-1's outlet loop sets a firing demand d, a fast inner loop makes the burner pressure follow the setpoint the demand asks for, and the heat passes a firebox lag and a coil to the outlet, as shown in Equation (14.7).

$$\begin{aligned} p_{b,SP} &= \sqrt{p_a^2 + (p_{b0}^2 - p_a^2)\,\big(d/(100\,r_W)\big)^2} \\ \tau_{pb}\frac{\mathrm{d}p_b}{\mathrm{d}t} &= \min(p_{b,SP},\ p_{b,\max}) - p_b \\ \tau_{fb}\frac{\mathrm{d}Q_f}{\mathrm{d}t} &= \eta Q - Q_f \\ \tau_p\frac{\mathrm{d}T_{out}}{\mathrm{d}t} &= \frac{1000\,Q_f}{\dot{m}c_p} - (T_{out} - T_{in}) \end{aligned} \tag{14.7}$$

where r_W is the Wobbe ratio the feedforward applies (1 with it off), $\tau_{pb} = 3$ s, $\tau_{fb} = 40$ s, $\tau_p = 90$ s, $\eta = 0.85$, $T_{in} = 220$ °C and $\dot{m}c_p = 113.33$ kW/K, sized so that $Q_0 = 20$ MW heats the coil to 370 °C. One percent of firing is 1.5 K at the outlet. H-1 is starved when $p_{b,SP}$ exceeds $p_{b,\max}$. The eighteen lumped heaters use the same burner law with $Q_0 = 230$ MW, and their outlet loops are one integral action on their heat, $\mathrm{d}h_L/\mathrm{d}t = \max(0.005, h_L)(Q_{L,dem} - Q_L)/(Q_{L,dem}\tau_L)$ with $\tau_L = 3$ min and $Q_{L,dem} = 230\,\mathrm{MW}\,d_L/100$, held when they starve. H-20, when lit to a fraction s, burns at a fixed burner pressure, $F_{su} = s\,(3600\,Q_{SU}/\mathrm{HHV}_0)\sqrt{SG_0/SG}$ with $Q_{SU} = 18$ MW, so it draws 1473 Nm^3/h at the design gas.

14.3.6 The Coker's Drum Switch

A delayed coker switches its effluent to a cold drum at the end of each cycle, and the new drum condenses part of the vapor. The application reproduces that as a teaching shape rather than importing the coker model of Chapter 11. At a switch the coker gas falls to 40 % of its rate over 30 s, stays there

20 min and recovers linearly over the next 40 min. Switches a scenario publishes are known from $t = 0$, so the MPC may preview them. One fired from the panel joins the plant's schedule when it happens and is never announced to the MPC.

14.3.7 Instruments and What Is Counted

Instruments sit between the plant and every controller. The header pressure has a 1 s lag and 0.003 bar of noise, the flare header 1 s and 0.3 mbar, every flow meter a 2 s lag and 0.3 % noise, and the ultrasonic meter at the flare tip 2 %. The Wobbe meter is a combustion calorimeter with a hydrogen cell that samples every 10 s, analyzes for 10 s behind a 10 s sample line and reports with 0.03 MJ/Nm^3 of noise. H-1's outlet reads through 15 s of delay, a 5 s lag and 0.1 K of noise, and the valve positioners report their stems. The flared volume, its carbon dioxide and the import-and-flare time are integrated from the true flows, as shown in Equation (14.8).

$$\begin{aligned} V_{flare} &= \int F_{flare}\, dt \\ m_{CO_2} &= \int \tfrac{44.01}{22.414} \left(F_{seal} C_{rel} + F_{FV} C\right) dt \\ t_{I\&F} &= \int \left[F_{MU} > 50\right] \left[F_{flare} > 50\right] dt \end{aligned} \tag{14.8}$$

where C is carbon atoms per molecule, 1.18 for the relief gas, and the brackets are 1 when true. Import and flare is therefore counted from the flows, both above 50 Nm^3/h at the same instant, and never inferred from the valve positions. The carbon dioxide assumes complete combustion with no destruction efficiency and no methane slip, a teaching assumption that measured flares do not meet [11]. Gas is valued at 6 $/GJ for natural gas bought and for fuel gas flared, a teaching price.

14.3.8 Parameters, Initialization and Numerics

Table 14.3 gives every parameter with its default. A case opens lined out at its own settings and disturbances. The flare header takes the smallest stage count its load and unload thresholds allow, a bisection on u finds the split-range output at which makeup or flare valve closes the header's balance at the setpoint, the heaters' demands are set so that each delivers its duty with the Wobbe meter lined out on the starting gas, and the manual stations are placed on those values so that Manual from $t = 0$ holds. At design the output sits at 41.07 % with the makeup valve 12.6 % open. Fourth-order Runge-Kutta steps of 0.5 s advance the states, the controllers execute every 1 s, and the log keeps a row every 5 s. Each lump, the flare header and the metered totals close over a disturbed hour to better than 10^{-9} kmol. Halving the step moves no verdict of the challenge and no figure of it by more than 0.002 bar or half a second of import and flare, and in case 1 it moves the time to 3.0 bar g from 45.0 to 44.8 s.

14.3.9 Assumptions

The header is one well-mixed isothermal volume, so pressure waves, transport delay from the drum to a burner and temperature changes on compression are absent, and a composition change reaches every burner within the 41 s residence time. Each source sends what its unit makes whatever the header pressure, which holds while the units control their own pressure upstream and fails if the header rose high enough to back gas into them. Gas is four lumps. Their heating values and molar masses are teaching values close to those of the pure components, and the C_{2+} lump's 92.0 MJ/Nm^3 sits nearer a mean carbon number of 2.7 than the 2.6 its molar mass implies, by the application's technical review. The standard calculation of heating value, density and Wobbe index from a full composition is ISO 6976, which the model does not use. The burner and valve law is an isothermal subsonic one, and the makeup and flare valves are treated as choked. Heater efficiency is constant on the higher heating value, so the latent heat in the water a hydrogen-rich fuel makes is not tracked. Only H-1 is simulated in full, the eighteen lumped heaters share one integral loop, and H-20 has no starvation check, which no case reaches. Every safety-related number, the seal, the stage thresholds,

Table 14.3 Parameters and Defaults. A, Adjustable in the Parameters Panel. C, a Control on a Card. F, Fixed in the Source

Symbol	Meaning	Value	Unit	
V	header volume (30, 60, 120, 240)	60	m^3	C
T, p_0	header temperature, design pressure	40, 3.5	°C, bar g	A
$F_{MU,\max}$, τ_{MU}	makeup valve capacity, lag	8000, 5	Nm^3/h, s	A
$F_{FL,\max}$, τ_{FL}	flare valve capacity, lag	15000, 3	Nm^3/h, s	A
r_v	valve stroking limit	10	%/s	A
V_F, p_{seal}	flare header volume, water seal	400, 60	m^3, mbar g	A
k_{seal}	flow over the seal	500	$Nm^3/h/mbar$	A
F_{st}	recovery capacity per stage	1200	Nm^3/h	A
$p_{\min}$, Δp_{sb}	spillback floor, span	5, 15	mbar	A
Q_0, η	H-1 firing at design, efficiency	20, 0.85	MW, —	A
T_{in}, $T_{out,0}$	H-1 inlet, design outlet	220, 370	°C	A
τ_{fb}, τ_p	firebox lag, coil time constant	40, 90	s	A
p_{b0}, u_0	design burner pressure, valve opening	2.2, 0.6	bar a, —	A
τ_{pb}	burner pressure loop	3	s	A
$Q_{L,0}$, τ_L	lumped heaters' firing, reset time	230, 3	MW, min	A
Q_{SU}	H-20 firing when lit	18	MW	A
F_k	FCC, coker, purge, reformer at design	8000, 5000, 4000, 1500	Nm^3/h	A
—	coker gas after a switch	40	%	A
—	natural gas price, fuel gas value	6, 6	$/GJ	A
H_i, M_i	lump heating values, molar masses	§14.3.2		F
y_k	stream compositions	Table 14.2		F
—	flowing threshold for $t_{I\&F}$	50	Nm^3/h	F

the makeup capacity and the WOBBE LOW alarm at 52.5 MJ/Nm^3, is a teaching value. API Standard 521 governs the design of real pressure-relieving and flare systems, and a plant's relief and flare studies set its numbers.

14.4 Controllers as Implemented

Within a physics step the controllers act first, then the states advance, then the plant-side bookkeeping (ramps, the coker's schedule, the import-and-flare count) runs, and the outputs are written last. The controllers execute every $\Delta t_c = 1$ s on the instruments of Section 14.3.7. One of the application's tests whitelists what the controllers may read, and another fails if the flow feedforward or the MPC reads the natural gas supply's availability, a disturbance no instrument measures.

14.4.1 The Header Pressure Loop and the Split Range

The pressure controller is the shared velocity-form PID of Section 1.6.2, used here without derivative action. Its measurement is the header pressure p_m, its output bias u_b in Equation (1.18) is the feedforward output u_{ff}, which is zero for the plain split range, its output is limited to 0 to 100 %, and its tracking time T_t equals τ_I. A high header raises the output toward the flare valve, so the loop is direct acting on u, with $a = -1$ and the error $p_m - p_{SP}$. The defaults are $K_c = 60$ % of output per bar and $\tau_I = 2$ min, an integral time of 60 min removes integral action, and the default dead band is 6 %. The output passes through Equation (14.4) to the two valve commands. In Manual the learner sets u and the PI tracks it, so a return to Auto is bumpless, as Section 1.6.2 describes.

14.4.2 Flow Feedforward

The second architecture adds a feedforward from the meters. The metered need is $\Phi = F_{h,m} + F_{L,m} + F_{su,m} - \sum_k F_{k,m} - F_{RC,m}$, and u_{ff} is the output that would deliver it, mapped back through the split range, $u_{ff} = 50 - \delta/2 - (50 - \max(0, \delta/2))\,\Phi / F_{MU,\max}$ for $\Phi \geq 0$ and the flare valve's mirror image for $\Phi < 0$. The capacity used is the makeup valve's declared one. Because u_{ff} jumps from one edge of the dead band to the other whenever Φ changes sign, the feedforward carries the meters' noise across the band. After the purge step of the challenge, when the header needs only about 125 Nm^3/h

of makeup and the summed meters carry several tens of $\mathrm{Nm^3/h}$ of noise, the output crosses 50 % 240 times in 3 h, against 4 times under the PI alone. The Manual helper *Balance the header* sets u to the same value once.

14.4.3 Model Predictive Control on Both Valves

The third architecture replaces the split range with the MPC option of Section 1.6.3. Every 2 s an MPC solves for both valve commands over 90 s on the header's integrator with its declared volume, each stem's lag and a disturbance, as shown in Equation (14.9).

$$\begin{aligned} \dot{p} &= K\left(F_{MU,\max}\, x_{MU} - F_{FL,\max}\,\tfrac{P}{P_0}\, x_{FL} + w\right) \\ \tau_{MU}\dot{x}_{MU} &= c_{MU} - x_{MU}, \qquad \tau_{FL}\dot{x}_{FL} = c_{FL} - x_{FL} \end{aligned} \tag{14.9}$$

where c are the commands and w is the metered imbalance (sources and recovered gas minus heaters) plus a bias. An observer runs the same model every 1 s on the commands sent and corrects it with the measured pressure, $\hat{p} \leftarrow \hat{p} + 0.2\,\varepsilon$ and $b \leftarrow b + 0.02\,\varepsilon/(K\Delta t_c)$ with $\varepsilon = p_m - \hat{p}$, which makes the controller offset-free. With the coker preview on, w over the horizon also carries the coker's metered gas scaled by the published switch shape. The continuous model becomes an exact zero-order-hold step of 2 s, and the quadratic program over 45 steps in 11 graduated move blocks is Equation (14.10).

$$\begin{aligned} \min\ & \textstyle\sum_k w_p (p_k - p_{SP})^2 + 10\, w_p (p_N - p_{SP})^2 \\ & + w_{\Delta u} \textstyle\sum_j \|\Delta c_j\|^2 + 10^4 s^2 \\ & + w_\$ \textstyle\sum_k (\gamma_{MU} c_{MU,k} + \gamma_{FL} c_{FL,k}) \\ \text{s.t. } & 0 \le c_{MU} \le \bar{c}_{MU}, \quad 0 \le c_{FL} \le \bar{c}_{FL} \\ & |p_k - p_{SP}| \le 0.15 + s, \quad s \ge 0 \end{aligned} \tag{14.10}$$

where $w_p = 100$ per $\mathrm{bar^2}$, $w_{\Delta u} = 1$, $w_\$ = 1$, and γ is the money each valve's gas is worth per decision step at full opening, natural gas bought for the makeup and fuel gas burned for the flare. Because both valves cost money and act on the pressure in opposite directions, the optimum never opens both. The stems lag the plan, however, so an interlock sets $\bar{c}_{FL} = 0$ while the makeup is commanded open or its positioner reads more than 0.2 % of travel, and $\bar{c}_{MU} = 0$ in the mirror case. The series' shared solver handles the program with the dual method of Goldfarb and Idnani [16], only the first planned move reaches the valves, as in industrial receding-horizon practice [17], and a solve that fails, is not finite or leaves its bounds is rejected, after which the tracked split-range PI takes the valves without a bump. On the challenge with the suggested staging and feedforward the MPC never failed in 8 h.

The interlock is necessary, and the measured reason is the stems. With the MPC's plan alone, which never opens both valves, the lagging stems pass gas both ways at the hand-overs, and the challenge counts 7.88 to 9.33 min of import and flare in 8 h across five seeds. An interlock on the commands alone still leaves 3.00 to 3.47 min. Both fail the challenge's 1 min limit, and the interlock on the positioners' feedback brings it to zero. An earlier record of the build gave half a second to 2.5 s for the case without interlock, which was not reproduced.

🕮 **Pause.** Try Exercise 14.3 before reading on.

14.4.4 Recovery Compressor Staging

The two stages load and unload on the measured flare header pressure, where the compressor takes its suction, and not on the fuel gas header controller's output. With k stages loaded the next loads when $p_{F,m}$ rises above its load threshold L_{k+1}, and the last unloads when $p_{F,m}$ falls below its unload threshold U_k. A sequencer from the series' shared core holds every start for a minimum run of 2 min and every stop for a minimum rest of 2 min, with 15 s between any two changes. The defaults are $L = 30$ and 45 mbar g and $U = 10$ and 12 mbar g, all below the seal, so the stages load before gas reaches the flare. An earlier design of the application put the stages on the controller output, loading

Table 14.4 The Nine Cases

#	Case and events	Teaching point
1	Manual, coker switch at 3 min	an integrator
2	$\delta = -4\,\%$, other heaters $-4\,\%$ at 5 to 20 min	overlap and dead band
3	FCC trips at 5 min	a balance the makeup cannot close
4	H-1 on Man, purge $+50\,\%$ at 5 min	hydrogen and the Wobbe index
5	purge $+50\,\%$ at 5, back at 30 min	Wobbe feedforward
6	relief to 1600 $\mathrm{Nm^3/h}$ at 5 to 50 min, $U_2 = 40$	staging and hysteresis
7	Manual, coker switch at 3 min, 60 or 240 $\mathrm{m^3}$	volume and gain
8	H-20 lights at 5 min	measured disturbance
9	Challenge: 8 h, coker at 1, FCC $-40\,\%$ at 3, purge at 5 h	the whole system

before its flare valve, but the compressor's suction is the flare header, and a stage on the fuel gas header's output would have no gas to take.

14.4.5 The Heater Loop and Wobbe Feedforward

H-1's outlet loop uses the same shared PI, reverse acting on the measured outlet, with $K_c = 1\,\%$ of firing per kelvin, $\tau_I = 2\,\mathrm{min}$ and an output of 0 to 130 %. Its output is a firing demand, not a fuel flow, and the burner-pressure loop of Equation (14.7) turns it into a burner pressure. With the Wobbe feedforward on, every heater's demand, H-1's and the lumped bank's, is divided by $r_W = \mathrm{clip}(W_m/W_{\mathrm{ref}},\ 0.7,\ 1.3)$ before it becomes a burner pressure, $W_{\mathrm{ref}} = 54.69\,\mathrm{MJ/Nm^3}$ being the design gas's index. Because the heat at a fixed burner pressure goes as W in Equation (14.6), this division is exact for a correctly measured W. On a volumetric fuel-flow loop the exact ratio would be the heating value instead. Chapter 3 meets an unmeasured change of heating value with its outlet loop alone, and Volume 3, Chapter 4 adds the air side of a heater burning fuel gas. The application has no Wobbe blending loop. Natural gas is already the pressure loop's makeup, so any more of it than the balance needs could only leave through the flare, a structural reason that does not depend on the coincidence of the two indices.

14.4.6 Three Models

Section 14.3 is the plant. What a learner identifies is a single number, the slope K read from case 1 or 7. The controllers' models are the PI's absence of one, the feedforward's meter balance, and the MPC's three-state integrator with the declared volume, stem lags and the makeup valve's declared capacity, which knows neither the composition nor a supply shortfall. Its numbers are the declared ones, fixed when the application loads, so another header volume or a Parameters edit is model mismatch: at 30 $\mathrm{m^3}$, twice the gain the model assumes, case 4 under the MPC cycles about 0.16 bar from its setpoint, where a model given the true volume held it within 0.015 bar. With the natural gas supply cut to half its capacity before case 8 starts, the header's worst error after the heater lights is 0.215 bar under the PI, 0.110 bar with flow feedforward and 0.015 bar under the MPC, against 0.118, 0.015 and 0.007 bar with full supply. That cost is the honest consequence of a disturbance no instrument measures.

14.5 Nine Cases, Measured

The nine cases appear in Table 14.4. All share one reference configuration (the defaults above, Auto with the plain split range, and the design disturbances), and a row lists only what differs. Every number below was taken headless from the engine of Section 1.6.1 at seed 20261016, advanced step by step as the page advances it, with a row's settings entered before the run. Each case produces the same run when loaded first, when picked after another case and after Reset. Pressure deviations are on the true header against its setpoint.

14.5.1 An Integrator With Nobody at the Valve

In case 1 the split-range output is held in Manual when the coker switches at 3 min. The pressure falls in a straight line and passes 3.0 bar g 45 s after the switch began, which is K times the 3000 Nm^3/h shortfall over the 0.5 min ramp and the quarter minute after it. Nothing stops it until H-1's fuel valve opens wide at 4.47 min. Then the heaters starve, burn less, and close the balance themselves, and the header settles at 1.74 bar g with H-1's outlet 27.6 K low at 15 min. Auto at 3.8 min, with the header at 2.94 bar g, brings it back within 0.1 bar from 8.0 min. Pressing *Balance the header* at 3.8 min only stops part of the fall, because the gas the coker stopped sending is the richest in the header. The Wobbe index falls from 54.69 to 53.67 MJ/Nm^3 by 7 min, the lumped heaters' meters rise from 20,634 to 21 178 Nm^3/h as their loops burn more of the leaner gas, and the heaters starve at 8.12 min. Case 7 repeats the switch with a 240 m^3 header, which passes 3.0 bar g after 133.5 s instead of 45 s and starves the heaters at 7.89 instead of 4.47 min, while both end near 1.75 bar g. In Auto with the same tuning the larger header dips 0.204 bar against 0.280 bar. Four times the volume quarters the rate of fall and buys time, but it does not change where the balance ends.

> **TRY IT IN THE APP** **1 · A coker drum switch, by hand (Manual)**
>
> Load the case and press Run. The split-range output stays in Manual, and at 3 min the coker sends 3000 Nm^3/h less gas. Watch the header pressure trend: it falls in a straight line and passes 3.0 bar g 45 s after the switch began. Nothing stops it until the fuel valve of H-1 opens wide at 4.47 min and the starved heaters burn less.

14.5.2 The Overlap and the Dead Band

Case 2 ships the −4 % overlap. When the other heaters fire 4 % less at 5 min, the makeup the header needs falls to a few hundred Nm^3/h and the output rises into the overlap, where both valves open. At 12 min the output is 48.5 %, the makeup stem 6.8 % open passing a metered 533 Nm^3/h, and the flare stem 1.2 % open passing 198 Nm^3/h, as shown in Figure 14.1b. Over the 30 min case the refinery imports and flares for 14.38 min, flares 49.8 Nm^3 and buys 362 Nm^3, while the header never leaves 0.025 bar of its setpoint. With any dead band from 0 to 20 % the import and flare is zero, nothing is flared, the makeup falls to 310 Nm^3, and the worst error is 0.020 to 0.026 bar, as shown in Figure 14.2b. Wider overlaps do worse, 30 min at $\delta = -8$ % with 210 Nm^3 flared, and a 1 % overlap happened to give none. The application's tests repeat the overlapped case at twelve tunings, gains of 30 to 240 % per bar and integral times of 0.5 to 10 min, and every one imports and flares for 14.2 to 15.3 min. On this integrating header with a PI loop, a positive dead band costs nothing measurable, because the integral carries the output across it once and does not return. In the challenge, with the other three suggested changes in place, a dead band of exactly zero still imports and flares for 4.5 s in 8 h, because a closing stem lags its command by seconds at each crossing, while 2 and 12 % give none.

14.5.3 A Balance No Tuning Can Close

In case 3 the FCC, the largest source at 8000 Nm^3/h and 48.9 % of the header's heat, trips at 5 min. The makeup valve, already passing about 1000 Nm^3/h, reaches its 8000 Nm^3/h 1.35 min after the trip began, and natural gas at 42.53 MJ/Nm^3 replaces only part of FCC gas at 55.00 MJ/Nm^3. The header keeps falling, the heaters starve at 6.82 min, and it settles at 1.838 bar g with H-1 23.8 K low. The tests sweep the gain from 30 to 240 % per bar and the integral time from 0.5 to 10 min, and every tuning lets the header fall below 3.0 bar g. By hand, the wide-open makeup supports 210.4 MW with the FCC gone, so the lumped heaters must fall to 82.8 % of their firing. Shed at 10 min to 84 % they change nothing, because starved they already burn less. At 82 % the header reaches only 2.83 bar g by 50 min, at 81 it is back within 0.1 bar at 38.8 min, and at 80 % at 30.9 min.

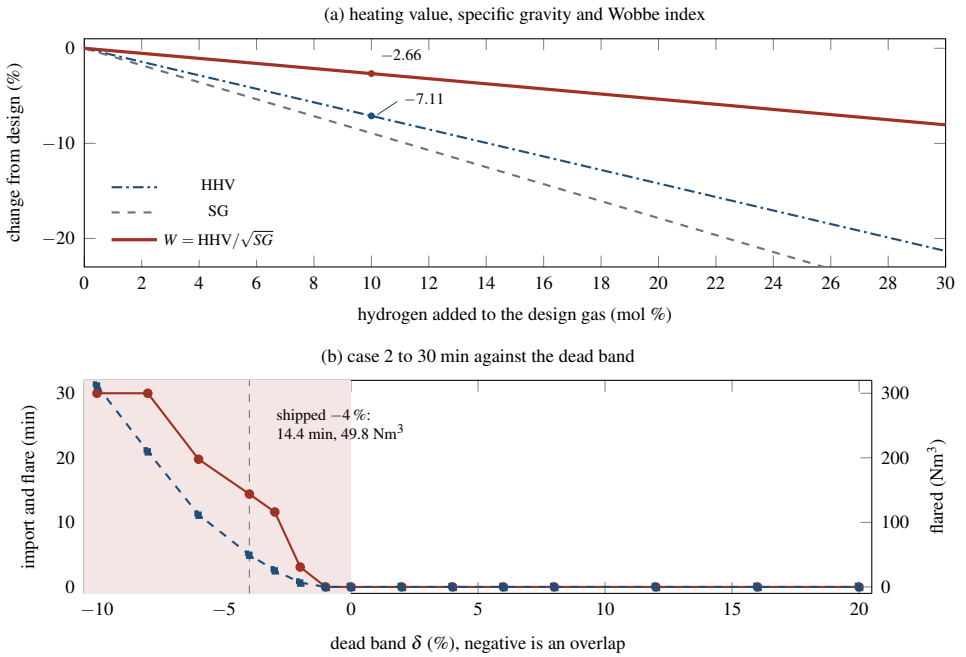


Figure 14.2 Two characteristics computed from the application. (a) Heating value, specific gravity and Wobbe index of the design gas as pure hydrogen is added: heating value dash-dotted, specific gravity dashed, Wobbe index solid. At 10 % the heating value falls 7.11 % and the Wobbe index 2.66 %. (b) Case 2 run to 30 min at each dead band, the import-and-flare time (solid with circles, left) and the volume flared (dashed with squares, right). Overlaps are shaded. Every $\delta \geq -1$ % gives zero.

14.5.4 Hydrogen, the Wobbe Index and the Feedforward

In case 4 the purge rises by half at 5 min with H-1's loop on Man. The header's hydrogen rises from 32.09 to 38.33 %, the heating value falls 5.33 % from 44.097 to 41.746 MJ/Nm^3, the specific gravity falls from 0.6501 to 0.6085, and the burners pass 3.37 % more gas at the same pressure, so the Wobbe index falls only 2.14 %, from 54.69 to 53.517 MJ/Nm^3. H-1 drops from 20.00 to 19.57 MW and its outlet settles 3.22 K low, below 368 °C at 8.49 min. The meter's 53.48 MJ/Nm^3 at 30 min is 2.21 % below W_{ref}, which is what the Quality view prints. With the loop in Auto the demand rises 2.19 % and the outlet returns. In case 5, the same step with the loop in Auto and back at 30 min, the outlet strays 1.23 °C 3.3 min after the step without the feedforward and 0.36 °C with it, and 1.23 against 0.33 °C on the way back. Across five seeds the two are 1.22 to 1.25 and 0.33 to 0.38 °C. What the feedforward leaves is the 20 s to 30 s by which the meter's report trails the gas at the burners.

14.5.5 Staging Against the Seal

Case 6 raises the relief load from 900 to 1600 Nm^3/h for 45 min, more than one 1200 Nm^3/h stage takes, and ships stage 2 unloading below 40 mbar g. Two loaded stages hold the flare header at 15.0 mbar g, well under that threshold, so stage 2 unloads as soon as its 2 min minimum run ends and loads again when one stage lets the header climb back to 45 mbar g. It starts 11 times, and the recovered gas, returned to a fuel gas header that is already nearly balanced, leaves through the header's flare valve, 51.3 Nm^3 in all. With the unload threshold at 12 mbar g, below where two stages hold the suction, stage 2 starts once and 3.0 Nm^3 are flared. Stage 1 loading above 70 mbar g, above the seal, never loads, and $900 \cdot 15/60 + 1600 \cdot 45/60 = 1425$ Nm^3 goes over the seal in the hour, exactly what the engine counts.

Table 14.5 Measured Cases. Ranges Cover Seed 20261016 and Seeds 0, 1, 2 and 3 Where Those Were Run

Case	Configuration	Result
1	Manual, 60 m^3	3.0 bar g after 45 s, starved 4.47 min
7	Manual, 240 m^3	3.0 bar g after 133.5 s, starved 7.89 min
7	Auto, 60 / 240 m^3	dip 0.280 / 0.204 bar
2	$\delta = -4$ %	$t_{I\&F}$ 14.38 min, 49.8 Nm^3 flared
2	$\delta = 0$ to 20 %	$t_{I\&F}$ 0, 0 Nm^3 flared
3	any of 12 tunings	below 3.0 bar g, settles 1.838 bar g
4	H-1 on Man	W −2.14 %, outlet −3.22 K
5	feedforward off / on	outlet 1.22–1.25 / 0.33–0.38 °C
6	U_2 40 / 12 mbar g	11 / 1 starts, 51.3 / 3.0 Nm^3 flared
8	PI / flow FF / MPC	dip 0.118–0.119 / 0.013–0.016 / 0.006–0.008 bar

Table 14.6 The Eight-Hour Challenge. Ranges Cover Five Seeds (20261016, 0, 1, 2, 3), Each at Δt and $\Delta t/2$

Settings	$\lvert p - p_{SP}\rvert$ (bar)	Flared (Nm^3)	$t_{I\&F}$ (min)	$\lvert T - T_{SP}\rvert$ (°C)	Verdict
shipped	0.303–0.305	7203	480	1.22–1.23	fails
suggested	0.153–0.156	6.1–6.3	0	0.32–0.35	passes
with $\delta = -4$ %	0.162–0.165	958–959	179.6–179.7	0.32–0.35	fails
with stages 65 / 80	0.153–0.156	7200	480	0.33–0.36	fails
with feedforward off	0.153–0.155	15.6–15.8	0	1.21–1.23	fails
with 60 %/bar	0.292–0.294	5.6–5.8	0	0.32–0.35	fails
flow FF, 60 %/bar	0.060–0.063	8.2–8.6	0–0.07	0.32–0.35	passes
MPC, preview	0.009–0.019	6.1–6.5	0	0.32–0.35	passes
MPC, no preview[a]	0.013–0.019	6.1–6.5	0	0.32–0.35	passes

[a] At Δt only, five runs.

TRY IT IN THE APP 6 · Recovery compressor staging and hysteresis

Load the case and press Run. At 5 min the relief load rises to 1600 Nm^3/h, more than one stage takes. Two stages hold the flare header at 15.0 mbar g, below the 40 mbar g at which stage 2 unloads, so it starts 11 times. Set Stage 2 unloads below to 12 mbar g, press Reset and run again: it starts once, and 3.0 Nm^3 are flared instead of 51.3.

14.5.6 A Heater Lights

Case 8 lights H-20 over a minute at 5 min, 1473 Nm^3/h more gas. The PI alone lets the header dip 0.118 bar, 1.26 min after the burners begin to light, and 0.118 to 0.119 bar at seeds 1 to 3. Flow feedforward sees the new heater's meter and opens the makeup at once, and the dip is 0.015 bar (0.013 to 0.016). The MPC, reading the same meters with its model of the stems, holds 0.007 bar (0.006 to 0.008). The panel's *Heater start-up* button is different, because it ramps the lumped bank's demand 8 % over 2 min, and their loops draw the extra gas over about ten minutes, 1473 Nm^3/h at 10 min and 1556 Nm^3/h at 20 min, so the header dips only 0.044 bar.

14.5.7 The Challenge

The challenge runs 8 h: a published coker switch at 1 h, the FCC gas 40 % down from 3 h for half an hour, and the purge up by half from 5 h. Its target, judged on the truth every physics step, is the header within ±0.2 bar of setpoint after the first minute, at most 500 Nm^3 flared, at most 1 min of import and flare, H-1's outlet within ±1 °C and no heater starved. It ships with the −4 % overlap, stages that load at 65 and 80 mbar g, above the seal, the Wobbe feedforward off and 60 % per bar. The outcomes over five seeds and two steps, 85 runs in all, are in Table 14.6, and Figure 14.3 sets them against the four targets.

As shipped, the stages never load because the seal holds the flare header at 61.8 mbar g, below the 65 mbar g at which stage 1 would, so the whole 900 Nm^3/h relief load burns while the header buys natural gas. The flare budget is gone in 33.3 min, 7203 Nm^3 are flared by 8 h, and the import-and-flare

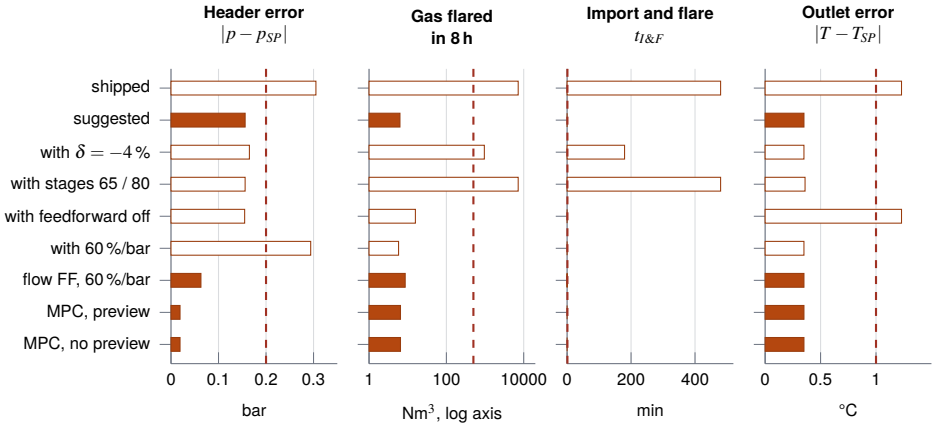


Figure 14.3 The eight-hour challenge of Table 14.6 against its four targets, one panel each: the header's worst pressure error, the gas flared (logarithmic axis), the import-and-flare time and H-1's worst outlet error. Solid bars pass the challenge and white bars fail it. Where the table gives a range the bar reaches its upper end, and a dashed line marks each target. Notice that dropping the dead band or the staging leaves the pressure error almost unchanged while the flaring and the import-and-flare time rise, and that dropping the feedforward or the higher gain fails only the outlet or only the pressure.

count runs the full 480 min. The coker switch takes the header 0.294 bar below setpoint at 60.95 min, the FCC upset 0.303 bar at 181.3 min, and the outlet strays as far as 1.23 °C. The suggested settings make four changes, a 6 % dead band, stages that load at 30 and 45 mbar g, the Wobbe feedforward on and 120 % per bar, and pass with the header within 0.154 bar, 6.1 Nm^3 flared, no import and flare and the outlet within 0.34 °C. Each change is necessary. Dropping any one fails at all ten of its runs, the overlap with 179.7 min of import and flare once the purge brings the header near balance, the staging with 7200 Nm^3 flared, the feedforward with the outlet 1.23 °C off, and the gain with the FCC upset taking the header 0.292 bar off. The 1 °C band is tighter than the 3 °C first planned for the application, for a measured reason. The worst outlet error without feedforward is 1.23 °C, so a 3 °C band would be met by any loop and the feedforward would teach nothing.

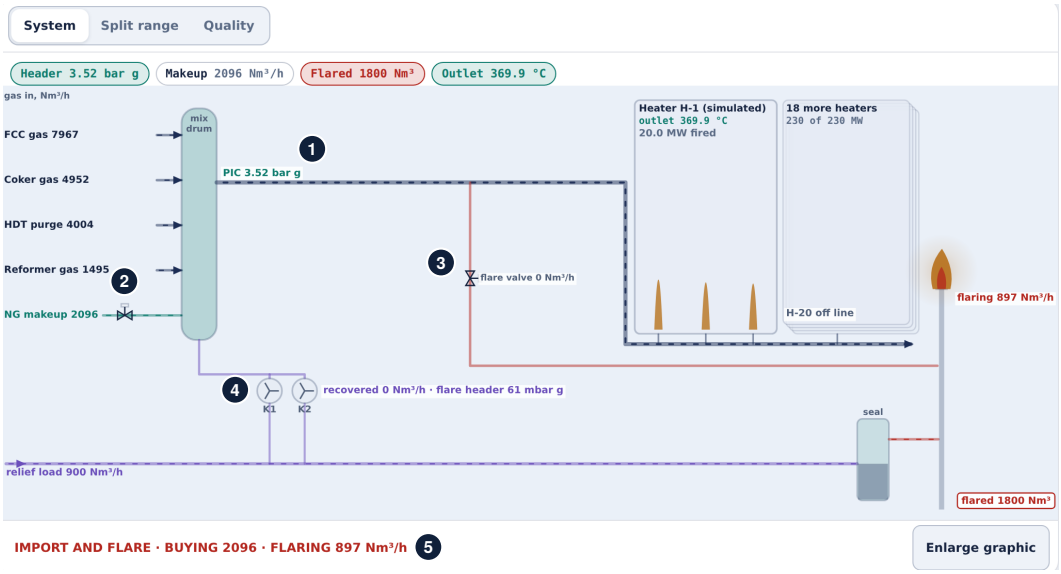


Figure 14.4 The challenge as shipped at $t = 2$ h on the System view, light theme, cropped to the plant. (1) The header reads 3.52 bar g on its 3.50 bar g setpoint. (2) The natural gas makeup passes a metered 2096 Nm^3/h. (3) The header's flare valve is shut. (4) The stages sit idle with the flare header at 61 mbar g, so the flame is lit at 897 Nm^3/h over the seal and the counter reads 1800 Nm^3. (5) The status strip reads IMPORT AND FLARE.

> **DID YOU KNOW? Stages Above the Seal Never Start**
>
> From the simulator (Table 14.6). With every suggested setting in place except the recovery stages, which load at 65 and 80 mbar g, above the 60 mbar g water seal, the challenge flares 7200 Nm^3 in 8 h against 6.1 to 6.3 Nm^3, and imports and flares for all 480 min. The header pressure error is the same, 0.153 to 0.156 bar.

Flow feedforward at the default 60 % per bar also passes, within 0.060 to 0.063 bar, but its deadband crossings import and flare for 0.5, 3, 0.5, 4 and 0.5 s at the five seeds and 0, 3, 0.25, 4 and 0.25 s at half the step, at most 4 s against the 1 min limit. The MPC with the coker preview holds the header within 0.009 bar at the default seed and 0.013 to 0.019 bar at the others, and without the preview within 0.013 to 0.019 bar. At the top speed of 120 min/s the 8 h play in 4.0 s in a browser (4.02 s measured in Chrome at both 1366×768 and 390×844, 120 min of simulated time per second of wall time). The shipped challenge at 2 h is shown in Figure 14.4.

14.6 Technology in Practice

Figure 14.5 draws the header pressure loop, the heater loop with its Wobbe feedforward and the recovery staging of Section 14.4 with the hardware a refinery installs.

What Measures It

The pressure loop reads one transmitter on the header, the feedforwards read the flow meters and the Wobbe meter, and the staging reads the flare header. The federal rule for refinery flares asks for the same pair of gas quality instruments that the chapter's Wobbe meter combines: a calorimeter that continuously records the heating value of the vent gas and, if the owner chooses, a hydrogen monitor, whose reading corrects the calorimeter's [18]. The flow to a flare is hard to meter, because the quiet

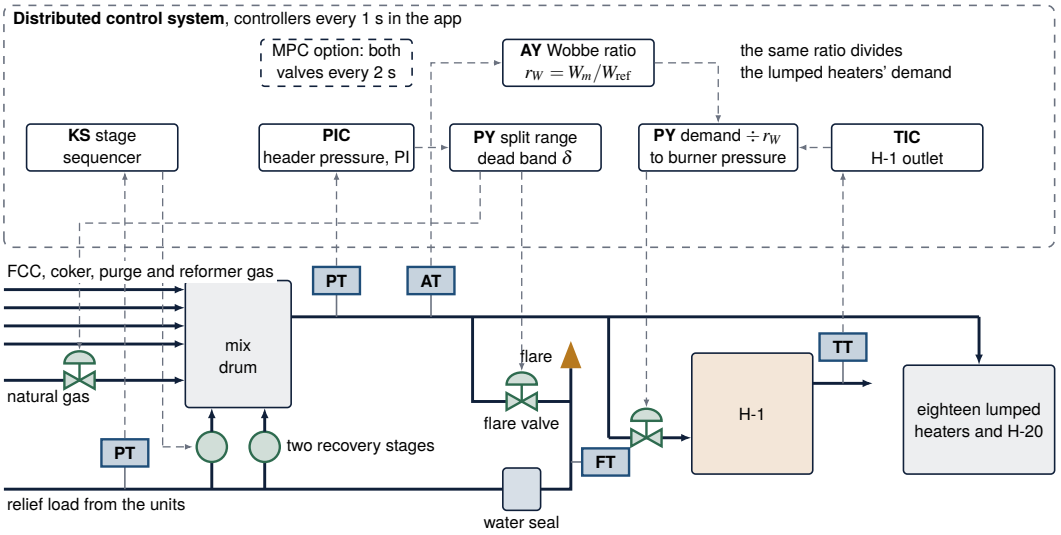


Figure 14.5 The loops of the fuel gas header with their hardware. Lettered boxes are sensing devices: the header pressure transmitter (PT), the Wobbe meter (AT), the H-1 outlet thermocouple (TT), the flare header pressure transmitter (PT) and the flare flow meter (FT). Valve symbols and the two compressor stages are final elements, and dashed arrows are signals. Notice that one controller output reaches two valves through one split-range block, and that the recovery stages answer the flare header, not the fuel gas header. The positioners' stem readings, the burner pressure transmitter and the MPC's signals are left out. Not to scale, and not a piping diagram.

flow of leaking relief valves and the full flow of an emergency call for a very large turndown, and the gas changes in composition. Velocity meters such as ultrasonic ones are less affected by composition than pressure-based or thermal meters, and multi-path transit-time ultrasonic meters show promise, at a high cost on large flare pipes [19]. The chapter's flare tip meter is ultrasonic with 2 % of noise.

> **TECHNOLOGY** **Sensor: The Header Pressure Transmitter**
>
> A common electronic pressure sensor is a taut metal diaphragm between two fixed plates, so that a pressure difference flexes the diaphragm and changes the two capacitances [19]. A needle valve between the process and the instrument damps pulsation, because the restriction and the instrument together form a low-pass filter. In gas service the transmitter is mounted above its tap, so that liquid drops cannot collect in the impulse line and add a false pressure [19]. A controller acts on such an error as on a true pressure, because it sees only the instrument, as the controllers of this chapter do. The chapter's transmitter has a 1 s lag and 0.003 bar of noise.

What Moves It

The pressure loop moves the makeup and flare valves, the heater loop moves the burner pressure, and the sequencer starts and stops the two compressor stages. A common way to split-range two valves is to give each its own 4–20 mA output from the same controller, which also leaves each smart positioner on its own circuit [19]. That positioner is the readback the chapter's MPC interlock waits for.

TECHNOLOGY Final Element: Two Valves on One Output

Valves that follow one controller output are split-ranged. In exclusive sequencing both are shut at an output of 50 % and one opens toward each end, so that the fluid passes one valve or the other but never both, which suits two valves whose effects cancel [19]. The ranges can be set in each valve's positioner or transducer, or in the controller, which then sends each valve its own signal. Where the sequencing sits decides the fail-safe state: sequencing after the controller makes the two valves fail in opposite states, and two controller outputs let both fail closed [19]. The chapter's makeup and flare stems lag their commands by 5 and 3 s and stroke at most 10 %/s.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisory level oversees local controllers and cycle times range from milliseconds to minutes [20]. The app executes its controllers every 1 s and its MPC every 2 s, inside that range. The pressure controller is a PI block whose output feeds a split-range block. One fieldbus standard has a function block for exactly this, which takes one control signal and drives a separate signal to each of two valves [19]. The Wobbe division is a calculation block ahead of each burner pressure loop, and the stage sequencer is logic with its own thresholds and timers.

IN PRACTICE A Written Plan for Every Flare

A United States refinery flare under Subpart Ja of 40 CFR Part 60 needs a written flare management plan. It lists the fuel gas systems connected to the flare, weighs a flare gas recovery system among the ways to cut flaring, and sets procedures to reduce flaring when the refinery has more fuel gas than it needs. For staged recovery compressors it states the longest time the next compressor takes to begin recovering gas [21]. Where the flare is watched by its line pressure against the liquid depth of its water seal, every period above that depth outside compressor staging calls for a root cause analysis [21]. The chapter's stages load below its seal for that reason.

What Is Modernizing It

Three developments touch this header. The first is continuous flare monitoring. Since January 30, 2019, a refinery flare under the federal rule has measured the flow in the flare header continuously, kept a pilot flame lit at all times and held the heating value of the gas in the combustion zone at or above a limit in each 15 minute block [18]. The second is the flare management plan of the in practice box above, which puts flare gas recovery and the staging of its compressors on paper. The third is predictive control, with more than 4600 applications reported and refining the largest share [17], which can move both valves of a header at once, as the chapter's MPC does.

14.7 Limitations

The sizes, compositions, prices and thresholds are generic teaching values chosen for a consistent design point, a 250 MW heater population on about 20 400 Nm^3/h of gas, and none was fitted to a refinery. The header is one isothermal volume with no transport delay, the gas is four lumps whose heating values are not computed by ISO 6976, and the C_{2+} lump's heating value is somewhat high for its molar mass. Only one heater is simulated in full, the eighteen others share one integral loop, and H-20 has no starvation check. The header's flare valve bypasses the seal and the compressor by construction, the coker's switch is a teaching shape, and the flare's carbon dioxide assumes complete combustion that measured flares do not achieve [11]. Combustion effects of hydrogen beyond the Wobbe index, such as flame speed and flashback, are outside the model.

The controllers carry the declared volume (the header as built), the declared valve capacities and the meters, so model error reaches them only through what a case changes, and a supply shortfall no instrument measures shows up only in the pressure. The MPC's model is linear in the valve positions, rebuilt every 2 s, and carries no guarantee beyond its horizon. There is no Wobbe blending loop and

no Wobbe band in the MPC or the challenge, and the equality of the natural gas and header indices that makes blending pointless here is a coincidence of the compositions. Every number is from one model at one default seed, with four more seeds and a halved step checked for the challenge and the switching cases.

OPEN CHALLENGE A Burner Measure Beyond the Wobbe Index

A 2026 review calls the Wobbe index one of the most important indicators of whether one gas can replace another, because it follows the heat a burner delivers. Hydrogen lowers the index but raises the burning speed sharply. Its laminar flame speed is about six times that of methane, and flashback becomes a critical factor limiting burner operation as its fraction rises [22]. The 1977 utility study had already found the interchangeability indexes of its day unable to predict hydrogen blends [15], and the review still lists combustion stability control among the breakthroughs required [22]. Progress would be a measured quantity, beside the Wobbe meter, that tells a controller how close a burner is to flashback.

14.8 Conclusion

A refinery fuel gas header is an integrator between sources and heaters that nobody schedules for it, and this simulator logs its pressure, its makeup, its flaring and its gas quality side by side from an ideal-gas balance on four lumps. A coker switch takes the 60 m^3 header from 3.5 to 3.0 bar g in 45 s, at the rate the declared volume predicts. A 4 % overlap in the split range imports and flares for 14.38 min of a 30 min case while the pressure stays within 0.025 bar, and any dead band of zero or more removes it. Ten percent hydrogen lowers the heating value 7.11 % and the Wobbe index 2.66 %, and the Wobbe feedforward holds the heater within 0.36 °C where feedback alone lets it stray 1.23 °C. Recovery stages placed above the seal never load, and the challenge needs all four suggested changes at every seed. An MPC whose plans never open both valves still imports and flares for minutes unless it waits for the positioners. The model is lumped and its compositions are teaching values. Further work would add transport delay along the header, a distributed heater population, a burner model that responds to flame speed as well as the Wobbe index, and destruction efficiency at the flare. The simulator shows that a split range is a statement about money as well as about pressure.

KEY POINTS

- The header holds well under a minute of its gas, so a coker drum switch takes it from 3.5 to 3.0 bar g in 45 s.
- A 4 % overlap in the split range imports and flares for 14.38 min of a 30 min case while the pressure stays within 0.025 bar, and any dead band of zero or more removes it.
- The Wobbe feedforward holds the heater within 0.36 °C where feedback alone lets it stray 1.23 °C.
- Recovery stages placed above the seal never load, and an MPC whose plans never open both valves still imports and flares for minutes unless it waits for the positioners.

Where else this applies. Volume 4, Chapter 5 splits one output between the heating and cooling of a barrel zone, and Volume 4, Chapter 9 uses valve-position control on a site with an import limit.

References

[1] A. Reyes-Lúa, C. Zotică, K. Forsman, and S. Skogestad, “Systematic design of split range controllers,” *IFAC-PapersOnLine*, vol. 52, no. 1, pp. 898–903, 2019.

[2] A. Reyes-Lúa and S. Skogestad, “Multi-input single-output control for extending the operating range: Generalized split range control using the baton strategy,” *Journal of Process Control*, vol. 91, pp. 1–11, 2020.

[3] F. G. Shinskey, *Process Control Systems: Application, Design, and Tuning*, 4th ed. New York, NY: McGraw-Hill, 1996.

[4] J. Klimstra, “Interchangeability of gaseous fuels—the importance of the Wobbe-index,” in *SAE Technical Paper Series*, no. 861578. SAE International, 1986.

[5] J. D. Zhang and G. Rong, "An MILP model for multi-period optimization of fuel gas system scheduling in refinery and its marginal value analysis," *Chemical Engineering Research and Design*, vol. 86, no. 2, pp. 141–151, 2008.

[6] L. Zhou, Z. Liao, J. Wang, B. Jiang, Y. Yang, and W. Du, "Energy configuration and operation optimization of refinery fuel gas networks," *Applied Energy*, vol. 139, pp. 365–375, 2015.

[7] G. Rong, Y. Zhang, J. Zhang, Z. Liao, and H. Zhao, "Robust engineering strategy for scheduling optimization of refinery fuel gas system," *Industrial & Engineering Chemistry Research*, vol. 57, no. 5, pp. 1547–1559, 2018.

[8] M. Zolfaghari, V. Pirouzfar, and H. Sakhaeinia, "Technical characterization and economic evaluation of recovery of flare gas in various gas-processing plants," *Energy*, vol. 124, pp. 481–491, 2017.

[9] E. Yazdani, J. Asadi, Y. H. Dehaghani, and P. Kazempoor, "Flare gas recovery by liquid ring compressors-system design and simulation," *Journal of Natural Gas Science and Engineering*, vol. 84, p. 103627, 2020.

[10] X. Lin, H. Ma, C. Liu, J. Zhang, Y. Zhang, and Z. Miao, "Experimental research on gas interchangeability indices for domestic fully premixed burners," *Fuel*, vol. 233, pp. 695–703, 2018.

[11] G. Plant, E. A. Kort, A. R. Brandt, Y. Chen, G. Fordice, A. M. Gorchov Negron, S. Schwietzke, M. Smith, and D. Zavala-Araiza, "Inefficient and unlit natural gas flares both emit large quantities of methane," *Science*, vol. 377, no. 6614, pp. 1566–1571, 2022.

[12] N. P. Cheremisinoff, *Industrial Gas Flaring Practices*. Hoboken, NJ: Wiley and Scrivener, 2013.

[13] M. R. Rahimpour and S. M. Jokar, "Feasibility of flare gas reformation to practical energy in Farashband gas refinery: No gas flaring," *Journal of Hazardous Materials*, vol. 209–210, pp. 204–217, 2012.

[14] U.S. Energy Information Administration, "Fuel consumed at refineries, U.S., annual," Petroleum and Other Liquids data series, with table definitions, 2026, eia.gov/dnav/pet/pet_pnp_capfuel_dcu_nus_a.htm, read November 2026.

[15] Public Service Electric and Gas Co., "Blending of hydrogen in natural gas distribution systems. Volume II. Combustion tests of blends in burners and appliances. Final report, June 1, 1976–August 30, 1977. [8, 11, 14, 20, 22, 25, and 31% hydrogen]," Public Service Electric and Gas Co., Newark, NJ, Tech. Rep., 1977, distributed by the Office of Scientific and Technical Information (OSTI).

[16] D. Goldfarb and A. Idnani, "A numerically stable dual method for solving strictly convex quadratic programs," *Mathematical Programming*, vol. 27, no. 1, pp. 1–33, 1983.

[17] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[18] U.S. Environmental Protection Agency, "National emission standards for hazardous air pollutants from petroleum refineries: requirements for flare control devices," Code of Federal Regulations, Title 40, Part 63, Subpart CC, section 63.670, edition of July 1, 2025, 2025, govinfo.gov.

[19] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[20] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[21] U.S. Environmental Protection Agency, "Standards of performance for petroleum refineries for which construction, reconstruction, or modification commenced after may 14, 2007: design, equipment, work practice or operational standards," Code of Federal Regulations, Title 40, Part 60, Subpart Ja, section 60.103a, edition of July 1, 2025, 2025, govinfo.gov.

[22] D. Guo, Y. Cui, P. Zhang, X. Wu, Z. Yang, and J. Bian, "Hydrogen blending in natural gas: Technologies and challenges," *ACS Omega*, vol. 11, no. 41, pp. 62 310–62 339, 2026.

14.9 Exercises

14.1 Concept check · 2 min

In case 1 the split-range output is held in Manual when the coker switches drums and sends 3000 Nm^3/h less gas. What stops the header pressure from falling?

(a) Nothing answers the pressure until the heaters' fuel valves open wide, and the starved heaters then burn less

(b) The makeup valve passes more natural gas as the header falls, so the header settles a little below its setpoint

(c) The four sources send more gas as the header falls, because their units see a lower back pressure

(d) The recovery compressor returns more relief gas, because its stages load as the header pressure falls

14.2 Concept check · 1 min

In case 4 the hydrotreater purge rises by half while H-1's outlet loop is on Man, so its burner pressure stays where it was. The heating value of the header gas falls 5.33 %. How does the heat that H-1 releases change?

(a) It falls by about the same 5.3 %, because a fixed burner pressure passes a fixed volume of gas

(b) It falls by less, about 2.1 %, because the lighter gas passes more volume at the same burner pressure

(c) It falls by more than the heating value, because the lighter gas passes less volume at the same burner pressure

(d) It does not change, because the burner-pressure loop opens the fuel valve until the heat is restored

14.3 Concept check · 2 min

The MPC of Equation (14.10) charges for the gas of both valves, so none of its plans opens the makeup and the flare valve together. Without the interlock on the positioners, the challenge still counts 7.88 to 9.33 min of import and flare in 8 h. Why?

(a) The model's header volume differs from the plant's, so the plan opens the wrong valve after each disturbance

(b) A solve fails at each hand-over, and the split-range PI with the shipped overlap takes the valves

(c) The stems lag their commands, so at a hand-over the closing valve still passes gas while the other opens

(d) The import-and-flare count reads the valve commands, and the plan sets both commands at each hand-over

14.4 Theory and calculation · 7 min

The header is one ideal-gas volume at 40 °C. Use Equation (14.2) with $R = 8.314\,\mathrm{J/(mol\,K)}$.

(a) Compute the integrating gain K of the 60 $\mathrm{m^3}$ header in bar per minute per $\mathrm{Nm^3/h}$, and the rate at which a 3000 $\mathrm{Nm^3/h}$ shortfall lowers it. *Checkpoint:* $K = 3.227 \times 10^{-4}$ and 0.968 bar a minute.

(b) The coker's gas falls by 3000 $\mathrm{Nm^3/h}$ along a straight ramp of 0.5 min and then holds. With nothing else changing, show that once past the ramp the header has fallen Δp at $t = T_r/2 + \Delta p/(K\,\Delta F)$ after the switch began, with T_r the ramp time, and find when the header passes 3.0 bar g from 3.5 bar g at 60 and at 240 $\mathrm{m^3}$. *Checkpoint:* 46.0 s and 139.0 s.

(c) The application measures 45.0 and 133.5 s. Subtract the half ramp from each measured time, compare their ratio with the ratio of the volumes, and name the flow that makes the real header fall slightly faster than the calculation. *Checkpoint:* 30 and 118.5 s, a ratio of 3.95. The heaters pass more volume of the lighter gas.

14.5 Theory and calculation · 6 min

The design gas of the header has a specific gravity $SG = 0.6501$ and a heating value HHV $= 44.097\,\mathrm{MJ/Nm^3}$. Pure hydrogen has a heating value $H = 12.75\,\mathrm{MJ/Nm^3}$ and a molar mass $M = 2.016\,\mathrm{kg/kmol}$, and air has a molar mass of 28.96 kg/kmol.

(a) Compute the Wobbe index of the design gas. *Checkpoint:* 54.69 MJ/Nm^3.

(b) A tenth of the gas is replaced by pure hydrogen, so the mixture is 0.9 design gas and 0.1 hydrogen by mole. Compute its heating value, specific gravity and Wobbe index, and the percentage change of the heating value and of the Wobbe index. *Checkpoint:* 40.96 MJ/Nm^3, 0.5921 and 53.24 MJ/Nm^3, then -7.11 and −2.66 %.

(c) Evaluate $\mathrm{d}\ln W/\mathrm{d}a$ at $a = 0$ for a fraction a of hydrogen added, and use it to estimate the change at $a = 0.1$. *Checkpoint:* $-0.711 + 0.447 = -0.264$, so about −2.64 %.

14.6 Theory and calculation · 6 min

A loaded recovery stage takes up to 1200 Nm^3/h, its spillback holds its suction above 5 mbar g, and it delivers its full flow 15 mbar above that, so N loaded stages with enough capacity settle the flare header at $p_{\min} + \Delta p_{sb} F_{rel}/(N F_{st})$. The water seal lifts at 60 mbar g and passes 500 Nm^3/h per mbar above it.

(a) Find where one stage holds the flare header with the design relief load of 900 Nm^3/h, and where two stages hold it with the 1600 Nm^3/h of case 6. *Checkpoint:* 16.25 mbar g and 15.0 mbar g.

(b) Case 6 ships stage 2 unloading below 40 mbar g and loading above 45 mbar g. Explain why it short cycles, and state the condition its unload threshold must meet. *Checkpoint:* Two stages hold 15.0 mbar g, below 40 mbar g, so the threshold must lie below 15.0 mbar g.

(c) With no stage loaded, where does the seal hold the flare header at 900 Nm^3/h, and what does that mean for the challenge's shipped load thresholds of 65 and 80 mbar g? *Checkpoint:* 61.8 mbar g, below 65, so no stage ever loads.

14.7 App challenge · 7 min

Load scenario 2, *Split range and its dead band*, which ships with Dead band at −4 %, an overlap. Run to 30 min at 1 min/s with the default seed.

(a) Read the Import and flare, Flared and Makeup tiles at 30 min. *Checkpoint:* 14.38 min, 50 Nm^3 and 362 Nm^3.

(b) Set Dead band to 6 %, press Reset and run to 30 min again. Read the same three tiles. *Checkpoint:* 0.00 min, 0 Nm^3 and 310 Nm^3.

(c) Compare the extra makeup bought in the first run with the volume it flared, and say what the pressure trend showed in each run. *Checkpoint:* 52 Nm^3 more makeup against 50 Nm^3 flared, and the pressure stayed near setpoint both times.

14.8 App challenge · 6 min

Load scenario 8, *A heater lights: a consumer surge*, with the default seed. H-20 lights at 5 min. Before pressing Run, choose Header control on the PIC card, and make three runs to 15 min at 1 min/s, one with each choice: Split range (PI), Split range + flow feedforward, and MPC on both valves. Export the CSV file after each run and read the largest $|\texttt{truth_p_barg} - 3.5|$ from 4 to 15 min, each within 0.0005 bar.

14.9 App challenge · 8 min

Load scenario 9, the challenge, with the default seed. Before pressing Run, make the four suggested changes: Dead band 6 %, the stages loading lower (Stage 1 loads above 30 mbar g and Stage 2 loads above 45 mbar g), Wobbe feedforward On, and Pressure gain 120 % per bar. Run the 8 h at 120 min/s.

(a) Read the verdict in the note at 8 h: the pressure band, the volume flared, the import-and-flare time and the outlet band. *Checkpoint:* Met: 0.154 bar, 6 Nm^3, 0.00 min and 0.34 °C.

(b) Set Pressure gain back to 60 % per bar, press Reset and run the 8 h again. Which target fails, by how much, and what upset causes it? *Checkpoint:* The pressure, 0.292 bar against 0.2 bar, in the FCC upset.

14.10 Application area · 4 min

In case 3 the FCC trips. Use the heating values and flows of Table 14.2 and a wide-open makeup of 8000 Nm^3/h. H-1 keeps its 20 MW, and the lumped heaters fire 230 MW at 100 %.

(a) Compute the heat the header can deliver with the FCC gone, on the higher heating value. *Checkpoint:* 210.4 MW.

(b) Find the firing to which the lumped heaters must be shed, and say why a shed to 84 % changes nothing once the heaters have starved. *Checkpoint:* 82.8 %. Starved, they already burn less.

14.11 Application area · 3 min

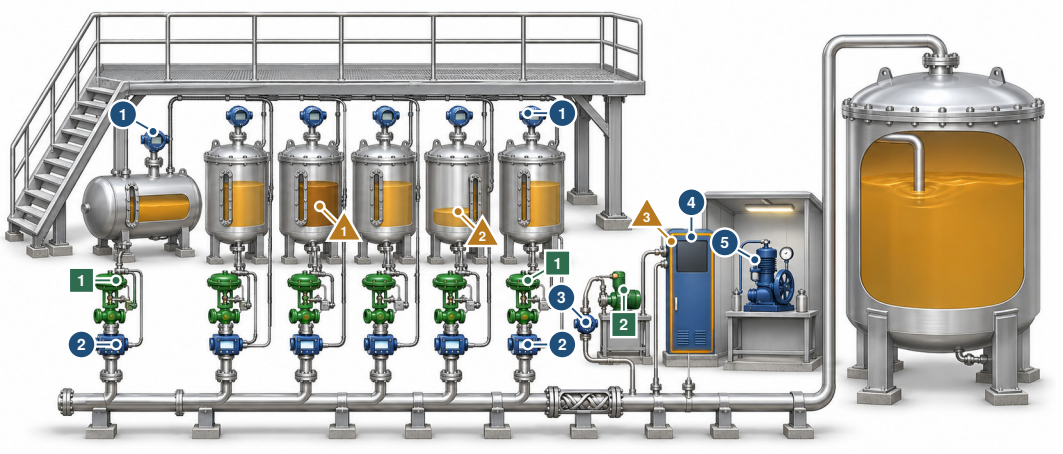
The simulated header's flare valve discharges past the water seal straight to the stack, and the Flare CO_2 tile assumes that the flared gas burns completely. Name what each simplification hides from a plant engineer.

14.12 Review the agent · 5 min

An AI assistant reviews case 2 and writes: "The header holds within 0.025 bar of setpoint, but the output keeps drifting into the 4 % overlap and the Import and flare tile counts minutes. The loop is too sluggish to keep the output out of the overlap. Raise the Pressure gain from 60 to 240 % per bar, and keep the overlap, since it ensures that no stretch of output leaves both valves shut." Find the flaw and say how you would check it in the app.

CHAPTER 15

Gasoline Blending Simulator Whose Octane Authority Shrinks as the Tank Fills



The controller moves 1 six component valves, which set the recipe 2 additive pump

The controller reads 1 level in each component tank 2 flow of each component 3 additive flow 4 online analyzer 5 laboratory knock engine

What disturbs it 1 a weaker reformate tank 2 alkylate taken for another blend 3 an analyzer that drifts or freezes

Every instrument sits on the feed legs and the header at the left, and none reads the quality of the product tank at the right, whose average is what the certificate is written on. Illustration, not to scale.

IN PLAIN TERMS A Tank Remembers Every Barrel

The main idea. Gasoline is mixed from six refinery streams in one pipe that fills a large tank. It is graded once, from a sample of the full tank, so what counts is the average of all that went in.

The challenge. Poor fuel in the tank can only be made up by better fuel still to come, and the fuller the tank, the better it must be. The instrument that checks the mix reports three minutes late and can drift while it reads on target.

Why it matters. A tank that fails must be blended again, and one that passes with room to spare gives costly ingredients away.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	This page, the History boxes in Sections 15.2 and 15.5, the Did you know? boxes in Sections 15.1 and 15.6, the Try it boxes in Sections 15.4 and 15.5, and Key points in Section 15.9. No equations.
Study	50 min	The whole chapter in order, stopping at the three pause points, then Exercises 1 to 12.
Practice	15 min	This page, Section 15.6 with Figure 15.4, Section 15.7 and Section 15.8 with the Open challenge box that closes it.

The control question Gasoline is certified from a sample of its full product tank, so the blender is judged on the average of everything it has made, while every loop works on the header. How should a late analyzer, a biased blend model and a recipe optimizer be layered so that the tank lands on its octane target, when a deficit already in the tank can only be repaid by the blend still to come?

Learning objectives

- Compute the vapor pressure and octane of a recipe from the blending index and the interaction model, including the octane boost and the vapor-pressure bump of ethanol.
- Derive the header target $T + \Delta/f$ of cumulative-average targeting and predict what its stopping floor and header clip cost the certificate.
- Explain from the four-report delay why an octane trim tuned like a flow loop oscillates, and what the gain and the integral time each do.
- Show by a run that a bias update follows a drifting analyzer, and that only an independent laboratory check brings the tank back.
- Judge which control layer can answer a quality change, a component shortage and a late deficit in one batch.

Prerequisites Sections 1.4.6, 1.4.7, 1.6.2 and 1.6.4. Chapter 9 for the reformate and its inferred octane, and Chapter 10 for the alkylate.

The app https://apmonitor.com/apps/gasoline-blending/

See also Volume 9, Chapter 4 for an estimator that carries a bias and Volume 9, Chapter 8 for an optimum that sits on a moving constraint. Volume 3, Chapter 7 for a ratio blend trimmed on an analyzer. Volume 5, Chapter 3 for a loop on a slow analyzer and Volume 5, Chapter 4 for a late, drifting analyzer checked by a laboratory.

15.1 Introduction

A refinery's gasoline is made in its last unit. An in-line blender draws butane, naphthas, reformate, cracked gasoline, alkylate and an oxygenate from their tanks, meters them into one header in the proportions of a recipe, and sends the blend to a product tank that is certified once, from a sample of its full contents [1]. Two of the properties on the certificate do not blend linearly. The octane number of a blend departs from the volume average of its components' octane numbers [2], and vapor pressure is blended through an index rather than through the pressures themselves [1]. Ethanol adds a third departure, because its contribution to octane depends on the base fuel it enters [3] and a small addition raises the vapor pressure of a hydrocarbon base although its own is low [4]. Model-based real-time optimization of such a blender re-solves the recipe as the operation proceeds [5], and the feedback control and optimization of fuel blending remain an active problem [6].

This chapter corrects one belief, that a blend is on specification whenever its header is. The certificate is written on the tank, and the tank remembers every barrel. In case 4 of this chapter a tank 80 percent full averages an anti-knock index (AKI) of 86.90, one tenth below its limit, and the remaining fifth of the batch must be blended at 88.40 for the tank to land on its 87.20 target. The same deficit found with 5 percent of the batch left would need five times the correction of one found with 25 percent left. A second belief follows, that the online analyzer is the truth. It reports each parcel 3.0 min after the blend point, it can drift, and a model corrected only against it inherits its error, so a drifting analyzer takes the tank off specification while it reads on target. A step test from a recipe fraction to an octane reading would return a gain and a delay and say nothing about either belief.

✦ DID YOU KNOW? A Real Tank Found Off Specification

An industrial gasoline blend reported in 2010 began with about 20 000 m^3 of off-specification product already in the tank and 15 000 m^3 still to add [6]. Its controller ran every five minutes, and the tank came near its targets after about six hours. Case 4 of this chapter starts with 6400 m^3 and adds 1600 m^3.

The class of problem is general. Any product certified by an accumulated average has the same shape, a quality that is the integral of what has already been made, a measurement that arrives late, and a correction whose weight in the average falls as the batch grows. Cement raw materials are blended to a target chemistry by blending systems with control technologies of their own [7], and the exponentially weighted correction of a model offset that the blender's optimizer uses is the estimator that run-to-run control applies between runs in semiconductor manufacturing [8]. Economic optimizers whose model is corrected by a bias alone have been analyzed as a class [9], and blend scheduling ties the recipe to the tanks that must supply it [10].

However, the refining literature treats blending as a problem of planning and optimization [10, 11], and the control literature treats ratio control, analyzer loops and bias updating as separate structures [9, 12], so neither offers a plant on which a tank can be watched falling off specification while every loop reads on target. This chapter studies that failure on an in-line gasoline blender that runs in a browser and in which the header and the tank are separate objects, the analyzer and the laboratory are delayed and fallible instruments, and the certificate is computed from the tank's composition by the same blending rules as everything else. The blending correlations, the linear program, the bias update and the PI trims are all established practice. Section 1.4.6 of Chapter 1 introduces the tank average and the bias update in general terms, and this chapter works them out on one plant.

Section 15.2 covers the history of knock ratings, blending indices and blend optimizers. Section 15.3 specifies the blender: a 39-state model with finite component inventories and volume-weighted property ledgers in the tank, an analyzer with transport, sampling, noise, drift, freezing and validation, and a bounded bias update, both introduced by this application and since shared by the framework with other apps. Section 15.4 shows why the specification does not add up, through the vapor-pressure index, ethanol and the cheapest recipe. Section 15.5 describes the controller layers and derives two results: a closed-form statement of what the stopping floor and the header clip of cumulative targeting cost the certificate, and a four-report delay model of the octane trim whose poles explain the tuning that oscillates. Section 15.6 reports the eight cases and the graded batch measured on the engine, each of which sets a header against its tank, an analyzer against the blend it reports, or a fast trim against a slow certificate. Section 15.7 sets the real analyzer, valves and control platform beside the model, Section 15.8 states the limitations and Section 15.9 draws the conclusions.

15.2 Knock Ratings, Blending Indices and Blend Optimizers

Octane numbers (ON) began as a way to rank knock, a story the History box below tells. A blend of rated components does not carry the volume average of their ratings, and correlations for blends transform or correct the component octanes before averaging them [2]. The application uses an interaction structure that it attributes to the Ethyl RT-70 method of 1959, the volume average plus a term in the covariance of octane and sensitivity and terms in the variances of olefin and aromatic content, with coefficients chosen for teaching.

Vapor pressure has the same difficulty and a simpler remedy. Butane, the cheapest stream in the pool, sets the vapor pressure of the blend, and refinery practice blends the Reid vapor pressure (RVP) through an index, $\mathrm{RVP}^{1.25}$, that is linear by volume [1], a form the application attributes to a 1959 trade article by Stewart and to Chevron. The summer ceiling on vapor pressure is an air-quality rule, and the effect of gasoline content regulation on ozone has been estimated from monitoring data [13]. Ethanol complicates both properties. Its contribution to octane depends on the base fuel and on the amount added [3, 14], a small addition raises the vapor pressure of a hydrocarbon base [4], and its octane has been argued to justify engines of higher efficiency [15].

> ⧖ HISTORY **1927: Two Hydrocarbons Make a Scale**
>
> In the 1920s knock was traced to the composition of the fuel and the compression ratio of the engine, and the Bureau of Standards asked a steering committee to work toward a standard reference fuel and a test method [16]. In 1927 Graham Edgar expressed the knock characteristics of a gasoline in terms of a standard fuel, a mixture of two pure hydrocarbons whose proportion defines the rating [17]. The research and motor tests that followed still rate a fuel against mixtures of normal heptane and isooctane [16]. United States rules define the octane rating posted at the pump as half the sum of the research and motor numbers, measured by ASTM D2699 and D2700 [18]. Regular grade is 87 on that scale and premium 92 [16]. The specification of this chapter asks for 87.0.

Linear programming came to blending early. The blending of aviation gasolines was posed as a linear program in 1952 [11], and a survey of 1957 described linear programming across the oil industry [19]. A recipe computed offline assumes component properties that are known only approximately, so blenders acquired online analyzers and model-based optimizers that re-solve the recipe as the analyses arrive [5]. The adequacy of a model corrected only by a bias for an economic optimizer was analyzed in 1994 [9], recipe optimization was later combined with the short-term scheduling of the blend [10], and feedback control of blend properties continues to be refined [6].

The certificate adds one more layer. A blender that meets the specification in the header at every instant certifies its tank, but one that recovers from an upset in the header has still put off-specification blend into the tank, and only the blend still to come can offset it. The simulator makes that layer explicit. It aims the rest of the batch at the tank's own target through the cumulative average, and it computes the certificate that no instrument measures during the batch from the volume of each component the tank holds.

15.3 The Blender as Modeled

The simulated blender fills an $8000\,\mathrm{m}^3$ product tank at $1000\,\mathrm{m}^3/\mathrm{h}$, an eight-hour batch, from six streams numbered $i = 1$ to 6: n-butane (C4), light straight-run naphtha (LSR), reformate (REF), FCC gasoline (FCC), alkylate (ALK) and denatured fuel ethanol (ETH). Each stream passes its own flow controller into one header, a detergent additive is dosed into the blend in parts per million, a near-infrared (NIR) analyzer samples the header $15\,\mathrm{m}^3$ downstream of the blend point, and a knock-engine laboratory samples it every half hour, as shown in Figure 15.1.

The 39 states in Table 15.1 are the six component flows and the additive pump's flow, the volume of each component and of the additive in the product tank, the six inventories, eighteen property ledgers and the money spent on components, listed beside the measured, manipulated and disturbing variables. Internally the model works in SI units on a clock in minutes, the display switches between field units (bbl, bbl/h, psi, \$/bbl) and SI, and the exported file has 161 columns. Each exported row pairs the true header and tank properties with the NIR and laboratory readings and adds the biases, the recipes, the header targets, the remaining fraction, the give-away and the certificate.

15.3.1 Component Data and the Specification

The component table and the specification are illustrative teaching data, as the application declares, and they are listed in Table 15.2. The specification is a teaching grade in the manner of the summer regular gasoline of ASTM D4814, as the application states, and not a certified grade. It asks for an AKI, $(\mathrm{RON}+\mathrm{MON})/2$, of at least 87.0, an RVP of at most 7.8 psi (13.5 psi in the winter setting), at most 25 vol % aromatics, 10 vol % olefins and 0.62 vol % benzene, at most 10 mg/kg sulfur, 10 vol % ethanol and 3.7 wt% oxygen, and at least 36 ppm of additive. Two entries were set with the lesson in mind. The LSR naphtha carries 0.5 vol % benzene, because at 1.5 the benzene limit caps it and octane never binds, and ethanol is priced at \$130 a barrel, because at \$70 it would be the cheapest octane in the pool and the optimizer would add it to every blend.

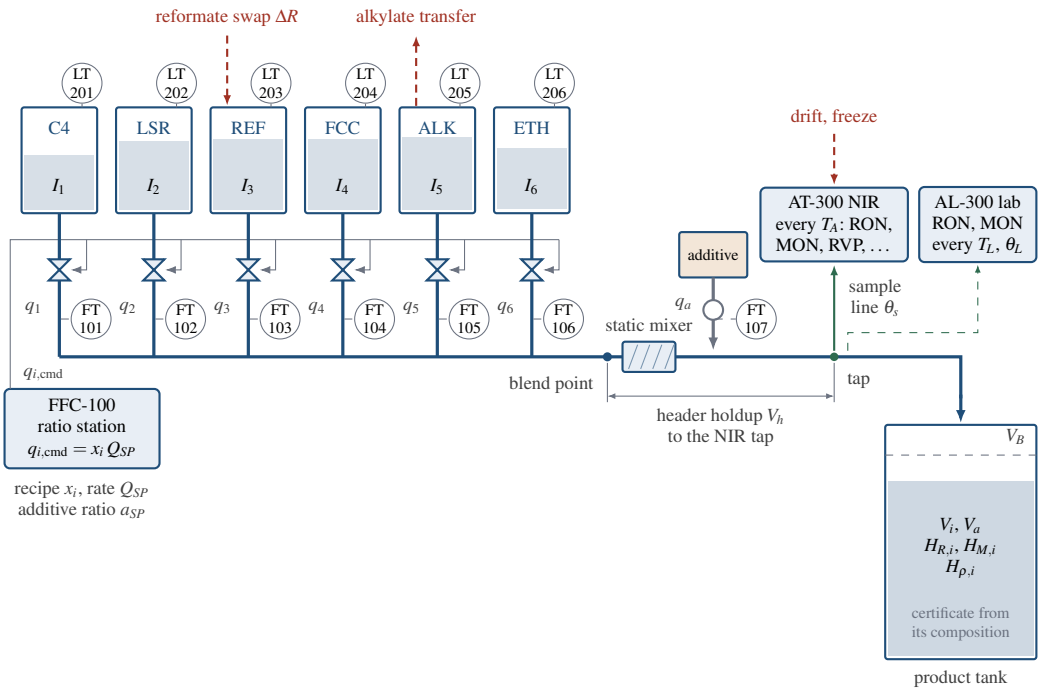


Figure 15.1 The in-line blender and its instruments. Six component tanks with inventories I_1 to I_6 and gauges LT-201 to LT-206 feed the header through valves commanded by the ratio station FFC-100, with each flow q_i metered by FT-101 to FT-106. The blend passes the static mixer, takes the additive q_a (FT-107), and travels the header holdup V_h to the tap of AT-300, which reports RON, MON, RVP, aromatics and benzene every T_A after the sample line's dead time θ_s. The laboratory AL-300 checks RON and MON every T_L with a turnaround of θ_L. The product tank holds each component's volume V_i with its volume-weighted RON, MON and density $H_{R,i}$, $H_{M,i}$ and $H_{\rho,i}$, up to the batch volume V_B. Thick dashed arrows (red) mark the disturbances.

15.3.2 Blending Rules

Every property of a blend is computed by one function from the volume fractions v_i of its components, and the header, the tank's certificate, the analyzer's truth, the optimizer and the operating map all call it. The compositional properties are averages by volume or by mass, as shown in Equation (15.1).

$$\begin{aligned} A &= \sum_i v_i A_i, \quad L = \sum_i v_i L_i, \quad Z = \sum_i v_i Z_i \\ \rho &= \sum_i v_i \rho_i, \quad S = \frac{1}{\rho}\sum_i v_i \rho_i S_i, \quad \omega = \frac{1}{\rho}\sum_i v_i \rho_i \omega_i \end{aligned} \tag{15.1}$$

where A, L and Z are the aromatics, olefins and benzene in vol %, the ethanol content is $E = 100\, v_6$, ρ is the density in kg/m^3, S the sulfur in mg/kg and ω the oxygen in wt%. Sulfur and oxygen are mass quantities, so they are weighted by volume times density. Vapor pressure is blended through its index, as shown in Equation (15.2).

$$\mathrm{RVP} = \left(\sum_i v_i R_i^{1.25}\right)^{0.8} \tag{15.2}$$

where R_i is the neat RVP of each hydrocarbon and $R_6 = 16.0$ psi is the blending RVP of ethanol, far above its neat 2.3 psi. The index is homogeneous of degree one in the pressures, so Equation (15.2)

Table 15.1 Classification of Variables, With Values at the Start of Case 2

Variable	Role	Start value	Range
AKI_N, RVP_N	CVs, AIC-301 and 302	87.50, 7.36 psi	—
$\overline{AKI}$, $\overline{RVP}$	CVs, FY-400	—	—
tank AKI, RVP	certificate	—	—
Q_m	CV, FFC-100	1000 m^3/h	0–1200
x_i	MVs, recipe (vol %)	2, 16, 24, 26, 32, 0	0–60[a]
Q_{SP}	MV, blend rate	1000 m^3/h	0–1200
Δ_O, Δ_B	MVs, trims (vol %)	0, 0	±6, ±3
a_{SP}	MV, additive ratio	40 ppm	0–120
AKI_h, RVP_h	SPs, from FY-400	87.20, 7.70 psi	see text
ΔR	DV, reformate RON	0	−5 to +2
$\Delta\rho$	DV, FCC density	0	±30 kg/m^3
$\dot{b}_N$	DV, NIR drift	0	±2 ON/h
—	DVs, freeze, transfer	off	1500 m^3
c_i	DVs, prices ($/bbl)	40, 75, 95, 85, 105, 130	10–200
q_i, q_a	states, flows	$x_i Q_{SP}$, 667 mL/min	—
V_i, V_a	states, tank	0	—
I_i	states (m^3)	400, 2500, 3000, 3000, 3000, 1500	0–20000
$H_{R,i}$, $H_{M,i}$, $H_{\rho,i}$	states, ledgers	0	—
C	state, cost	0	—

[a] butane 0–12.

Table 15.2 Component Data. Illustrative Teaching Values, With RVP in psi, Aromatics, Olefins and Benzene in vol %, Sulfur in mg/kg, Density in kg/m^3, Oxygen in wt % and Price in $/bbl

	RON	MON	RVP	Aro	Ole	Bz	S	ρ	O	Price
C4	93.8	89.6	52	0	0	0	0	584	0	40
LSR	68	66	11.5	3	1	0.5	5	670	0	75
REF	98	88	3.5	62	1	0.6	1	815	0	95
FCC	92	80	6.5	28	30	0.8	30	745	0	85
ALK	95	92	4.5	0.5	0.5	0	1	700	0	105
ETH	109	90	2.3[a]	0	0	0	0	794	33.7	130

[a] neat, blended at 16.0 psi.

gives the same result in psi or kPa, and the ethanol value is calibrated so that E10 lifts a 7.0 psi base by 1.00 psi.

Octane is computed first for the hydrocarbon base, with $w_i = v_i/(1-v_6)$ for $i = 1$ to 5, as shown in Equation (15.3).

$$\begin{aligned} \mathrm{RON}_b &= \bar{R} + a_1\left(\overline{Rs} - \bar{R}\,\bar{s}\right) + a_2\left(\overline{L^2} - \bar{L}^2\right) \\ &\quad + a_3\left(\overline{A^2} - \bar{A}^2\right) \\ \mathrm{MON}_b &= \bar{M} + b_1\left(\overline{Ms} - \bar{M}\,\bar{s}\right) + b_2\left(\overline{L^2} - \bar{L}^2\right) \\ &\quad + b_3\left(\overline{A^2} - \bar{A}^2\right) \end{aligned} \tag{15.3}$$

where each bar is a w-weighted average over the base, R_i and M_i here denote the components' RON and MON, $s_i = \mathrm{RON}_i - \mathrm{MON}_i$ is the sensitivity, and the coefficients are $a_1 = 0.0322$, $a_2 = 0.00101$, $a_3 = 0$, $b_1 = 0.0445$, $b_2 = 0.00081$ and $b_3 = -0.00050$. The three interaction terms are a covariance and two variances, so they vanish for a single component. Ethanol then blends into the base with a

blending value that falls as the base octane rises, as shown in Equation (15.4).

$$\begin{aligned} \mathrm{RON} &= (1-v_6)\,\mathrm{RON}_b + v_6\,\beta_R \\ \mathrm{MON} &= (1-v_6)\,\mathrm{MON}_b + v_6\,\beta_M \\ \beta_R &= 128 - 1.5\,(\mathrm{RON}_b - 90) \\ \beta_M &= 100 - (\mathrm{MON}_b - 82) \end{aligned} \tag{15.4}$$

where β_R and β_M are ethanol's blending RON and MON, and $\mathrm{AKI} = (\mathrm{RON}+\mathrm{MON})/2$. The denaturant brings the oxygen of the ethanol from 34.7 to 33.7 wt%. The shipped recipe, 2 vol % butane, 16 LSR, 24 reformate, 26 FCC gasoline and 32 alkylate, blends to RON 91.39, MON 83.61 and AKI 87.50, with an RVP of 7.36 psi, 22.8 vol % aromatics, 8.36 vol % olefins, 0.43 vol % benzene, 9.24 mg/kg sulfur and a density of 732 kg/m^3, at \$91.30 a barrel.

15.3.3 Flows, Inventories and the Tank

Each flow controller is a closed loop of first order around its command, and the tank accumulates volumes and property ledgers, as shown in Equation (15.5).

$$\begin{aligned} \tau_F \frac{\mathrm{d}q_i}{\mathrm{d}t} &= q_{i,\mathrm{cmd}} - q_i, & \frac{\mathrm{d}V_i}{\mathrm{d}t} &= -\frac{\mathrm{d}I_i}{\mathrm{d}t} = q_i \\ \frac{\mathrm{d}H_{R,i}}{\mathrm{d}t} &= q_i\,\mathrm{RON}_i^*, & \frac{\mathrm{d}H_{M,i}}{\mathrm{d}t} &= q_i\,\mathrm{MON}_i^* \\ \frac{\mathrm{d}H_{\rho,i}}{\mathrm{d}t} &= q_i\,\rho_i^*, & \frac{\mathrm{d}C}{\mathrm{d}t} &= \sum_i c_i\,q_i \\ \tau_a \frac{\mathrm{d}q_a}{\mathrm{d}t} &= q_{a,\mathrm{cmd}} - q_a, & \frac{\mathrm{d}V_a}{\mathrm{d}t} &= q_a \end{aligned} \tag{15.5}$$

where flows are in m^3 per minute of the model's clock, $\tau_F = 5\,\mathrm{s}$, $\tau_a = 2\,\mathrm{s}$, and the starred properties are the true ones. A reformate tank swap moves the true reformate to $\mathrm{RON}_3^* = 98 + \Delta R$ and $\mathrm{MON}_3^* = 88 + 0.833\,\Delta R$, and a feed change moves the FCC density to $\rho_4^* = 745 + \Delta\rho$, while the optimizer keeps the declared table. Because the ledgers carry the octane each component had when it went in, the reformate already in the tank keeps its octane after a swap. The certificate evaluates Equations (15.1) to (15.4) on the tank's composition V_i with each component at $\mathrm{RON}_i^T = H_{R,i}/V_i$, $\mathrm{MON}_i^T = H_{M,i}/V_i$ and $\rho_i^T = H_{\rho,i}/V_i$, so the tank's average is exact under nonlinear blending and no analyzer reading enters it. Two events are applied after each step. A component tank that runs dry stops its flow at once, the overdraw of the last step is returned so that the balance still closes, and COMPONENT SHORT is raised. The batch completes when $\sum_i V_i$ reaches V_B.

The header holdup is represented only as the analyzer's transport delay. The tank receives the blend at the blend point, since everything in the header reaches the tank anyway, which keeps the balance $V_i(t) - V_i(0) = I_i(0) - I_i(t)$ exact. Give-away, the quality paid for and not needed, is scored on the tank as shown in Equation (15.6).

$$G = V\left[(\mathrm{AKI} - 87.0)^+\,v_{oct} + (\mathrm{RVP}_{\max} - \mathrm{RVP})^+\,v_{rvp}\right] \tag{15.6}$$

where the tank volume is V, $(\cdot)^+ = \max(0,\cdot)$, and v_{oct} and v_{rvp} are \$1.00 per octane number and \$1.00 per psi for each barrel.

15.3.4 Instruments

The flow meters FT-101 to FT-106 lag their flows by 1 s and carry 0.2 % white noise, and the tank gauges LT-201 to LT-206 lag by 10 s with a resolution of 0.5 m^3. The NIR tap sees the header composition V_h behind the blend point, a transport delay keyed by volume, so it lengthens as the

Table 15.3 Parameters and Where They Can Be Changed. Visible Ones Appear as Sliders in the Model Dialog, Hidden Ones Do Not, and Fixed Ones Are Written Into the Source

Symbol	Meaning	Value	Class
V_B	batch volume	8000 m^3	visible
—	already in the tank	0, case 4 6400 m^3	visible
$I_i(0)$	starting inventories	Table 15.1	visible
$q_{i,\max}$	valve capacities (m^3/h)	60, 400, 450, 450, 450, 150	hidden
R	valve rangeability	30:1	visible
τ_F	flow-loop time constant	5 s	visible
—	meter noise, lag	0.2 %, 1 s	hidden, fixed
—	gauge lag, resolution	10 s, 0.5 m^3	fixed
V_h	header holdup to the tap	15 m^3	visible
θ_s	sample-line dead time	1.1 min	visible
T_A	NIR cycle and analysis	1.0 min	visible
—	NIR noise, octane, RVP	0.10 ON, 0.35 kPa	hidden
—	bad analyses	2 % of reports	hidden
T_L, θ_L	laboratory interval, turnaround	30 min, 20 min	visible
—	laboratory repeatability	0.15 ON	hidden
$q_{a,\max}$	additive pump capacity	1500 mL/min	visible
N_a	additive pump turndown	10:1	visible
τ_a	additive pump time constant	2 s	hidden
—	additive required	36 ppm	visible
v_{oct}, v_{rvp}	give-away, per bbl	\$1.00/ON, \$1.00/psi	visible
—	NIR gate, octane, RVP, aro, Bz	1.0, 3.0 kPa, 2.0, 0.1	fixed
$b_{\max}$	bias bounds, same order	4, 14 kPa, 4, 0.3	fixed
—	laboratory correction gain	0.7	fixed
—	inventory reserve	20 m^3	fixed
—	targeting clip, AKI, RVP	1.5, 1.0 psi	fixed

blend rate falls. The analyzer draws a sample every T_A, the fluid now at its inlet after the sample line's dead time θ_s, and publishes the result one cycle later, as shown in Equation (15.7).

$$y_k = P\big(x(t_k - V_h/Q - \theta_s - T_A)\big) + b_N(t_k - T_A) + \varepsilon_k \tag{15.7}$$

where P is the property from Equations (15.1) to (15.4) on the true table, x the header composition and Q the blend rate. The noise ε_k has standard deviations of 0.10 ON on RON and MON, 0.35 kPa (0.05 psi) on RVP, 0.15 vol % on aromatics and 0.01 vol % on benzene, a bad analysis in 2 % of reports moves every channel the same way by 2.0 ON, 6.0 kPa, 3.0 vol % and 0.15 vol %, and the instrument's own error b_N integrates a drift in RON and MON. At 1000 m^3/h the header adds 0.9 min, so each report describes blend made 3.0 min earlier. A frozen instrument keeps publishing its last reading, and validation flags the analyzer STALE after three identical reports, or after three cycles without a fresh one. The knock-engine laboratory draws a header sample every $T_L = 30$ min and reports its RON and MON $\theta_L = 20$ min later, with a repeatability of 0.15 ON.

15.3.5 Initial States, Clock and Integration

Table 15.3 lists the parameters. Each case starts with the flows on the commands of its recipe, the additive pump on its ratio, the inventories at their starting values and the tank holding whatever the case places in it. Case 4 starts with 6400 m^3 of the optimizer's recipe blended from a reformate 1.108 RON weaker, an AKI of 86.90. The plant advances by fourth-order Runge-Kutta integration with a fixed step of 0.5 s, a tenth of the flow loops' time constant, the controllers execute every second, and the historian logs a row every 5 s, twelve per analyzer cycle and 5760 for an eight-hour batch. Every random sequence is seeded at 20261102, so a run replays bit for bit (Section 1.6.1), and the speed selector spans 1 to 300 simulated minutes per second of wall time, 5 by default, without changing the trajectory, so a full batch takes under two seconds at the fastest speed. The application passes 44 checks. Among them, halving the step moves the true header RVP through the butane step of case 1

and the tank AKI of case 2 at 240 min by less than 0.001, the tank holds the integral of its component flows to 10^{-9} through a transfer and a tank run dry, and the certificate is unchanged when a frozen, biased analyzer reading is altered.

15.3.6 Assumptions

The component table, the coefficients of Equations (15.2) to (15.4) and the specification are teaching values, not the properties of any refinery's streams, a fitted correlation or a certified grade. One blending RVP of 16 psi serves ethanol at every concentration, while real ethanol lifts the vapor pressure most near 5 to 10 vol % and less above, and the ethanol octane terms are meant for less than about 15 vol %.

Volumes are additive, so the shrinkage of a few tenths of a percent that butane causes on blending is ignored, and sulfur and oxygen are mass averages through fixed densities. The product tank is well mixed and receives the blend at the blend point, which is exact for the certificate at the end of a batch and shifts the analyzed record of the tank by the 15 m^3 in the header during it.

Each flow controller is a first-order closed loop with a capacity and a rangeability, without valve dynamics, stiction or meter calibration error. The NIR is a sample-and-hold instrument with white noise, occasional bad analyses, a drift that is a ramp and a freeze, and it reads the header by the same blending rules as everything else, so it has no chemometric error of its own. The laboratory is an unbiased instrument with a repeatability. All sizes, prices and limits are teaching values for one grade, not a design basis.

15.4 Why the Specification Does Not Add Up

Vapor pressure is the property that the recipe sets most sharply. In index form butane counts for 139.6 $\text{psi}^{1.25}$ against 21.2 for LSR, 10.4 for FCC gasoline, 6.6 for alkylate, 4.8 for reformate and 32.0 for ethanol at its blending value, while the shipped blend's index is 12.1. Differentiating Equation (15.2) with butane replacing the rest of the recipe in its own proportions gives the slope of Equation (15.8).

$$\frac{\mathrm{dRVP}}{\mathrm{d}x_1} = 0.8\,\mathrm{RVP}\,\frac{R_1^{1.25} - I_r}{I} \tag{15.8}$$

where I is the blend's index and $I_r = 9.52\,\text{psi}^{1.25}$ is the index of the non-butane part. At the shipped recipe the slope is 0.632 psi per 1 vol % of butane, against 0.461 psi for a volume average of the pressures, so the index rule weighs butane 1.37 times as heavily. Three more percent of butane lift the RVP from 7.36 to 9.20 psi, a rise of 1.84 psi where the volume average predicts 1.38, as shown in Figure 15.2. A blender who sized butane by the volume average would put 4.06 vol % in the blend and ship 8.64 psi against the 7.8 psi limit, while the index rule allows 2.70 vol %. Octane barely moves with butane, 0.037 AKI per percent, because butane is itself of high octane, so butane is run against the vapor-pressure ceiling and nothing else.

> **TRY IT IN THE APP** **1 · Manual recipe: butane and RVP**
>
> Load the case and press Run. Nothing else has to be set. At 5 min the butane in the recipe steps from 2 to 5 vol %. Watch the header RVP climb from 7.36 to 9.20 psi while the AKI barely moves. The analyzer first shows the step at 9.0 min. By then the tank has already passed its 7.8 psi limit, at 6.6 min.

Ethanol is the opposite case, a stream whose own properties mislead in both directions. In the shipped base Equation (15.4) gives ethanol a blending RON of 125.9 and MON of 98.4, far above its neat 109 and 90, so E10 raises the AKI from 87.50 to 89.96. Blended at its neat 2.3 psi, E10 would lower the RVP to 6.91 psi. At its blending value of 16 psi it raises the RVP to 8.31 psi, over the summer limit, and it brings 3.62 wt% oxygen against the 3.7 wt% limit. At $130 a barrel the E10 recipe costs $95.17 against $91.30, and room under the vapor-pressure ceiling for the ethanol has to come from butane and LSR naphtha.

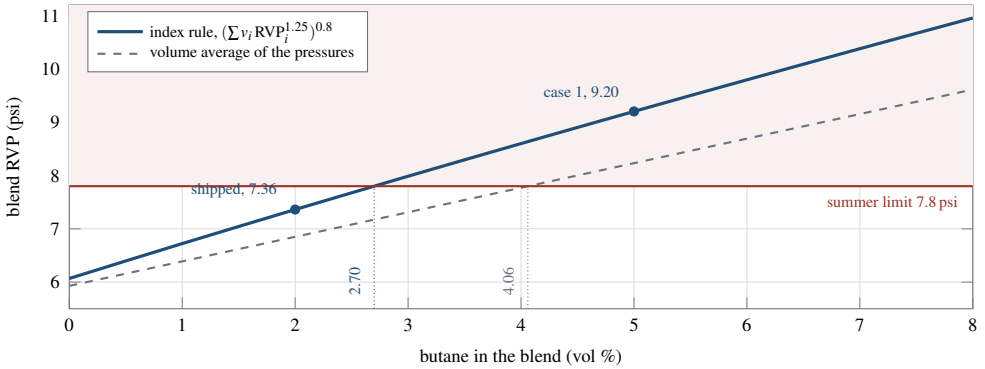


Figure 15.2 RVP of the shipped recipe against its butane fraction, with the other five streams in their shipped proportions. The index rule of Equation (15.2) (thick solid line) crosses the 7.8 psi limit at 2.70 vol % and the volume average of the pressures (dashed) at 4.06 vol %, and the two marked points are the shipped recipe and case 1 after its step.

The cheapest recipe is pressed against several of these limits at once. With every specification but AKI and RVP held a certification margin inside its limit (0.3 vol % aromatics and olefins, 0.02 vol % benzene, 0.3 mg/kg sulfur and 0.05 wt% oxygen), the linear program of Section 15.5, solved once from the shipped recipe at the base prices and targets of 87.20 and 7.70 psi, puts 2.40 vol % butane, 16.99 LSR, 26.36 reformate, 27.55 FCC gasoline and 26.70 alkylate in the blend. That recipe blends to AKI 87.21, 7.70 psi, 24.70 vol % aromatics and 9.70 mg/kg sulfur and costs \$90.19 a barrel against \$91.30 for the shipped recipe. Re-linearized at its own solution the program moves each fraction by less than 0.03 vol % onto AKI 87.20 exactly, and perturbing its targets there gives marginal values of \$1.004 per octane number and \$0.997 per psi for each barrel, the values that Equation (15.6) rounds to \$1.00. The feasible region around that vertex is a small pocket, which is why a recipe held by hand keeps a margin everywhere.

15.5 Implemented Controller

The control system reads only its instruments, the flow meters, the tank gauges, the NIR, the laboratory, the pumps' low-suction trips and the product tank's high level, and never the true header or the tank. It executes every $\Delta t_c = 1$ s, every second physics step, in the order of Section 1.6.1, and a row is logged every 5 s. A selector on the FY-100 card chooses where the recipe comes from, the recipe sliders with the trims' manual stations (fixed), the PI trims, or the optimizer, and a switch turns cumulative targeting on or off, as shown in Figure 15.3. In global manual the operator holds the recipe and both trim stations, every PI tracks, and the optimizer is still solved so that the map can show where it would go. In automatic the recipe sliders follow the optimizer's recipe and the trim stations follow the trims after every execution, so a switch to manual is bumpless.

15.5.1 Ratio Station, Flow Limits and Additive

FFC-100 multiplies the master rate by the recipe, and each flow command is limited by its valve, as shown in Equation (15.9).

$$q_{i,\mathrm{cmd}} = \mathrm{clamp}\big(x_i Q_{SP},\; q_{i,\max}/R,\; q_{i,\max}\big) \tag{15.9}$$

with $q_{i,\mathrm{cmd}} = 0$ when $x_i Q_{SP} < q_{i,\max}/(2R)$ or tank i is empty, so a valve controls no lower than a thirtieth of its capacity and shuts below half of that. Moving one recipe slider scales the other five to keep 100 percent. The additive pump is set in ratio to the metered header flow, $q_{a,\mathrm{SP}} = 10^{-6}\, a_{SP}\, Q_m$, and clamped between $q_{a,\max}/N_a$ and $q_{a,\max}$. At 40 ppm and 1000 m³/h it runs at 667 mL/min, and below 225 m³/h it sits on its 150 mL/min turndown, so at 150 m³/h it doses 60 ppm.

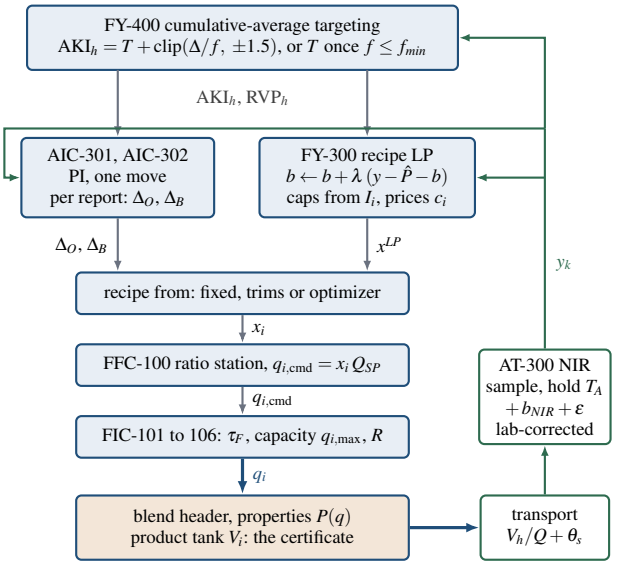


Figure 15.3 Control layers of the blender. Cumulative targeting sets the header targets, the PI trims or the recipe optimizer turn them into a recipe, the ratio station and the six flow loops deliver it, and the NIR reading y_k, delayed by the header, the sample line and the analysis cycle, returns to the trims, to the optimizer's bias update and to the analyzed tank average.

15.5.2 Property Trims

The two trims act on the recipe rather than on a valve. AIC-301 swaps reformate for LSR naphtha and AIC-302 adds butane with everything else scaled to make room, as shown in Equation (15.10).

$$\begin{aligned} d &= \text{clamp}(\Delta_O/100, -x_3, x_2) \\ x_3' &= x_3 + d, \qquad x_2' = x_2 - d \\ x_1'' &= \text{clamp}(x_1 + \Delta_B/100, 0, 0.5) \\ x_i'' &= x_i' \frac{1 - x_1''}{1 - x_1}, \qquad i \neq 1 \end{aligned} \tag{15.10}$$

Both are PI controllers in the framework's velocity form (Section 1.6.2), executed once for each fresh, valid NIR report while blend is flowing, as shown in Equation (15.11).

$$\begin{aligned} e_k &= \text{SP}_k - \text{PV}_k, \qquad P_k = K_c\,e_k \\ S_k &= S_{k-1} + P_k - P_{k-1}, \qquad v_k = S_k + I_{k-1} \\ u_k &= \text{clamp}(v_k, u_{min}, u_{max}) \\ I_k &= I_{k-1} + \left[\frac{K_c}{\tau_I}\,e_k + \frac{u_k - v_k}{\tau_I}\right]\Delta t_k \end{aligned} \tag{15.11}$$

where Δt_k, the time since the previous execution, is normally 1 min. These are Equations (1.16) to (1.19) of Chapter 1 with no derivative term, no output bias and $T_t = \tau_I$. Windup is handled by back-calculation on the integral alone, the clamp limits only what is emitted, and neither output has a rate limit. In manual, with a fixed recipe or under the optimizer, each trim tracks by setting $I_k = u_{op} - S_k$ for the value it is held at, so the return to the trims is bumpless. AIC-301 has SP = AKI$_h$, PV = AKI$_N$, output limits of ± 6 vol %, $K_c = 1.5$ vol % per ON and $\tau_I = 2$ min. AIC-302 has SP = RVP$_h$, PV = RVP$_N$, limits of ± 3 vol %, $K_c = 0.06$ vol % per kPa (0.41 per psi) and $\tau_I = 1.5$ min. Under the ISA convention both are reverse acting, since more reformate raises the octane and more

butane raises the vapor pressure, and positive gains with $e = \mathrm{SP} - \mathrm{PV}$ deliver that action. Volume 3, Chapter 7 trims a ratio blend of syrup into water on a refractometer in the same way. The split of Equation (15.10) adds one limit that the PI law does not know, the butane floor at zero, which would let the RVP trim's integral wind below -2 vol % on the shipped 2 vol % of butane. It never binds in the cases measured here, where the butane trim never falls below –0.94 vol %.

15.5.3 Validation, Laboratory Correction and the Bias Update

Each fresh report is first compared with the model's prediction for the very sample it analyzed. The controller keeps its own history of metered compositions keyed by metered volume, finds the metered volume V_m at the moment the sample left the tap, and predicts $\hat{P}_k = P(\hat{x}(V_m - V_h))$ on the declared table. A channel further than its gate width from the prediction plus its bias is rejected, unless three reports in a row are rejected on the same side, which is taken as a real change. With the laboratory correction on, each knock-engine result moves an offset c on RON and MON 70 percent of the way to the NIR's error on the laboratory's own sample, $c \leftarrow c + 0.7\,(y_N - y_L - c)$, and the controllers read $\tilde{y}_k = y_k - c$. The correction is off in the shipped cases. Chapter 9 corrects an analyzer and an inferred octane from its own laboratory engine. The bias of each measured property then follows Equation (15.12), the update of Equation (1.14) with a bound.

$$b \leftarrow \mathrm{clamp}\big(b + \lambda\,(\tilde{y}_k - \hat{P}_k - b),\; -b_{\max},\; b_{\max}\big) \tag{15.12}$$

with $\lambda = 0.3$ for RON, MON, RVP, aromatics and benzene, skipped when the reading is stale, rejected or taken while nothing flows. For a constant model error δ the bias after k updates is $\delta\,[1 - (1-\lambda)^k]$, 97 percent of the error in ten reports. The bias belongs to the model, and it cannot tell a drifting analyzer from a wrong model. The laboratory, an independent method, is the only element that can. An estimator that carries a bias as a state is the offset-free predictive controller of Volume 9, Chapter 4.

Pause. Try Exercise 15.2 before reading on.

15.5.4 Cumulative-Average Targeting

FY-400 keeps the tank's analyzed average. Each report stands for the metered volume that passed the tap since the previous one and is booked at its corrected AKI and RVP index, or at the model's prediction plus the bias when the reading cannot be trusted. With $\bar{V}$ the analyzed volume and $\overline{\mathrm{AKI}}$ its average, the targets follow Equation (15.13).

$$\begin{aligned} f &= \frac{V_B - \bar{V}}{V_B}, \quad \Delta = (T - \overline{\mathrm{AKI}})(1 - f), \quad D = \frac{\Delta}{f} \\ \mathrm{AKI}_h &= \begin{cases} T + \mathrm{clamp}(D, -1.5, 1.5), & f > f_{min} \\ T, & f \le f_{min} \end{cases} \end{aligned} \tag{15.13}$$

where T is the AKI target, 87.20 by default, Δ is the tank's deficit spread over the whole batch and $f_{min} = 0.05$. The header target $T + D$ is Equation (1.15) of Chapter 1 with $\overline{\mathrm{AKI}}$ as the average so far and f as the share still to blend. The RVP target is set the same way in index form, with a clip equivalent to 1.0 psi. The few tens of cubic meters in the header and the sample line that have not yet been analyzed count as remaining.

Pause. Try Exercise 15.1 before reading on.

The law has a closed-form cost. Blending a small share $\mathrm{d}\phi$ of the batch at a header AKI H changes the deficit by $\mathrm{d}\Delta = (T - H)\,\mathrm{d}\phi$. While the header delivers $T + D$ exactly, Δ falls in proportion to f, so D stays at its first value $D_0 = \Delta_0 / f_0$. While the clip holds the header at $T + c$, the deficit falls by c for each unit of batch blended. Either way the rest of the batch after f_{min} is blended at T, and the tank

finishes at the value of Equation (15.14).

$$\mathrm{AKI}_{end} = \begin{cases} T - f_{min}\,\Delta_0 / f_0, & D_0 \le c \\ T - \Delta_0 + c\,(f_0 - f_{min}), & D_0 > c \end{cases} \tag{15.14}$$

The floor therefore costs a shortfall of $f_{min} D_0$ even when the header meets every target, and without targeting the tank finishes at $T - \Delta_0$.

> **TRY IT IN THE APP** **4 · A late upset with 20 % left**
>
> Load the case and press Run. The tank starts 80 percent full at AKI 86.90, one tenth below its limit. Watch the header target rise to 88.40 and hold there while the tank average creeps up. At 75.0 min the targeting stops and the header returns to 87.20. The batch completes at 96.0 min and certifies at 87.15, short of its 87.20 target.

15.5.5 Recipe Optimizer

FY-300 solves a linear program on every NIR report around the recipe x^0 it last applied, as shown in Equation (15.15).

$$\begin{aligned} \min_{x,s\ge 0} \; & \sum_i c_i x_i + \sum_k w_k s_k \quad \text{subject to} \\ & \sum_i \bar{a}_i x_i + \tfrac{1}{2}(b_{\mathrm{RON}} + b_{\mathrm{MON}}) + s_1 \ge \mathrm{AKI}_h \\ & \sum_i R_i^{1.25} x_i - s_2 \le (\mathrm{RVP}_h - b_{\mathrm{RVP}})^{1.25} \\ & \sum_i A_i x_i - s_3 \le A'_{\max} - b_A, \qquad \sum_i L_i x_i - s_4 \le L'_{\max} \\ & \sum_i Z_i x_i - s_5 \le Z'_{\max} - b_Z \\ & \sum_i \rho_i (S_i - S'_{\max}) x_i / \rho_r - s_6 \le 0 \\ & \sum_i \rho_i (\omega_i - \omega'_{\max}) x_i / \rho_r - s_7 \le 0 \\ & \sum_i x_i = 1, \qquad \ell_i \le x_i \le h_i \end{aligned} \tag{15.15}$$

where the primed limits carry the certification margins of Section 15.4, $\rho_r = 740\,\mathrm{kg/m^3}$, and $\bar{a}_i$ are the blending AKIs of the components, the forward-difference gradient g_i of Equations (15.3) and (15.4) at x^0 shifted so that $\bar{a}_i = \mathrm{AKI}(x^0) + g_i - \sum_j x^0_j g_j$ and the linear model is exact at the recipe it starts from. The RVP row is exact, since the index is linear in volume. The elastic slacks s_k carry penalties of \$20 000 per cubic meter of blend for each unit of violation (\$200 000 for benzene and \$5882 per index unit for RVP), so they are zero whenever the specification can be met and report by how much it cannot. The bounds are $h_i = \min(\kappa_i, x^0_i + \Delta x_{\max})$ and $\ell_i = \max(0, x^0_i - \Delta x_{\max})$ with a step limit $\Delta x_{\max}$ of 2 vol % per analysis, and the caps are $\kappa_i = \min\left[1,\; (I_{i,m} - 20)/V_{rem},\; q_{i,\max}/Q_{SP}\right]$, so each component's gauged inventory less a $20\,\mathrm{m^3}$ reserve is spread over the volume V_{rem} still to blend (at least $20\,\mathrm{m^3}$), with ethanol further capped at 10 vol %. A dense two-phase simplex with Bland's rule solves the problem twice, once with the step limit for the recipe it applies and once without it for the map, the cost trend and the FY-300 card.

> ⧖ HISTORY **1980s: Blend Control Goes Online**
>
> Blend controllers that close the loop on online analyzers drew research effort for some forty years before 2010 [6]. Commercial blending control packages were available by the late 1970s, but none had a satisfactory answer to component properties that were poorly known. One refiner dates its first online applications to the early 1980s. In the late 1990s it took up the uncertainty problem, and the version in service since 2001 joins an optimization problem to a feedback loop with an observer of the component properties. By 2010 it ran on 17 blenders in six refineries [6]. The optimizer of this section answers the same uncertainty with the bias update of Equation (15.12).

15.5.6 What the Trims See and What the Controllers Assume

Section 15.3 describes the simulation model, and the application fits no control model of the plant. Manual steps of the trim stations in case 2, each made from the shipped recipe at 10 min, before the reformate swap, or at 40 min, after it, measure what the trims see. One percent of reformate for LSR raises the true header by 0.226 AKI within seconds (0.198 after the reformate swap), and half a percent of butane raises its RVP by 0.314 psi. A move made at one report first appears in the fourth report after it, because 3.0 min of transport and analysis and the phase of the sampling separate the blend point from the result. Volume 5, Chapter 3 meets the same limit on a column composition loop. The trim therefore sees $y_k = K u_{k-4}$. Equation (15.11) emits u_k with the integral carried in from the previous report, so at $\Delta t_k = 1$ min the closed loop has the characteristic polynomial of Equation (15.16).

$$z^5 - z^4 + K K_c z - K K_c (1 - \Delta t_k / \tau_I) = 0 \tag{15.16}$$

For $K_c = 4$ and $\tau_I = 1$ min its dominant roots lie at $|z| = 1.13$ to 1.16 for $K = 0.198$ to 0.226, a growing oscillation with a period of 12 reports. For the shipped $K_c = 1.5$ and $\tau_I = 2$ min the slowest root is real, at 0.72 to 0.77, and at that integral time the loop turns unstable above $K_c = 3.9$ to 4.5. The shipped RVP trim, at $K = 4.33$ kPa per percent of butane, has its slowest root at 0.76. The controllers' internal models are the linearized blend model with its biases in Equation (15.15), the aligned prediction $\hat{P}_k$ on the declared table, the volume average in Equation (15.13), and the recipe split of Equation (15.10). None of them knows the true reformate, the analyzer's drift or the composition of the tank.

🕮 **Pause.** Try Exercise 15.3 before reading on.

15.6 Scenarios and Measured Results

The cases of Table 15.4 share one reference configuration, that of Table 15.1 with the outlier gate on, the laboratory correction off, $\lambda = 0.3$, a step limit of 2 vol % and a floor of $f_{min} = 0.05$, and each departs from it only where its row says so. Every number in this section was measured headlessly on the engine at its default seed, with each configuration set on the case as loaded before the run starts. The application pauses a run when the batch completes, and Table 15.5 gathers the scorecard of each batch at that pause. Times are simulated minutes counted from the start of each case. Reloading a case after a run, or pressing Reset with the same settings, reproduces these numbers exactly, and the spread over six noise seeds is given beside the results that the noise moves most. Figure 15.4 draws the certificate AKI and the give-away of Table 15.5 as a chart.

15.6.1 Butane, Ethanol and the Analyzer's Delay

In case 1 the butane goes from 2 to 5 vol % at 5 min, in manual. The butane flow reaches 90 % of its step 11.5 s after the recipe changes, as its 5 s loop requires, and the true header RVP rises from 7.36 to 9.20 psi while the AKI moves only from 87.50 to 87.61. The first NIR report that shows the step arrives at 9.0 min. Its sample left the tap at 6.9 min, carrying blend made at 6.0 min, so the report is 4.0 min after the step and 3.0 min after the blend it describes. The tank passes the 7.8 psi limit at 6.6 min, before the analyzer has reported anything, and holds 8.74 psi at 20 min. Case 5 adds 10 vol % of ethanol at 5 min, in manual, and the header goes from AKI 87.50 to 89.96 (RON 91.39 to 94.84, MON 83.61 to 85.09) and from 7.36 to 8.31 psi, the numbers of Section 15.4.

Table 15.4 Cases, Their Events and What Each Teaches

#	Scenario and event	Teaching point
1	manual, butane 2 to 5 vol % at 5 min	index rule, analyzer delay
2	fixed, reformate −3 RON at 30 min	a recipe is its data
3	trims at $K_c = 4$, $\tau_I = 1$, swap at 10 min	PI on a dead time
4	optimizer, tank 6400 m^3 at 86.90	Δ/f and the floor
5	manual, ethanol 0 to 10 vol % at 5 min	octane and RVP together
6	butane \$25 at 5 min, optimizer at 10	an LP sits on its limits
7	optimizer, NIR drift +0.2 ON/h at 10	bias follows the analyzer
8	optimizer, 1500 m^3 alkylate out at 60	plan the inventory
9	fixed, swap at 120 min, 1600 m^3 alkylate	graded batch

Table 15.5 Scorecard of Full Batches: Certificate AKI and RVP, Result, Give-Away and Blend Cost

Case and configuration	AKI	RVP	Certificate	G (\$)	\$/bbl
2 fixed	86.79	7.36	off, AKI	22 062	91.30
2 trims	87.19	7.70	certified	14 769	91.27
2 optimizer	87.19	7.70	certified	14 396	90.95
4 targeting on	87.15	7.70	certified	12 610	91.14
4 targeting off	86.96	7.70	off, AKI	5 131	90.18
4 trims, targeting on	87.15	7.70	off, aromatics	12 473	90.72
6 full batch	87.20	7.70	certified	15 493	89.85
7 correction off	86.45	7.70	off, AKI	4 854	89.43
7 correction on	87.06	7.70	certified	7 835	90.04
8 optimizer	87.18	7.71	certified	14 001	90.86
8 fixed[a]	86.67	8.02	off, four	0	88.60
9 as shipped[a]	85.97	7.84	off, four	0	88.88
9 trims[a]	86.69	7.70	off, three	4 793	89.50
9 optimizer	87.19	7.71	certified	14 041	91.56
9 suggested	87.19	7.70	certified	14 511	91.57

[a] alkylate ran dry, and the batch completed after 480 min, at 532, 565 and 564 min. RVP in psi.

15.6.2 A Recipe Is Only as Good as Its Data

In case 2 the blender runs the shipped recipe into an empty tank, and at 30 min the reformate is drawn from a tank 3 RON and 2.5 MON weaker. The header falls from AKI 87.50 to 86.74 within a minute, and the NIR first reads below 87.0 at 33.0 min. Nothing acts on it. The tank average crosses 87.0 at 87.5 min, and the batch completes at 480 min at AKI 86.79, off specification, with about 309 m^3 of alkylate needed to correct it. With the PI trims, selected before the run, the octane trim ends at 2.21 vol % of reformate for LSR and the butane trim at 0.82 vol %, and the tank certifies at 87.19 with \$14 769 of give-away (\$14 769 to \$15 133 over six noise seeds). The optimizer certifies at the same AKI for \$90.95 a barrel against \$91.27 for the trims, because it re-optimizes the whole recipe while the trims move only reformate against LSR and butane against the rest.

15.6.3 An Octane Trim on a Dead Time

Case 3 ships the octane trim tuned like a flow loop, 4 vol % per ON with a 1 min integral time, and swaps the reformate at 10 min. Three corrections go out before the first is seen, the true header swings between AKI 85.59 and 87.93 in the hour, crossing 87.20 downward every 11.3 min (11.0 min to 11.6 over six noise seeds), and the trim sits on its ±6 vol % limits 38 % of the time from 12 min to 60. The growth that Equation (15.16) predicts for this tuning is held by those limits in a limit cycle, and its period of 11.3 reports lies near the 12 of the linear model. With the shipped 1.5 vol % per

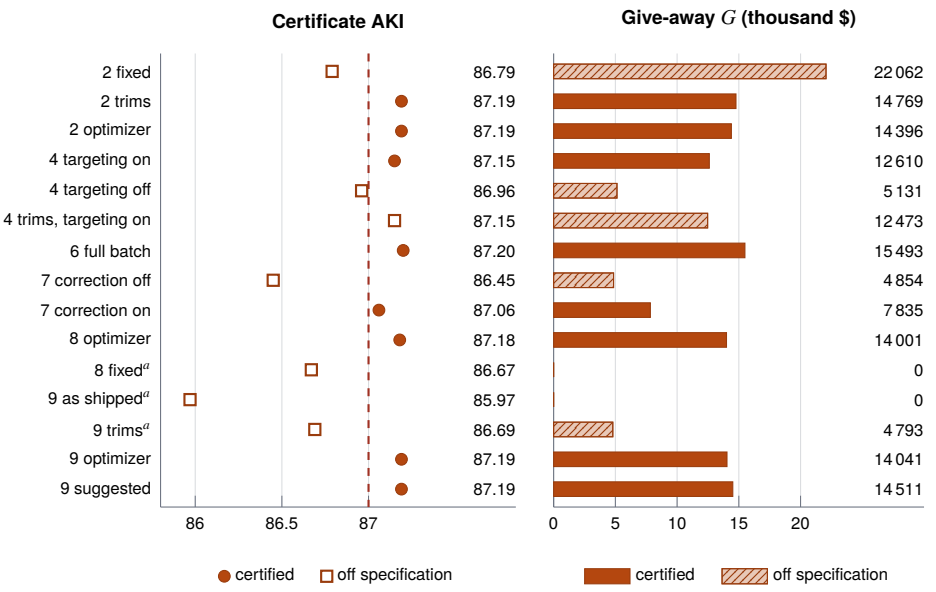


Figure 15.4 The certificate AKI and the give-away of Table 15.5 as a chart. The dashed line is the 87.0 limit. Filled circles and solid bars are certified batches, open squares and hatched bars are batches off specification. Notice that every certified batch lands between the limit and the 87.20 target and gives away $7 835 or more, that the batches with the least give-away are all off specification, and that one batch, the trims of case 4 with targeting on, clears the octane limit and still fails, on aromatics. Rows marked *a* ran the alkylate dry.

ON and 2 min set before the run, the header dips to 86.47 and is back within 0.1 AKI of its setpoint 9.0 min after the swap (6.1 to 9.3 min over the seeds). At 0.5 and 2 min it dips only to 86.62 but takes 33.1 min (28.3 to 33.1), the slow root of 0.94 to 0.95. Holding 2 min and raising the gain, the header's noise-driven swing over each half hour from 60 to 240 min grows from 0.18 to 0.26 AKI peak to peak at $K_c = 2.5$ to between 0.30 and 0.72 at 4.5, where Equation (15.16) puts the root for the weaker reformate at $|z| = 1.00$. On a loop that is almost pure delay the integral action does the work, and the gain mainly decides how hard the first correction is.

15.6.4 A Late Deficit and Its Floor

Case 4 starts with 6400 m^3 in the tank at AKI 86.90 and blends the last 1600 m^3 with the optimizer and cumulative targeting on. The deficit is $\Delta_0 = 0.30 \times 0.8 = 0.24$ at $f_0 = 0.2$, so $D_0 = 1.20$ and the header target rises to 88.40, then to between 88.40 and 88.44 as blend made before the optimizer reached the new target enters the analyzed average. The optimizer meets it within 3 min, since it may move each fraction 2 vol % per analysis. While its steps are limited it borrows up to 2.0 vol % of ethanol, gone again by 4 min, and from 10 to 70 min its recipe averages 31.3 vol % of alkylate instead of 26.7 and 3.1 of butane instead of 2.4, with LSR cut from 17.0 to 11.0. The demand D stays between 1.21 and 1.24 from 5 to 75 min, as the analysis of Equation (15.14) requires. The targeting stops at 75.0 min, when the analyzed share left falls below 5 percent, and the header returns to 87.20. The batch completes at 96.0 min and certifies at AKI 87.15, on specification but short of target, with $12 610 of give-away, as shown in Figure 15.5. Without targeting, switched off before the run, the last fifth is blended at 87.20 and the tank certifies off specification at 86.96, exactly $T - \Delta_0$, and needs about 63 m^3 of alkylate.

Equation (15.14) predicts 87.14 for this case, and the engine lands 0.011 higher because the targeting stops on the analyzed volume, 7615 m^3, while the tank already holds 7650 m^3, 50 m^3 past



Figure 15.5 Tank view of case 4, listed in the scenario menu as 4 · A late upset with 20 % left, at $t = 1.60$ h (96.0 min), the instant the batch completes and the application pauses the run, with the optimizer writing the recipe and cumulative-average targeting on, cropped to the plant view with numbered callouts. The scene reads batch complete with 50 318 bbl in a tank of 50 318 bbl and the header still metering 6278 bbl/h, since the valves close at the next control execution. The badges (1) give the tank at AKI 87.15 and 7.70 psi, $12 610 of give-away, $91.14 a barrel, 0.46 vol % benzene and 0 bbl remaining. Beside the product tank (2) its record shows the 6400 m^3 that started the case hatched below the 87.0 limit and the brighter slices blended near 88.4 above them. The remaining-authority chart draws Δ/f for the 0.049 AKI still missing against f from 0.5 to 0, with the present point (3) at the top of its scale, marked stopped below $f = 0.05$, and the floor at $f = 0.05$ crossing the +1.5 header limit (4). The certificate (5) passes all nine rows, with $7 578 of octane and $5 031 of RVP give-away, and the status line reads batch complete, tank certified. In the full window, outside this crop, the FY-400 card reports the analyzed AKI of 87.15 below the 0.05 floor, stopped, and the trends show the header AKI target near 88.42 until the stop at 1.25 h and 87.20 after it.

the floor, and the header needs another 131 s, about 36 m^3 of blend, to come back down. Four variants of the case, run headlessly with the starting tank made weaker by lowering only the reformate octane of its fill, a change the application does not offer, test the clipped branch. For starting averages of 86.80, 86.70, 86.67 and 86.60 the header target runs on its +1.5 clip at 88.70 from start to stop, the formula predicts 87.105, 87.025, 87.001 and 86.945, and the engine finishes at 87.115, 87.035, 87.011 and 86.955. The shipped floor and clip can therefore certify a tank found 80 percent full at an average as low as 86.67 and no lower. With the floor moved to 0.02, 0.10 and 0.15 instead, the formula gives 87.176, 87.080 and 87.020 for the shipped case and the engine finishes at 87.187, 87.089 and 87.027. The trims can meet the same octane demand, but by swapping reformate for LSR, and with targeting on they certify the octane at 87.15 while the aromatics end at 25.17 vol %, over the limit, a constraint that only the optimizer knows about.

15.6.5 An Optimizer on Its Constraints

In case 6 butane drops from $40 to $25 a barrel at 5 min and the optimizer takes over at 10 min. Moving each fraction 2 vol % per analysis, it reaches its optimum at the analysis at 12.0 min and then

follows it as the bias moves with the analyzer noise. From 30 to 60 min the applied recipe averages butane 2.40, LSR 17.05, reformate 26.31, FCC gasoline 27.54 and alkylate 26.69 vol %, and the header averages AKI 87.19 and 7.70 psi, on its targets. The optimum is pressed against AKI, RVP, aromatics and sulfur at once. At $25 butane the recipe costs $89.82 a barrel against $91.00 for the shipped one, a saving of $1.18 a barrel or about $59 000 on an 8000 m^3 batch, and the give-away of the blend falls from $0.94 to $0.29 a barrel. Run to the end, the batch certifies at 87.20 for $89.85 a barrel. Over the same half hour the bias-corrected linear model judges the true blend with a mean error of 0.008 AKI and −0.005 psi, inside the 0.05 AKI and 0.03 psi of one of the application's own checks.

15.6.6 An Analyzer That Drifts

In case 7 the optimizer blends the on-target recipe and, from 10 min, the NIR's octane drifts upward by 0.2 ON an hour, with the laboratory correction off. The bias update cannot tell the analyzer from the model. The model's RON bias follows the NIR to +0.68, and the optimizer cuts the octane it believes it has in excess, alkylate falling to 23.9 vol %. By 240 min the NIR reads 87.15 while the header is truly 86.50, the NIR's own error is 0.77, and the tank has fallen to AKI 86.85, off specification. The laboratory sees the discrepancy, the NIR reading 0.74 high on its last result, at 230 min. With the correction set on before the run, each result moves the NIR most of the way back, and the tank is at 87.03 at 240 min (87.03 to 87.15 over six noise seeds, against 86.84 to 86.85 without it), on specification, with the header lagging the drift by one laboratory interval. Run to 480 min, the tank certifies at 87.06 with the correction and fails at 86.45 without it, needing about 840 m^3 of alkylate.

15.6.7 An Inventory That Runs Short

In case 8 a premium blend takes 1500 m^3 of alkylate at 60 min. The optimizer's cap on alkylate falls from 38.8 vol %, above the 26.7 it was using, to 17.4 vol % at the next analysis, the 1213 m^3 left above the reserve spread over the 6983 m^3 still to blend, and the optimizer then holds alkylate at 17.3 vol %. From the analysis at 63 min it has more than 5 vol % of ethanol in the blend, 5.5 vol % at 75 min, and it pays for ethanol's vapor-pressure bump by cutting butane from 2.4 to 0.7 vol %, a new vertex of the program. For the three analyses at 61, 62 and 63 min the step-limited program applied to the blender needs its elastic slacks, while the unlimited program shown on the FY-300 card stays feasible, and the header dips only to AKI 87.07. The batch certifies at 87.18 with 20 m^3 of alkylate left, at $90.86 a barrel. The fixed recipe, selected before the run, keeps drawing the case's 26.7 vol % of alkylate, runs the tank dry at 337 min, blends on without it and has only 7364 m^3 in the tank at 480 min. It completes the batch at 532 min at AKI 86.67 and 8.02 psi, off specification on AKI, RVP, aromatics and sulfur with no single stock left that could correct it.

> **+ DID YOU KNOW? The Cheapest Batches Give Nothing Away**
>
> From the simulator (Table 15.5). The two batches with $0 of give-away are also the two cheapest, at $88.60 and $88.88 a barrel, and both fail four specifications. Every certified batch gives away between $7 835 and $15 493. A scorecard of cost and give-away alone would rank the failures first.

15.6.8 The Graded Batch

The challenge blends the 8000 m^3 batch through a reformate swap at 120 min (−3 RON) with 1600 m^3 of alkylate in its tank, less than the shipped recipe draws. The target asks for the batch to finish within 8 h, certified on every specification with no reblend, with no component tank run dry and with less than $16 000 of give-away. As shipped, on the fixed recipe, the alkylate runs dry at 300 min and the blend goes on without it at 680 m^3/h, so only 7040 m^3 are in the tank at 480 min, at AKI 86.21 with 25.84 vol % aromatics and 10.31 mg/kg sulfur, and no single stock left can correct it. The batch completes only at 565 min, at AKI 85.97 and off specification on RVP as well. The trims

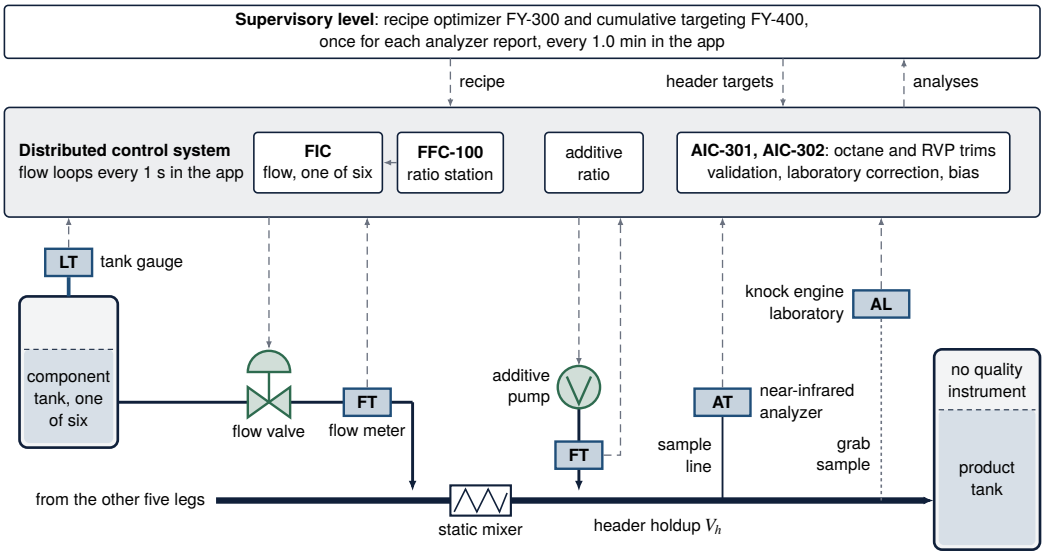


Figure 15.6 The loops of the blender with their hardware. Lettered boxes are sensing devices, namely the tank gauge (LT), the flow meters (FT), the online analyzer (AT) on its sample line and the laboratory (AL). Valve and pump symbols are final elements. Dashed arrows are signals. Notice that every quality reading is taken from the header, late, and that nothing reads the quality of the product tank, so the tank average that the certificate is written on exists only as a calculation at the supervisory level. Not to scale, and not a piping diagram.

do no better, with the alkylate dry at 301.9 min and the tank at 86.80 at 480 min, because they react to the octane and not to the inventory. With the optimizer and cumulative targeting, the suggested configuration, the optimizer spreads the alkylate over the whole batch from the start and covers the octane with ethanol, 5.45 vol % in the finished tank, its bias catches the reformate swap, and the batch completes at 480 min certified at AKI 87.19 and 7.70 psi with \$14 511 of give-away, at \$91.57 a barrel. The give-away limit was calibrated on that run with a margin of 10.3 %, and six noise seeds give \$14 511 to \$14 861, all passing. The optimizer does most of the work, and without targeting it certifies at 87.19 with \$14 041 of give-away. The batch is judged on the case as given, so a changed model parameter or disturbance voids the attempt.

15.7 Technology in Practice

Figure 15.6 draws one of the six component legs and the quality loops of Section 15.5 with the hardware a refinery installs.

What Measures It

The blend is read twice, by an online near-infrared analyzer (Figure 15.7) and by a laboratory knock engine that rates a sample by the research and motor methods, ASTM D2699 and D2700 [18]. Refiners install a limited number of analyzers, at the blender outlet. For cost and reliability the component properties are generally not measured on line, and the readings of the blend can suffer large and variable delays that depend on the sampling loop and the measurement technology [6]. The holdup V_h, the dead time θ_s and the cycle T_A of Equation (15.7) stand for that loop. Laboratory data have a very low time resolution and can be late or inaccurate, through a time stamp set half by hand, the location of the sampling point or a faulty entry [6]. The model's laboratory is late but unbiased.

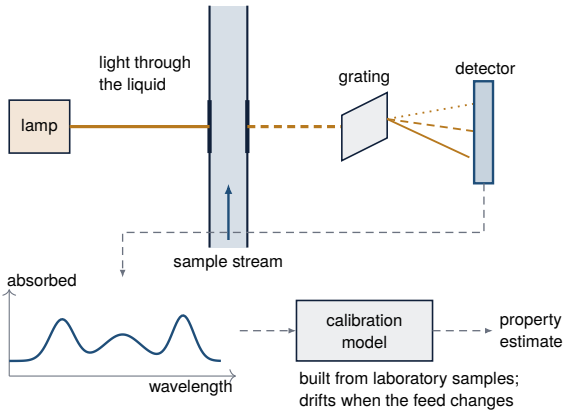


Figure 15.7 Principle of a near-infrared analyzer on a liquid stream. Notice that the detector delivers a spectrum and not a property. The octane and the vapor pressure come out of the calibration model, which is why Equation (15.7) gives the reading its own error b_N apart from the blend it describes.

> **TECHNOLOGY** **Sensor: Near-Infrared Analyzer**
>
> A near-infrared analyzer passes light through the fuel and records how much of each wavelength is absorbed. The bands are overtones and combinations of the vibrations of bonds such as carbon to hydrogen, and information on the different types of hydrocarbon is present throughout the region [20]. A measurement takes one minute or less and suits in-line use, and optical fibers let one instrument watch points up to 100 m away [20]. The analyzer does not measure octane. A multivariate model turns the spectrum into a property, and the method depends on a good reference method for the samples that teach it [20], here the knock engine. The 1.0 min cycle T_A of Table 15.3 stands for the scan. An analyzer fed by a blocked sample line goes on reporting stale sample with no way of knowing it [21], the frozen reading that validation flags STALE.

What Moves It

The recipe reaches the plant through six flow loops, each a meter, a controller and a valve. In ratio control a flow that is not controlled is called wild, and a flow held in proportion to it is called captive [21]. In this blender all six flows are captive to the master rate Q_{SP} of Equation (15.9), and the additive pump is captive to the metered header flow.

> **TECHNOLOGY** **Final Element: Flow Control Valve**
>
> Each component leg ends in a valve under a flow controller, and refiners call the set of these loops the ratio control system [6]. On an air-operated valve a positioner watches the stem and drives it to the position the controller's signal asks for [21]. The usual trouble is friction at the stem. The stem holds until enough force has built up and then jumps, so a smooth signal becomes a staircase, an effect called stiction. With integral action in the loop the controller output then ramps up and down without end, a slip-stick cycle, and detuning only lengthens its period [21]. The model's valves have a rangeability of 30 to 1, a 5 s flow loop and no friction.

Where the Controller Runs

A refinery runs such loops in a distributed control system, in which a supervisor sends set points to distributed controllers that act through actuators from sensor feedback, with cycle times from milliseconds to minutes [22]. The app executes its flow loops every 1 s, and the trims, the optimizer FY-300 and the targeting FY-400 once for each analyzer report, every 1.0 min. An industrial blend controller built from an optimizer and an estimator was reported to run on standard control systems

every 1 to 5 minutes [6]. The ratio station FFC-100 and the six flow loops would sit in the distributed controllers as function blocks, and the optimizer and the targeting at the supervisory level.

> **IN PRACTICE** **Certifying a Blend in the Line**
>
> United States fuel rules have a manufacturer test each batch of gasoline for sulfur, and each summer batch for vapor pressure, before it leaves the plant [23]. An in-line blender may ask for a waiver. The request describes how the automated system monitors and controls the proportions and the properties, and how composite samples are taken to ASTM D4177 and D5854. The operation is audited every calendar year, and a new request is due 60 days before a material change, such as new analyzer hardware or programming, a moved analyzer or new mixing control logic. The summer ceiling is 9.0 psi, or 7.8 psi in listed areas, with 1.0 psi more for blends of 9 to 10 percent ethanol [23].

What Is Modernizing It

Three developments bear on this chapter. The first is certification from the line. The waiver of the In practice box lets composite samples and an audited automated system stand in for a test of every tank [23]. The second is control that does not trust the component table. The industrial controller cited above estimates the component properties from the blend analyses in closed loop, and its estimates need not converge to the true values for the blend to reach its target [6]. The bias update of Equation (15.12) takes the same stance. For a tank that is not empty at the start, that controller uses a heel management strategy, a reference for the header computed from the tank's balance [6], which is the cumulative targeting of Equation (15.13). The third is the link upward. In many modern systems the distributed control system is interfaced with the corporate network to give business operations a view of production [22].

15.8 Limitations

The blending rules are correlations with illustrative coefficients, and the component table and the specification are teaching values for one summer grade, not the data of a refinery or a certified blend. Ethanol blends at one RVP whatever its concentration, and volumes blend without shrinkage.

The analyzer reads the header by the same rules that make the header, so it has noise, bad analyses, drift and freezes but no chemometric error, and the laboratory is unbiased. The targeting averages octane linearly over the analyzed volume, which ends case 4 within 0.001 of the tank's true AKI, and it fills an analyzer outage with the model's prediction plus its last bias, so a long outage is only as good as the model. The laboratory correction lags a drifting analyzer by about one laboratory interval.

The optimizer is a successive linear program with a step limit on one grade, and the valuation of give-away at $1.00 per octane number and per psi is fixed at the base prices. The map's feasible region is a small pocket around the optimum, there is no season switch within a batch, and blend scheduling across several tanks and grades is outside the model. The reformate swap and the alkylate transfer are applied as steps, which a refinery would seldom see, because a step isolates the effect of one cause.

> **OPEN CHALLENGE** **An Analyzer Model That Ages**
>
> The analyzer of this chapter drifts and freezes, but its model of the fuel is never wrong. Real calibration models are less sure. A near-infrared method depends on the quality of its reference values, a model may need a hundred to a thousand samples that span the natural variability, and a sample of different composition shows up as an outlier [20]. A 2025 study states that the usual linear model relies heavily on the match between calibration samples and new ones, that varied crude sources and processes compromise its accuracy, and that automated adaptive modeling needs further development [24]. Progress would be an analyzer that reports when a blend has left its calibration, before the laboratory of case 7 finds out.

15.9 Conclusion

The simulator of this chapter certifies the product tank from its own composition while every controller works from the header, which is the arrangement of a real in-line blender. Butane's index of 139.6 psi$^{1.25}$ makes three percent of it lift the RVP by 1.84 psi where the volume average predicts 1.38, and ethanol raises the AKI by 2.46 and the RVP by 0.95 psi at once. A fixed recipe through a reformate 3 RON weaker certifies at 86.79, and an octane trim tuned at 4 vol % per ON and 1 min oscillates with an 11.3 min period, near the 12 reports that a four-report delay model predicts. A tank 80 percent full at 86.90 needs 88.40 from the rest of the batch and certifies at 87.15 with cumulative targeting and at 86.96 without it, and the closed form for the floor's shortfall predicts the engine within 0.011 AKI down to the boundary at a starting average of 86.67. A drifting analyzer takes the tank to 86.85 through the bias update while the laboratory correction holds it at 87.03, and the graded batch goes from 7040 m^3 off specification at 480 min to a certified 8000 m^3 with \$14 511 of give-away. The model has no chemometric analyzer error, no shrinkage and one grade. A fuller model would add a season switch, blending across several product tanks, a moving-horizon estimate of the tank average and an analyzer whose calibration fails differently from the blend model. The blender shows why a batch that is certified by its average has to be steered from its first barrel.

KEY POINTS

- The certificate is written on the tank, and the tank remembers every barrel. A header on specification does not make a blend on specification.
- A deficit Δ found with a share f of the batch left asks the header for Δ/f more, five times as much at 5 percent left as at 25.
- Each analyzer report describes blend made 3.0 min earlier, so an octane trim tuned like a flow loop oscillates.
- The bias update cannot tell a drifting analyzer from a wrong model. Only the laboratory, an independent method, can.
- Butane is run against the vapor-pressure ceiling and nothing else.

Where else this applies. A late, drifting analyzer checked by a laboratory returns in Volume 5, Chapter 4, a ratio blend trimmed on an analyzer in Volume 3, Chapter 7, a loop on a slow analyzer in Volume 5, Chapter 3, and an estimator that carries a bias in Volume 9, Chapter 4.

References

[1] J. H. Gary and G. E. Handwerk, "Product blending," in *Petroleum Refining*. CRC Press, 2001, pp. 233–249.

[2] M. H. Rusin, H. S. Chung, and J. F. Marshall, "A "transformation" method for calculating the research and motor octane numbers of gasoline blends," *Industrial & Engineering Chemistry Fundamentals*, vol. 20, no. 3, pp. 195–204, 1981.

[3] T. M. Foong, K. J. Morganti, M. J. Brear, G. da Silva, Y. Yang, and F. L. Dryer, "The octane numbers of ethanol blended with gasoline and its surrogates," *Fuel*, vol. 115, pp. 727–739, 2014.

[4] V. F. Andersen, J. E. Anderson, T. J. Wallington, S. A. Mueller, and O. J. Nielsen, "Vapor pressures of alcohol–gasoline blends," *Energy & Fuels*, vol. 24, no. 6, pp. 3647–3654, 2010.

[5] A. Singh, J. F. Forbes, P. J. Vermeer, and S. S. Woo, "Model-based real-time optimization of automotive gasoline blending operations," *Journal of Process Control*, vol. 10, no. 1, pp. 43–58, 2000.

[6] M. Chèbre, Y. Creff, and N. Petit, "Feedback control and optimization for the production of commercial fuels by blending," *Journal of Process Control*, vol. 20, no. 4, pp. 441–451, 2010.

[7] J. E. Bond, R. Coursaux, and R. L. Worthington, "Blending systems and control technologies for cement raw materials," *IEEE Industry Applications Magazine*, vol. 6, no. 6, pp. 49–59, 2000.

[8] E. Sachs, A. Hu, and A. Ingolfsson, "Run by run process control: combining SPC and feedback control," *IEEE Transactions on Semiconductor Manufacturing*, vol. 8, no. 1, pp. 26–43, 1995.

[9] J. F. Forbes and T. E. Marlin, "Model accuracy for economic optimizing controllers: The bias update case," *Industrial & Engineering Chemistry Research*, vol. 33, no. 8, pp. 1919–1929, 1994.

[10] K. Glismann and G. Gruhn, "Short-term scheduling and recipe optimization of blending processes," *Computers & Chemical Engineering*, vol. 25, no. 4–6, pp. 627–634, 2001.

[11] A. Charnes, W. W. Cooper, and B. Mellon, "Blending aviation gasolines–a study in programming interdependent activities in an integrated oil company," *Econometrica*, vol. 20, no. 2, p. 135, 1952.

[12] D. E. Seborg, T. F. Edgar, D. A. Mellichamp, and F. J. Doyle, III, *Process Dynamics and Control*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2016.

[13] M. Auffhammer and R. Kellogg, "Clearing the air? The effects of gasoline content regulation on air quality," *American Economic Review*, vol. 101, no. 6, pp. 2687–2722, 2011.

[14] J. E. Anderson, U. Kramer, S. A. Mueller, and T. J. Wallington, "Octane numbers of ethanol– and methanol–gasoline blends estimated from molar concentrations," *Energy & Fuels*, vol. 24, no. 12, pp. 6576–6585, 2010.

[15] J. E. Anderson, D. M. DiCicco, J. M. Ginder, U. Kramer, T. G. Leone, H. E. Raney-Pablo, and T. J. Wallington, "High octane number ethanol–gasoline blends: Quantifying the potential benefits in the United States," *Fuel*, vol. 97, pp. 585–594, 2012.

[16] D. Splitter, A. Pawlowski, and R. Wagner, "A historical analysis of the co-evolution of gasoline octane number and spark-ignition engines," *Frontiers in Mechanical Engineering*, vol. 1, p. 16, 2016.

[17] G. Edgar, "Measurement of knock characteristics of gasoline in terms of a standard fuel," *Industrial & Engineering Chemistry*, vol. 19, no. 1, pp. 145–146, 1927.

[18] U.S. Federal Trade Commission, "Automotive fuel ratings, certification and posting," Code of Federal Regulations, Title 16, Part 306, section 306.0, edition of January 1, 2026, 2026, govinfo.gov.

[19] W. W. Garvin, H. W. Crandall, J. B. John, and R. A. Spellman, "Applications of linear programming in the oil industry," *Management Science*, vol. 3, no. 4, pp. 407–430, 1957.

[20] C. Pasquini, "Near infrared spectroscopy: fundamentals, practical aspects and analytical applications," *Journal of the Brazilian Chemical Society*, vol. 14, no. 2, pp. 198–219, 2003.

[21] T. R. Kuphaldt, *Lessons in Industrial Instrumentation*, version 2.33 ed. Open textbook, Creative Commons Attribution 4.0 license, ibiblio.org/kuphaldt/socratic/sinst, 2024.

[22] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[23] U.S. Environmental Protection Agency, "Regulation of fuels, fuel additives, and regulated blendstocks," Code of Federal Regulations, Title 40, Part 1090, sections 1090.215, 1090.1310 and 1090.1315, edition of July 1, 2024, 2024, govinfo.gov.

[24] B. Du, H. Yuan, L. Hao, Y. Wu, C. He, Q. Wang, and C. Chen, "Determining olefin content of gasoline by adaptive partial least squares regression combined with near-infrared spectroscopy," *Molecules*, vol. 30, no. 24, p. 4742, 2025.

15.10 Exercises

15.1 ⌕ Concept check · 2 min

With cumulative-average targeting, a deficit found with 5 % of the batch left needs five times the header correction of the same deficit found with 25 % left. Why?

(a) The analyzer delay grows as the tank fills, so a late correction must be larger to show in time

(b) The optimizer may move each fraction only 2 vol % per analysis, a limit that binds harder late in the batch

(c) A nearly full tank mixes slowly, so late blend must carry extra octane to reach the sample point

(d) The header adds $D = \Delta / f$, the deficit spread over the batch divided by the share still to blend

15.2 ⌕ Concept check · 1 min

In case 7 the NIR's octane drifts upward and the bias update of Equation (15.12) follows it. Which element of the blender can tell a drifting analyzer from a wrong blend model?

(a) The knock-engine laboratory, an independent method that tests the analyzer on its own sample

(b) The outlier gate, which rejects a reading far from the model's prediction plus its bias

(c) The bias bound $b_{\max}$, which stops the bias once it grows past four octane numbers

(d) The cumulative-average targeting, which books every analysis into the tank's average

15.3 Concept check · 2 min

Case 3 ships the octane trim at 4 vol % per ON with a 1 min integral time, and the header swings between AKI 85.59 and 87.93 with a period of about 11 min. What causes the oscillation?

(a) The six flow loops resonate with the reporting interval, since their lag is close to one analysis

(b) A move first appears in the fourth report after it, so three more corrections go out before the first is seen

(c) The trim's ±6 vol % limits wind up the integral, since the velocity form has no anti-windup

(d) The NIR noise of 0.10 ON is amplified by the high gain into a sustained swing

15.4 Theory and calculation · 8 min

The shipped recipe has 2 vol % butane and an RVP of 7.36 psi. Butane's neat RVP is 52 psi, and the other five streams in their shipped proportions have an index $I_r = 9.52\,\text{psi}^{1.25}$. Use Equation (15.2).

(a) Compute the blending index of butane and of the shipped blend. *Checkpoint:* 139.6 and 12.1 $\text{psi}^{1.25}$

(b) Butane is raised to 5 vol %, replacing the other streams in their proportions. Compute the new RVP. *Checkpoint:* 9.20 psi

(c) Find the largest butane fraction that keeps the blend at 7.8 psi. The other streams' volume-average RVP is 5.93 psi. What fraction would a blender who used the volume average of the pressures allow, and what is wrong with it? *Checkpoint:* 2.70 vol % by the index, 4.06 vol % by the volume average

15.5 Theory and calculation · 6 min

Ten percent of ethanol is added to the shipped recipe, with the five hydrocarbons scaled down in their proportions, so the base keeps RON 91.39, MON 83.61 and RVP 7.36 psi. Use Equation (15.4) and Equation (15.2) with ethanol's blending RVP of 16.0 psi.

(a) Compute ethanol's blending RON and MON in this base. *Checkpoint:* 125.9 and 98.4

(b) Compute the RON, MON and AKI of the E10 blend. *Checkpoint:* 94.84, 85.09, AKI 89.96

(c) Compute the E10 blend's RVP and compare it with the 7.8 psi summer limit. *Checkpoint:* 8.31 psi, 0.51 psi over

15.6 Theory and calculation · 8 min

Case 4 starts with 6400 m^3 of an 8000 m^3 batch in the tank at AKI 86.90, with a target of $T = 87.20$, a header clip of $c = 1.5$ and a stopping floor $f_{min} = 0.05$. Use Equations (15.13) and (15.14).

(a) Compute f_0, Δ_0, D_0 and the header target. *Checkpoint:* 0.2, 0.24, 1.20, 88.40

(b) Predict the certificate AKI with the floor at 0.05 and at 0.10, and without targeting. *Checkpoint:* 87.14, 87.08, 86.96

(c) Find the lowest starting average, with the tank still 80 % full, that the shipped floor and clip can bring to the 87.0 limit. *Checkpoint:* 86.67

15.7 App challenge · 8 min

Load scenario 1, "1 · Manual recipe: butane and RVP", at the default seed with no change, and run 20 min at the default speed of 5 (about 4 s of wall time). Export the CSV. RVP columns are in kPa, and 1 psi is 6.895 kPa.

(a) In column `rvp_nir_kPa`, find the first row after the butane step at 5 min that shows the step, and give its time and value in psi. *Checkpoint:* 9.08 min, 9.21 psi

(b) In column `rvp_tank_true_kPa`, find the first row above the 7.8 psi limit. *Checkpoint:* 6.67 min

(c) Explain why the tank is over the limit about two minutes before the analyzer shows anything, using the header holdup, the sample line and the analysis cycle. *Checkpoint:* 3.0 min from blend to report

15.8 App challenge · 10 min

Load scenario 4, "4 · A late upset with 20 % left", at the default seed. Use speed 20, at which each batch takes about 5 s. The run pauses by itself when the batch completes at 96.0 min.

(a) Run as shipped. Read the tank AKI and the give-away from the badges. *Checkpoint:* 87.15, \$12 610

(b) Press Reset, set "Stop targeting below f =" on the FY-400 card to 10 % before Run, and run to completion. Read the tank AKI and compare it with the prediction of the theory exercise on case 4. *Checkpoint:* 87.09

(c) Press Reset, set cumulative-average targeting Off on the FY-100 card, and run to completion. Read the tank AKI, the certificate and the corrective volume tiles. *Checkpoint:* 86.96, off specification on AKI

15.9 App challenge · 8 min

Load scenario 7, "7 · Analyzer drift, bias update and the lab", at the default seed, with "Correct the NIR by the lab" Off as shipped. Run at speed 20 until the clock passes 240 min (about 12 s) and export the CSV.

(a) In the row at 240.0 min, read `aki_tank_true`, `aki_nir`, `aki_hdr_true` and `nir_error_RON`. Is the tank on specification? *Checkpoint:* 86.85, 87.15, 86.50, 0.77

(b) Press Reset, set "Correct the NIR by the lab" On on the AT-300 card before Run, run past 240 min again and read `aki_tank_true` at 240.0 min. *Checkpoint:* 87.03

(c) Explain why the bias update did not stop the drift in part (a), and what the laboratory adds.

15.10 Application area · 3 min

The shipped recipe blends to AKI 87.50 and RVP 7.36 psi against limits of 87.0 and 7.8 psi, and give-away is valued at \$1.00 per octane number and \$1.00 per psi for each barrel. An 8000 m^3 batch is about 50 318 bbl.

(a) What is the give-away per barrel, and over a batch? *Checkpoint:* \$0.94 a barrel, about \$47 300

(b) Why does a recipe held by hand keep such margins, and what must an optimizer have before it can remove them?

15.11 Application area · 4 min

The model blends ethanol at one RVP of 16 psi at every concentration, while the chapter's assumptions note that real ethanol lifts the vapor pressure most near 5 to 10 vol % and less above.

(a) With Equation (15.2) and the shipped base at 7.36 psi, what RVP does the model give for E5? *Checkpoint:* 7.84 psi

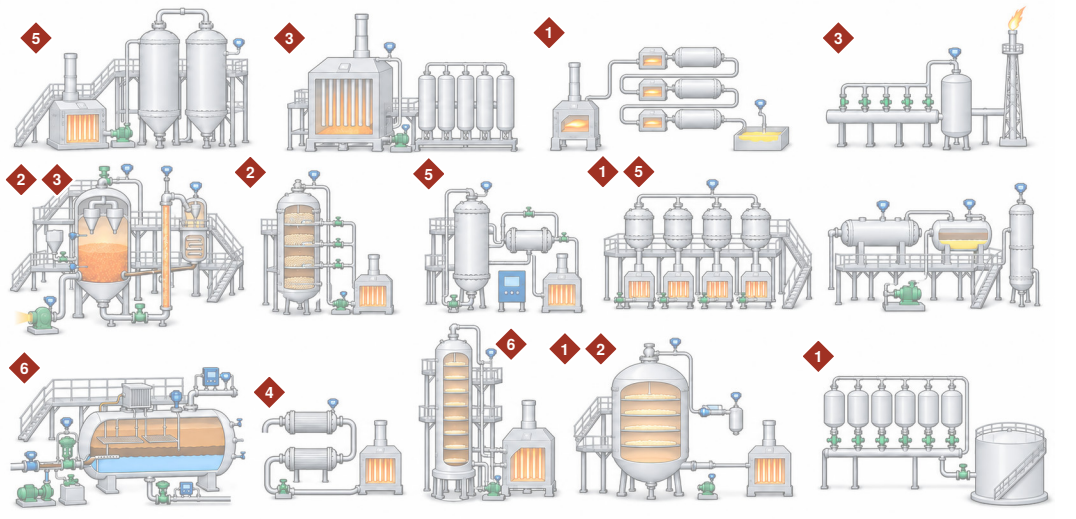
(b) In case 8 the optimizer brings 5.5 vol % of ethanol into the blend and cuts butane to make room. If a real blend's bump at this concentration is already close to its E10 size, is the real RVP above or below the model's, and what in the control system would notice?

15.12 Review the agent · 5 min

An AI assistant reviewing case 3 writes: "The oscillation at 4 vol % per ON and 1 min comes from the integral time. On a loop that is almost pure delay the integral does the work and the gain only sets the first correction, so keep $K_c = 4$ for a fast first move and lengthen τ_I to 2 min. That removes the oscillation." Find the flaw and say how you would check it.

CHAPTER 16

Summary and Outlook



Open gaps 1 inferred qualities 2 estimated limits 3 shared utilities 4 cleaning schedules 5 quality margins 6 changing feed

Notice that each number stands beside every unit where a chapter met that gap, so the first gap appears four times and the fourth once. Illustration, not to scale.

IN PLAIN TERMS **What Fourteen Refinery Units Taught**

The main idea. Fourteen chapters followed crude oil from its first wash to the gasoline tank. In every unit the same three things pulled against each other: how much oil goes through, how good the product is, and how hard the equipment is pushed.

The challenge. The limit that stops a unit is often worked out from other readings, not measured, so it can be late or wrong. The instrument that judges the product can drift while it still reads on target.

Why it matters. A unit that trusts a wrong reading makes bad fuel or harms its equipment while every display looks normal.

Fourteen chapters followed crude oil from the desalter to the gasoline tank. Table 1.2 listed, for each unit, the handle that sets throughput, the quality the unit holds and how it is known, and the equipment limit with the structure that guards it. This chapter adds what worked, what failed and the measured result, and reads the comparison for the question of the volume, how each unit balances throughput, product quality and equipment limits. It then gives a short guide for choosing a design, looks at where practice is heading, reports how the volume was made with AI agents and what they got wrong, and lists the problems that remain open.

16.1 What the Chapters Showed

Tables 16.1 and 16.2 compare the chapters, split between the crude and conversion units and the units that finish, supply and blend. The second column condenses Table 1.2 into the three quantities each

Table 16.1 How Each Unit Balanced Throughput, Quality and Limits, Chapters 2 to 8

Chapter	Throughput, quality, limit	What worked	What failed	Measured result
2, desalter	Charge set upstream. Salt sampled, the interface misread by the rag. Grid current and its trip.	A map optimizer on the mixing valve, a rag trim from samples and 15 ppm of demulsifier.	Design settings on a heavier crude. A level loop on a probe that reads half the rag.	Challenge salt 0.819 PTB, a pass, against a grid trip at 493.3 min as shipped (Table 2.5). True water 750 to 632 mm with the reading on setpoint (Table 2.4).
3, preheat train	Charge is the last handle. Column inlet measured. Firing limit as the train fouls.	A valve position override on the charge at a 95 % target. Cleaning outages kept apart.	Feedforward alone on the firing stop. Two outages 8 h apart.	With an exchanger out, the inlet falls to 340.3 °C, or holds 360.0 °C for 555 t of 19 200 (Table 3.5). IAE 559 against 48 °C h with one clean moved (Table 3.6).
4, crude tower	Charge set upstream. Four handles for eight inferred specifications. Overflash and heater duty.	Four loops along the one-way interaction, the gas oil draw on the overflash.	The top loop alone with cautious tuning, as shipped.	Challenge worst 0.28 vol% and 0.65 °C, dry 1.0 min, against 4.01, 1.92 and 174.2 min (Section 4.5.6). One reflux moves two specifications (Table 4.4).
5, vacuum tower	Lift by pressure and heater outlet. Gas oil 95 % point inferred. Ejector break and tube skin.	A lift maximizer that watches the ejector load and cuts the heater outlet as the critical load falls.	Fixed design setpoints. The same maximizer blind to the ejector load.	Challenge 1153.7 t, no break, against 1097.3 t with a break at 352.4 min, and 1161.1 t with a break when the load is ignored (Table 5.6).
6, catalytic cracker	Feed at a low selector. Severity as riser outlet. Blower, compressor, cyclone and valve limits.	A low selector over four constraint controllers with external reset.	The unit left alone through the shift.	$95,391 and 0 min above the cyclone limit, against $84,532 and 296 min (Table 6.9). Feed held at 284.5 t/h of 330 asked (Table 6.8).
7, hydrocracker	Feed. Conversion on a 25 min analyzer. Bed rise, runaway index 0.538 on bed 1.	An override on a rise compensated by one lag on the inlet.	The same override on the raw rise.	On a feed cut from 250 to 220 t/h the raw rise trips at 28.25 min, the compensated rise peaks at 29.9 °C and holds (Table 7.4).
8, diesel hydrotreater	Feed. Sulfur on a 30 min analyzer. Outlet temperature, heater duty and catalyst life.	Mid-ranging the bypass with the heater, a Smith predictor, a 7.5 wppm setpoint.	Mid-ranging off as the exchanger fouls. A 6.5 wppm setpoint.	Challenge worst 9.11 wppm at 1.21 catalyst days a day, against 386.97 shipped and 1.30 at 6.5 wppm (Table 8.5).

unit trades. The last column gives the measured result with the table or section of the chapter that reports it.

The columns make five points about the balance the volume asked about. Figure 16.1 draws one pair of results from each of twelve chapters.

Throughput is the handle of last resort, given up through a selector. On the preheat train the override held the column inlet at 360.0 °C through an exchanger outage at the cost of 555 t of 19 200, while feedforward alone held the full charge and delivered 340.3 °C with the fuel valve on its stop. On the catalytic cracker the low selector held 284.5 t/h against a request of 330 and gave back 23 t/h when the afternoon warmed from 15 to 38 °C (Table 6.8). The vacuum maximizer cut its heater outlet as the cooling water warmed and finished the challenge with more gas oil than the fixed setpoints. In

Table 16.2 How Each Unit Balanced Throughput, Quality and Limits, Chapters 9 to 15

Chapter	Throughput, quality, limit	What worked	What failed	Measured result
9, reformer	Feed. Octane from an analyzer, the laboratory and an inference. Tube metal and regenerator.	A trim on the inferred octane, feedforward from the naphtha, the aged heater distribution.	Any trim that trusts the shifted analyzer. Inlet temperature alone, as shipped.	In band 100 % of the judged time with tube metal at 606 °C, against 77.9 on the analyzer and 13.6 shipped (Table 9.5).
10, alkylation	Olefin feed. Acid strength titrated every 4 h. Refrigeration and the strength floor.	Loops with feedforward, a density soft sensor and moisture feedforward. A supervisor on the capacity rows.	Fixed flows. A strength loop on the titration alone.	Pass at $21,995, supervisor $23,839, fixed flows fail at a low of 93.52 RON (Table 10.7). Strength swing 0.86 points on titration, 0.11 on the soft sensor (Table 10.5).
11, delayed coker	Charge on a drum cycle. Heavy gas oil cut. Tube skin, foam and the drum switch.	Feedforward from the schedule and the shell thermocouples. Antifoam on the foam level.	Loops alone at a switch onto a cold drum. A fixed antifoam ratio.	Pressure dip 0.024 against 0.186 bar, cut within 1.0 against 9.7 K, antifoam 6.8 against 18.9 kg as shipped (Table 11.5).
12, hydrogen plant	Rate on header pressure. Steam to carbon floor, purity from the adsorber. Tube metal.	A lead-lag on demand, a firm header loop, a cycle that follows the feed, tail-gas feedforward.	A plain ratio on a ramping steam supply. The firm loop without the feedforward.	Passes at five seeds, header within 0.395 bar. Without the lead-lag the ratio falls to 2.711, without feedforward the tube reaches 927.1 °C (Table 12.4).
13, sulfur recovery	Acid gas set upstream. Air demand on a 3.2 min analyzer. Stack permit and dew point.	Flow and analysis feedforward under a PI trim.	The trim alone on a flow step. A trim that trusts a drifting analyzer.	Air demand IAE 7.3 with the flow ratio, 1227 without (Table 13.6). A 0.5 mol% drift puts 9.60 t/d up the stack (Section 13.5).
14, fuel gas header	Off-gas set by other units. Pressure and Wobbe index measured. Importing and flaring at once.	A split range with a dead band, Wobbe feedforward, recovery stages set below the seal.	A 4 % overlap. Stages above the seal. Feedforward off.	Challenge flares 6.1 to 6.3 Nm^3 with no time importing and flaring at once, against 7203 and 480 min shipped (Table 14.6).
15, gasoline blender	Blend rate into a tank. Tank-average octane and vapor pressure. Component stocks and give-away.	An optimizer with cumulative targeting and the laboratory correction.	A fixed recipe. Targeting off. The analyzer's bias update alone.	Certified at 87.19 with $14,511 of give-away, against 85.97 off specification shipped. Targeting on and off 87.15 and 86.96 (Table 15.5).

each case the unit gave up rate only while a limit bound and only as much as the limit asked. A unit that keeps its rate and lets the limit decide loses the quality instead.

A guard protects only as well as the signal it reads. The hydrocracker's override tripped the unit 8.25 min after a feed cut when it read the raw bed rise. The inlet answers a quench valve within 0.30 min and the outlet lags by a 3.5 min catalyst time constant, so the raw difference was inverted during the correction (Section 7.10). One lag on the inlet fixed it. The vacuum maximizer blind to the ejector load ran the heater outlet to 425 °C and the ejectors to 149 % of their critical load, against 85 % when the load was watched (Table 5.5). A limit is usually inferred, from a difference of two temperatures or from a load, and the inference has dynamics of its own.

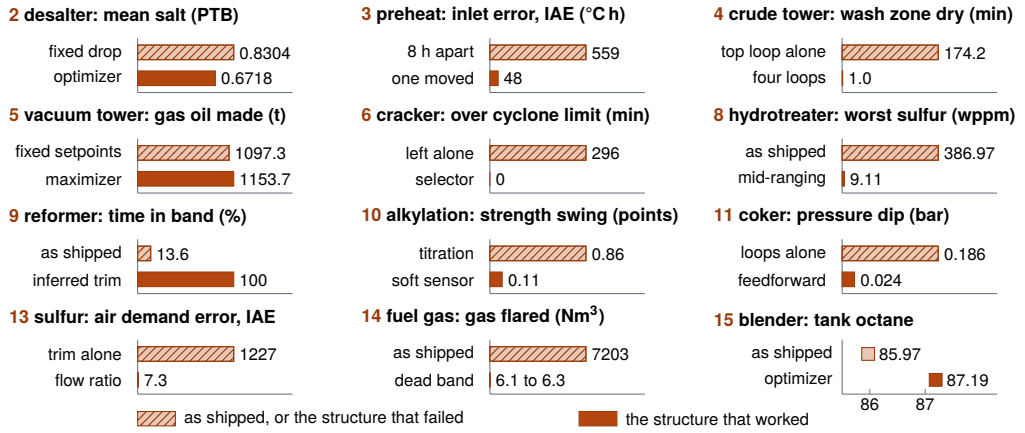


Figure 16.1 One pair of results from each of twelve chapters, with the numbers as Tables 16.1 and 16.2 print them (the desalter pair as this section quotes it from Table 2.6). Each panel has its own scale from zero. The blender alone is drawn with markers, on an axis that starts at 85.5. A larger value is better in the panels of Chapters 5, 9 and 15. The tables give the hydrocracker and the hydrogen plant no like pair. Notice that where a measured load was fed forward or a limit was guarded, the solid bar nearly vanishes.

Quality margin is paid for in catalyst, yield or tube life. The diesel hydrotreater passed its graded day at a 7.5 wppm setpoint with 1.21 catalyst days used per day and failed at 6.5 wppm on 1.30, although its worst sulfur reading improved. On the reformer 10 K of inlet temperature bought 2.01 RON for 1.51 points of liquid yield (Section 9.4). On the coker the same 10 K traded 1.9 points of coke for coils that coke 2.2 times as fast (Table 11.4). The setpoint of a quality loop is therefore an economic choice. A margin wide enough for the analyzer's noise and no wider is the target, and only a sweep over seeds or feeds shows where it lies.

A late quality is held by what is measured early, and lost to an analyzer that is trusted. A flow ratio on the sulfur plant cut the air demand error after a flow step from 1227 to 7.3. The coker's schedule feedforward cut the pressure dip of a cold drum from 0.186 to 0.024 bar. The alkylation soft sensor held the strength within 0.11 points where the four-hourly titration let it swing 0.86. The opposite failure recurs as well. The reformer trim that believed a shifted analyzer held its reading while the true octane fell to 94.24 (Table 9.4). The sulfur trim that believed a drifting analyzer sent 9.60 t/d up the stack with the air demand on target. The blender certified only with its laboratory correction on, at 87.06 against 86.45 (Table 15.5). Feedback holds the reading. A second source is needed to hold the product.

Optimization paid where it set the target. Prediction paid less than it first appeared to. The interior optima were found by an optimizer. On the desalter the map optimizer reached a mean salt of 0.6718 PTB where extremum seeking reached 0.7842 and the fixed pressure drop 0.8304 (Table 2.6). The predictive controllers were mixed. On the crude tower's end point steps the plan cut the freeze point IAE from 123.9 to 18.1 °C min and lost the heavier crude (Table 4.6). On the alkylation unit it held the strength between titrations at an IAE of 1.20 against the supervisor's 6.18 (Table 10.6). On the hydrocracker its higher conversion, 0.5501 against 0.5427, came from the bed 1 target, and the loops reproduced it with that setpoint lowered by hand (Table 7.5). On the preheat train its extra tonnes came from the valve's last five points, and at the override's 95 % ceiling it processed less (Table 3.8). On the cracker it crossed no limit in seven cases at about one percent less feed (Table 6.10). Volume 9 treats that choice in full.

> ⧖ HISTORY **2003: What a Survey of Predictive Control Left Open**
>
> The survey that counted the predictive controllers of Chapter 1 also listed what they could not yet do [1]. Most were built on models from a plant test, which ran day and night and could take up to 5 to 15 days, and no tool told the engineer when the data were adequate. No method verified whether a model had deteriorated after commissioning. Only two adaptive algorithms had reached the market, despite a strong incentive for a controller that tunes itself. The practitioners' questions were about engineering, not algorithms: when a model is good enough, and how to design control for an entire plant. In 2023 a review found new applications still built mainly on linear models from step responses [2]. Gaps 3 and 6 of Section 16.5 are those questions.

16.2 Choosing a Method

The chapters suggest an order of questions for a refinery unit. Each names the chapter where its answer decided the result. Figure 16.2 sets the same questions out as a flowchart.

1. *Which limit binds first, and how is it known?* List the limits and the signal behind each. A limit read from a difference of two temperatures or from a load needs its own dynamics checked, as the compensated rise of Chapter 7 and the ejector load of Chapter 5 show.
2. *Which handle may give up throughput?* Put it at a selector with external reset, with a valve or capacity target below the stop, so that rate falls only while a limit binds (Chapter 3, Chapter 6, Volume 4).
3. *Does the unit feed back its own heat?* Compute the loop gain of Table 1.1 before tuning. A gain near one slows and amplifies every move, and a fouling exchanger can push the inlet out of reach of the heater (Chapter 8).
4. *Is the optimum on a limit or inside the range?* An optimum on a limit is held by a selector. One inside is found by an optimizer, as on the desalter, or turned into a setpoint at zero, as the air demand of Chapter 13 is.
5. *How late is the quality, and what is measured ahead of it?* Feed forward the measured load first, then trim on the analyzer, and keep an inference or a laboratory between the analyzer and the product (Chapter 13, Chapter 9, Chapter 10).
6. *Is the product judged as an average?* Steer the remainder, not the instant, and expect the authority to shrink as the tank fills (Chapter 15).
7. *What does the margin cost?* Choose the quality setpoint from a sweep over noise and feeds against the catalyst, yield or tube budget (Chapter 8).
8. *What does a predictive controller add over these structures?* Compare it with the structure given the same target and the same ceiling. In this volume most of its first apparent gain belonged to one or the other.

16.3 Where Practice Is Heading

Optimization joins planning and scheduling. Enterprise-wide optimization was described in 2005 as a new frontier that integrates planning, scheduling, real-time optimization and inventory control across the functions and decision levels of a company [3]. A review of refinery operations gives that integration of decision levels as the emphasis for the industry, from crude unloading and mixing through the conversion units to blending and distribution, and surveys the methods for scheduling and planning [4]. Model-based real-time optimization of gasoline blending was reported in 2000 [5]. In the editor's view the coordinating layer of Section 1.4.7 will reach further across units, and the blender's cumulative target, set from the tank and not the instant, is the pattern it will follow.

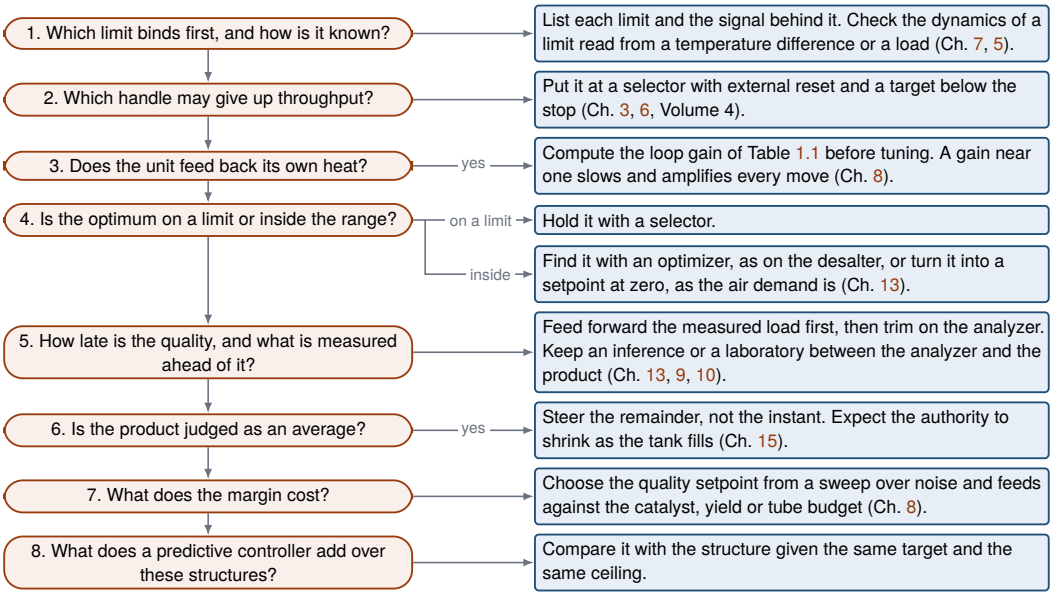


Figure 16.2 The eight questions of this section in order, each with the answer the list gives and the chapters where that answer decided the result. The third and the fourth question branch. Notice that seven of the eight are settled before a predictive controller is considered.

Models are maintained, not only commissioned. A 2015 review of industrial predictive control reports that practitioners now focus on easy commissioning, monitoring and the automated maintenance of applications, with continuous performance monitoring and automated reidentification of models [6]. A review of soft sensors for regression treats their maintenance alongside input selection, training and validation [7]. A perspective on machine learning for process data gives three attributes that process data analytics needs for such methods to apply in the process industries, and stresses a culture of data analytics [8]. The reformer's shifted analyzer and the vacuum tower's inferential that read 558 °C against a true 554 °C at a stripping steam it was not calibrated for (Section 5.5.4) are the cases such maintenance exists for.

Fouling and cleaning are scheduled from models. Optimal cleaning schedules for crude preheat trains have been posed as mixed-integer nonlinear programs, with fouling models reconciled from plant data on a network of 27 exchangers [9]. A review of threshold fouling models finds the most activity in using them to predict fouling, and gives the need to include the chemistry of fouling as one reason the models themselves have advanced little [10]. On the preheat train of Chapter 3, moving one of two colliding outages 32 h later cut the integrated error of three weeks from 559 to 48 °C h, which no tuning in the application could match.

The feed changes. Co-processing of bio-based feedstocks with petroleum fractions, through hydroprocessing and catalytic cracking, has been demonstrated industrially in some cases and reviewed as an option that uses existing refinery units [11]. In the editor's view such feeds will move the heat balances of the cracker and the hydrotreaters and the hydrogen they draw from the plant of Chapter 12, so the limits of these chapters will bind at new places, and models fitted to one crude will need the maintenance described above.

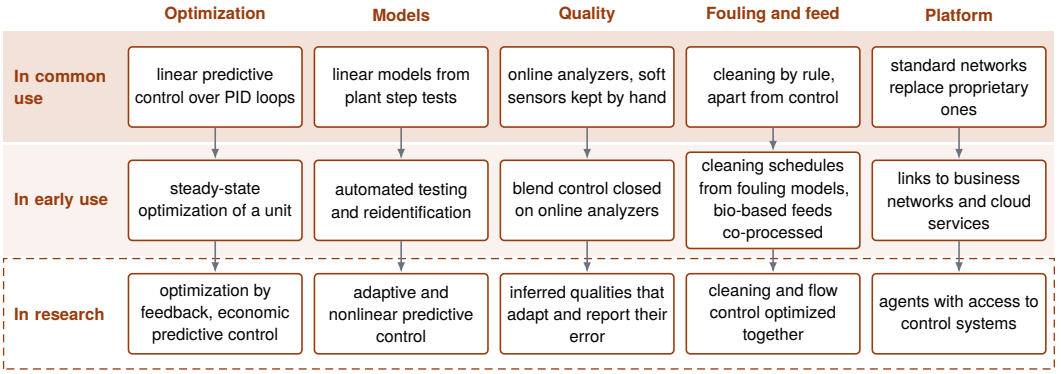


Figure 16.3 The developments of this section in five threads. The top lane holds what the cited source describes as standard practice, the middle lane what is in service at some sites or shown on plant data, and the bottom lane research. The placing is the editor's reading of the sources, and no box carries a date. Sources by thread: optimization [1, 2, 12], models [1, 2, 6], quality [13, 16], fouling and feed [9, 11, 17] and platform [15]. The last platform box is the fourth gap of Section 16.5.1. Notice that every thread ends in something that has to be kept right, a model or a measurement.

Optimization moves into the control layer. Most commercial real-time optimizers repeat two steps, an update of a steady-state model from plant data and a numerical optimization of the setpoints. A 2022 review states that this form is not as widely used as one would expect, and gives as reasons the cost of the model, the wait for steady state before each update and conflicts with the planning layer [12]. The same review surveys two decades of research on the alternative, in which the economic objective is translated into controlled variables and feedback loops bring the unit to its optimum without an optimization solved online. Joining control and optimization in one economic predictive controller is reported as rarely a good solution for practical implementation and tuning [2]. In the editor's view the selector that holds a unit on its binding limit, and the air demand held at zero in Chapter 13, are small cases of that alternative.

Analyzers close more loops. A blend controller that closes the loop on online analyzers downstream of the blender has been in service at one refiner in its present form since 2001, and by 2010 it ran on 17 blenders in six refineries [13]. Near-infrared spectroscopy, the analyzer of Chapter 9 and Chapter 15, takes one minute or less per sample and suits in-line use, and it stays dependent on a good reference method [14]. In the editor's view each loop closed on an analyzer adds a case of the first gap of Section 16.5, because the loop holds the reading whether or not the reading is right.

Control systems join business networks. A federal guide records that control systems, at first isolated on proprietary protocols and specialized hardware, are taking up Ethernet, Internet Protocol and wireless devices and are becoming more integrated into business networks and cloud infrastructures [15]. The integration leaves them much less isolated from the outside world and creates a greater need to secure them. In the editor's view the optimizers and the model maintenance of this section depend on that connection, and the security gap of Section 16.5.1 grows with it. Figure 16.3 sets the developments of this section on three lanes, by how far each has come into use.

16.4 Engineering With AI Agents

An AI agent is a language model that works toward a goal in a loop. It plans, calls tools such as a compiler, a simulator or a browser, reads the results and continues [18]. In the editor's view, agents

can now carry out much of the work that used to fill the hours of a control project. They derive a model from balances or fit one to plant data. They write the simulator and its tests, code the controller, tune it against stated criteria, script commissioning checks, write the documentation and watch running loops for degradation. The apps and chapters of this series were produced that way under the editor's direction, and the account that follows this section reports what the agents got wrong in this volume and how it was caught.

In the editor's view, this changes where an engineer's value lies. When a model, a controller and a report can each be produced in minutes, producing them is no longer the scarce skill. The work moves to five things.

Quality control. An agent's output arrives as a claim: this model fits, this tuning settles in four minutes, this code implements the equation. The engineer checks claims against the plant or the simulator. A reported number is measured again on a fresh run at the stated seed. The code is read for what it clamps, what it initializes and what it reads, because a display that shows the controller's model in place of the plant looks correct until the two differ. A tuning is tried on a disturbance the agent did not test. The most useful checks are independent of the author, made by a second agent or a person who has not seen the first answer, and they are checks that can fail.

Supervision. The engineer sets the scope of each task and decides what an agent may touch. Reading a historian and proposing a tuning is one level of access. Writing a setpoint, downloading a configuration or changing an interlock is another, and it passes through the plant's management of change with a named approver. Supervision also fixes the order of work: a baseline before a change, one change at a time and a comparison that shows what moved.

Creativity. An agent answers the question it is given. Choosing the question remains the engineer's task: which loop is worth improving, whether a disturbance calls for feedforward, a cascade or a change to the process, which step test the operation can tolerate, and which comparison would show that a predictive controller earns its place over a retuned PID loop. The experiment that separates two explanations is designed by someone who can imagine both.

Process insight. A model contains what its author put in. The equipment does more. A valve sticks, the dead time of an analyzer grows when its sample flow drops, a thermowell lags, a heat exchanger fouls and an operator intervenes at the same hour each night. The engineer who knows the unit recognizes a result that is too good, asks what the model left out and knows which measurement to distrust first.

Accountability. Responsibility does not transfer to a tool. Functional safety standards assign the activities of the safety lifecycle to persons and organizations of stated competence, and they require an assessment that includes someone independent of the design team [19]. An agent may draft the hazard analysis or the test procedure. A person owns the decision to accept it, to start up and to keep running.

These five are the professional form of the habit introduced in Section 1.7: check before trusting, and know enough to check.

16.4.1 How This Volume Was Made

Each of the fourteen apps was built and tested by agents. An agent then wrote the paper on it, keeping verification notes that trace every number to a run, and a referee report was written on each paper. At least ten of the fourteen reports were written by the agent that wrote the paper, reading its own draft as a referee would. Seven papers had a further referee pass on a controller added later, a predictive

controller on five, extremum seeking on one and a Smith predictor on one. Defects went into a log for each app with a severity, a cause, a suggested fix and a status. For this book one agent per chapter reran the numbers its exercises use, after a repair round on 2026-11-15 had applied the book's earlier defect list to apps and papers. The editor set the scope of each round and the rules on what each agent could change, and statements that runs contradicted were left for the editor to decide.

The defect logs. The fourteen logs written beside the papers hold 88 entries with a severity. Of these, 30 are wrong behavior or output of an app, 44 are documentation, 9 are cosmetic or display and 5 are other kinds, such as a fragile test or a session blocker. The counts understate the work, because nine apps keep a second log of their build in their own folders, which is not counted here. Four cases show the errors that matter for a volume about throughput, quality and limits.

An optimizer that reported optimal beyond its own limit. The alkylation unit's economic supervisor kept its targets beyond its own reboiler row once a feed increase put them there, and its card read "optimal". In the challenge under the supervisor the targets sat beyond the row for 1,290 of 1,440 minutes, the reboiler was at its steam limit for 1,158 minutes and the alkylate carried up to 21.8 % butanes. The cause was a linearized row that only stopped the step from moving further out. The agent adding the predictive controller found it, reported it without fixing it, and a separate fix agent made the row pull the targets back and report "infeasible" where it could not. The fix moved the supervisor's challenge row to the $23,839 of Table 16.2. A regression test fails on the old code. The scored octane had not shown the fault, because the verdict read the octane correlation and not the butanes.

Protections that read what no instrument measures. The hydrocracker's override cut the inlet setpoint on a runaway index computed from the plant's true kinetics and catalyst activity, so it reacted to an activity change before any thermocouple moved. The fixed override estimates the index from the rise it measures, and no shipped run moved. The catalytic cracker's afterburn and bed fences were computed from the true temperatures and not from the thermocouples the app models, a difference of up to 20.4 °C at the promoter loss of the challenge. That fix moved the challenge margin from $95,402 to $95,391, the number in Table 16.1. The fuel gas header's predictive controller took the valve capacities from the run's settings, so a makeup capacity edited from 8000 to 12 000 Nm^3/h moved its first move from 20.94 to 7.59 % with nothing measured changing. Each was caught by an audit that changed a hidden value at fixed measurements and asked whether a controller moved.

A limit computed from a lagged quality. On the hydrogen plant the predictive controller's ceiling on the adsorber rate scaled the rate by a purity reading that lags it through a drum and an analyzer, and applied it to a rate already cut. With the cycle fixed, the rate fell from 130 to 45 % in 9 min, the true steam to carbon ratio swung from 1.91 to 7.35 and the exercise stopped at 83 min. On the hydrocracker the first price on the bed 1 inlet limit let the plan warm the bed 2.1 °C past that feed's fold for 5.3 min after a trip reset. The hydrocracker case was found by the technical review of the added controller, which ran restart cases the author had not. The hydrogen case was fixed in the same audit round as the reads of the previous paragraph, with a regression test for it.

Gains credited to the wrong layer. Three referee passes on the predictive controllers corrected the credit. The hydrocracker draft credited the higher conversion to the controller, but the rise constraint bound in none of its 480 plans, and the shipped loops with bed 1's setpoint lowered by hand reproduced the result to four figures. The preheat draft said the plan "cuts only the charge it must", but at the override's valve ceiling it processed 92 t less, so the tonnes were the valve's last five points. The crude tower draft led with the case the plan won on every tile and left the heavier crude,

which it lost on every quality, for later. Each correction added a like-for-like row to the chapter's table.

The book round. Each chapter agent first confirmed that its paper had not changed since conversion, and none of the fourteen had. The fourteen chapter records list 216 rows of numeric verification, each a stated value or group of values set beside its reproduction, and all fourteen report that every rerun number reproduced. The book round's defect list holds eleven entries for this volume. None reports a new fault in control behavior. They are text, display and presentation, and one start-up limitation the owner had already chosen to keep. Two papers named a controller's action against the convention of Chapter 1, and the chapters now state it. Two chapter summaries contradicted their own measurements and were aligned. The fuel gas chapter said four times the header volume buys four times the time, against its own 45 and 133.5 s, and the crude tower chapter called a ratio of 0.61 fixed where its own runs give 0.63 over a wider range. The agent writing Chapter 1 first wrote that each optimum lies on a limit, and two chapters contradicted it. The book's tools were checked as well. Three chapter records noted wide drawings, scaled down to the text width, that the build's check report had not listed. The orchestrator traced the cause to a look-ahead error in the width log of the book's style, which had never fired in a real build, so an earlier recheck of four volumes was void. The log was fixed, the earlier volumes were checked again, and three more drawings were redrawn.

Every error above was caught by a check that could fail: a hidden value moved at fixed measurements, a structure given the controller's own target, a case run that the author had not run, a log that should have listed a known wide figure. Four working rules follow. Read every limit a controller enforces from an instrument and say which. Compare a new controller with the old structure at the same target and the same ceiling. Ask an optimizer what it does when it starts outside its constraints. Let a person decide what changes.

16.5 Gaps in Refinery Control

Six gaps belong to the control of refining units, and the gaps common to agent-era practice follow them in Section 16.5.1. In the editor's view none is closed. Each item gives the evidence of the chapters, the state of the literature, why the gap is hard and what progress would look like. A sentence on difficulty or progress that carries no citation is the editor's view. Figure 16.4 places the gaps by the chapters that met them and by the maturity of the method available, and the same numbers mark the units in the picture that opens this chapter.

1. *Inferred qualities that know when they are wrong.* The vacuum tower's inferential misread the gas oil end point by 4 °C at a stripping steam it was not calibrated for, the reformer's trim followed a shifted analyzer, and the sulfur plant's drift was seen only by the stack alarm. The blender certified only with its laboratory correction on. In the editor's view routine detection of a drifting analyzer or inferential against a second, independent source is not standard practice, and a review of soft sensors found that most have no automated mechanism for their maintenance [16]. The Open challenge box after this list takes up why the gap is hard and what progress would look like.

2. *Limits that are estimated, with known lag and bias.* The ejector's critical load, the bed's runaway index and the cracker's fences are computed, not measured. A protection built on such an estimate needs a statement of how late and how wrong it can be, and the hydrocracker's raw rise shows what happens without one. The investigation of a refinery explosion drew the same lesson: operating parameters that give only indirect information on a critical condition can leave a unit unable to tell that it is unsafe [20]. The gap is hard because the estimate has dynamics of its own and its error shows only when the limit is near. Progress would be a stated delay and error band for every computed limit, with the margin of the protection set from them.

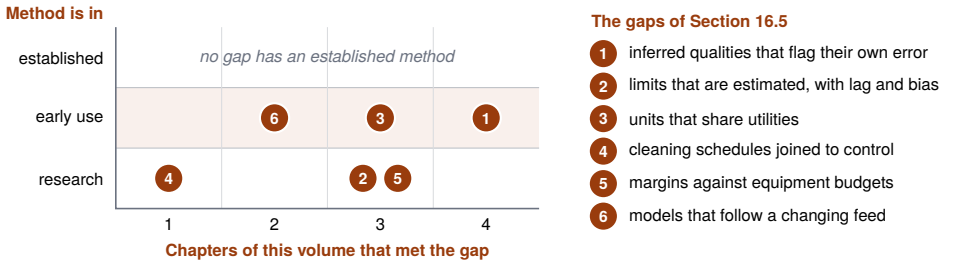


Figure 16.4 The six gaps of this section. Across is the number of chapters that each item names as evidence: four for gap 1 (Chapters 5, 9, 13 and 15), three for gap 2 (5, 6 and 7), three for gap 3 (6, 12 and 14), one for gap 4 (3), three for gap 5 (8, 9 and 11) and two for gap 6 (2 and 4). Up is the maturity of the method available, which is the editor's reading of the sources cited in each item and not a measurement. Notice that the top band is empty.

3. *Coordination of units that share utilities.* The fuel gas header is an integrator that nobody schedules, the hydrogen plant takes its rate from the pressure of the header it supplies, and the cracker gives back 23 t/h of feed on a warm afternoon. A coordinator that knows which unit's limit binds across the refinery, and moves rate there, remains an engineering problem. A 2022 review of real-time optimization counts the conflict between the optimum of one unit and that of the whole operation among the reasons the method is used less than expected [12]. The gap is hard because a shared header couples the units within seconds or minutes, while the layer that knows the prices acts each hour or each day (Figure 1.10). Progress would be a coordinating layer that reads the binding limit of each unit and moves rate between units within minutes.

4. *Cleaning schedules joined to control.* On the preheat train the phasing of two outages mattered more than any tuning. The cleaning schedule, the override's target and the crude slate are still decided by different people on different time scales. In the literature the cleaning schedule and the flow control of a preheat train have usually been treated independently and solved by heuristics, and a closed-loop scheme that optimizes both has been shown on a simulated industrial train [17]. The gap is hard because the decisions span weeks to years and cleaning crews are contracted with long notice, so operators ask for a schedule that does not change at every update [17]. Progress would be an override target and a cleaning date computed from one fouling model and proven on a plant.

5. *Margins chosen against equipment budgets.* The hydrotreater's 7.5 wppm setpoint came from a sweep over 31 seeds against a catalyst budget. The reformer and the coker paid for the same kind of margin in liquid yield and in the life of the heater coils. Any back-off from an active constraint is an economic loss [2], and how large the back-off must be was named in 2000 as an area in clear need of more research [21]. The gap is hard because the margin depends on disturbances that have not yet happened and on a price for catalyst or tube life that no instrument reports. A method that sets each quality back-off from the analyzer's noise, the feed's variability and the cost of the margin in catalyst, yield or tube life would replace that sweep.

6. *Models that follow a changing feed.* On the desalter no mixing pressure drop met the water specification on a heavier crude at the design demulsifier dose, and the crude tower's predictive controller lost the heavier crude on every quality. A review of catalytic cracking names light, tight oils as a challenge because their low carbon residue makes the coke deposition, and with it the heat balance of the unit, harder to control [22]. The gap is hard because new predictive controllers are still built mainly on linear models from plant tests, which are costly and slow to repeat [2]. A model that tracks the crude slate, and says when it no longer can, is needed before predictive control is trusted on a varying feed.

⚠ OPEN CHALLENGE An Analyzer That Reports Its Own Error

Gap 1 is the one most chapters met. A review of soft sensors found that there is often no objective measure of their quality, so the judgment rests on a person who compares the predictions with the correct values by eye [16]. The difficulty is that a drift of the instrument and a drift of the process are hard to tell apart in the data, can progress very slowly and take expert knowledge to separate [16]. A spectroscopic analyzer adds its dependence on a good reference method [14]. Progress would be an inferred quality delivered with its own error bound and checked against each laboratory result without a person.

16.5.1 Gaps Common to Agent-Era Practice

The final chapter of each volume lists gaps specific to its topic. Seven more are common to every topic in the series. In the editor's view none of them is closed.

1. *Verifying agent-written control code and models before they reach a plant.* Tests show that a controller behaves on the cases tested. For controllers learned from data, guarantees of stability and constraint satisfaction rest on assumptions about the model and its uncertainty, and open problems remain [23]. In the editor's view, code written by an agent has no comparable body of method yet. Practice needs a routine path from an agent's draft to code a plant can trust: independent derivation, tests of what the code reads and clamps, and trials with hardware in the loop.

2. *Trustworthy models from sparse or poor plant data.* Plant records are short, taken in closed loop, noisy and marked by unmeasured disturbances. A fitting routine returns a model from any record. What is missing is a dependable statement, delivered with the model, of where it can be trusted, and a method that says so when the data cannot identify a parameter.

3. *Safety cases and standards for agents in commissioning and operation.* Functional safety standards define a lifecycle of specified requirements, verification and assigned human roles [19]. In the editor's view, how an agent that writes configuration, runs commissioning checks or advises an operator is qualified, limited and audited under them is not yet settled.

4. *Security of agent access to control systems.* An agent with credentials to a control system is a new path into it. Industrial security standards state the requirements a control system must meet at each security level [24]. An agent that reads documents, messages or web pages can also be steered by text placed there. In the editor's view, least privilege, read-only access by default and human approval of every write are a starting point and not a solution.

5. *Benchmarks and reproducibility.* An agent asked twice may deliver two different controllers, and a result reported without its plant, seed and scoring cannot be checked. Control needs shared benchmark plants with fixed disturbances and published measures, on which agent-built and conventional designs are compared like for like. The apps of this series, with a default seed and an exported record of every run, are a start.

6. *Transfer from simulation to hardware.* A controller developed on a simulator inherits what the simulator omits. The robotics literature documents the gap between simulated and real performance and the methods used to narrow it, such as randomizing the simulated plant and adapting on the real one [25]. Process plants add slow dynamics, rare upsets and little freedom to experiment, so each transfer still needs a staged plan from the simulator to a bench rig to the unit.

7. *Keeping engineers' judgment sharp when agents do the routine work.* Automation removes the routine work through which skill is built and leaves the person the hard cases, an irony described for process operators in 1983 [26]. The same now applies to engineers. An engineer who has never tuned a loop by hand or closed a balance on paper is poorly placed to review an agent that has. Deliberate practice, of the kind the exercises in this book provide, has to be designed into education and into work.

KEY POINTS

- A unit gives up oil flow only while a limit binds, and only as much as the limit asks.
- A guard protects only as well as the signal it reads.
- The setpoint of a quality loop is an economic choice.
- Feedback holds the reading. A second source is needed to hold the product.
- A predictive controller is compared with the older structure at the same target and the same ceiling.

References

[1] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

[2] S. Skogestad, "Advanced control using decomposition and simple elements," *Annual Reviews in Control*, vol. 56, p. 100903, 2023.

[3] I. Grossmann, "Enterprise-wide optimization: A new frontier in process systems engineering," *AIChE Journal*, vol. 51, no. 7, pp. 1846–1857, 2005.

[4] N. K. Shah, Z. Li, and M. G. Ierapetritou, "Petroleum refining operations: Key issues, advances, and opportunities," *Industrial & Engineering Chemistry Research*, vol. 50, no. 3, pp. 1161–1170, 2011.

[5] A. Singh, J. F. Forbes, P. J. Vermeer, and S. S. Woo, "Model-based real-time optimization of automotive gasoline blending operations," *Journal of Process Control*, vol. 10, no. 1, pp. 43–58, 2000.

[6] M. G. Forbes, R. S. Patwardhan, H. Hamadah, and R. B. Gopaluni, "Model predictive control in industry: Challenges and opportunities," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 531–538, 2015.

[7] F. A. A. Souza, R. Araújo, and J. Mendes, "Review of soft sensor methods for regression applications," *Chemometrics and Intelligent Laboratory Systems*, vol. 152, pp. 69–79, 2016.

[8] S. J. Qin and L. H. Chiang, "Advances and opportunities in machine learning for process data analytics," *Computers & Chemical Engineering*, vol. 126, pp. 465–473, 2019.

[9] F. Smaïli, V. S. Vassiliadis, and D. I. Wilson, "Mitigation of fouling in refinery heat exchanger networks by optimal management of cleaning," *Energy & Fuels*, vol. 15, no. 5, pp. 1038–1056, 2001.

[10] D. I. Wilson, E. M. Ishiyama, and G. T. Polley, "Twenty years of Ebert and Panchal—what next?" *Heat Transfer Engineering*, vol. 38, no. 7–8, pp. 669–680, 2017.

[11] S. Bezergianni, A. Dimitriadis, O. Kikhtyanin, and D. Kubička, "Refinery co-processing of renewable feeds," *Progress in Energy and Combustion Science*, vol. 68, pp. 29–64, 2018.

[12] D. Krishnamoorthy and S. Skogestad, "Real-time optimization as a feedback control problem – a review," *Computers & Chemical Engineering*, vol. 161, p. 107723, 2022.

[13] M. Chèbre, Y. Creff, and N. Petit, "Feedback control and optimization for the production of commercial fuels by blending," *Journal of Process Control*, vol. 20, no. 4, pp. 441–451, 2010.

[14] C. Pasquini, "Near infrared spectroscopy: fundamentals, practical aspects and analytical applications," *Journal of the Brazilian Chemical Society*, vol. 14, no. 2, pp. 198–219, 2003.

[15] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson, "Guide to operational technology (OT) security," National Institute of Standards and Technology, Gaithersburg, MD, Special Publication 800-82r3, 2023.

[16] P. Kadlec, B. Gabrys, and S. Strandt, "Data-driven soft sensors in the process industry," *Computers & Chemical Engineering*, vol. 33, no. 4, pp. 795–814, 2009.

[17] F. Lozano Santamaria and S. Macchietto, "Stability of optimal closed-loop cleaning scheduling and control with application to heat exchanger networks under fouling," *Processes*, vol. 8, no. 12, p. 1623, 2020.

[18] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, W. X. Zhao, Z. Wei, and J. Wen, "A survey on large language model based autonomous agents," *Frontiers of Computer Science*, vol. 18, no. 6, p. 186345, 2024.

[19] International Electrotechnical Commission, "Functional safety – Safety instrumented systems for the process industry sector – Part 1: Framework, definitions, system, hardware and application programming requirements," IEC 61511-1:2016, Edition 2.0, 2016.

[20] U.S. Chemical Safety and Hazard Investigation Board, "ExxonMobil Torrance refinery electrostatic precipitator explosion, Torrance, California, incident date February 18, 2015," U.S. Chemical Safety and Hazard Investigation Board, Washington, DC, Investigation Report 2015-02-I-CA, 2017.

[21] S. Skogestad, "Plantwide control: The search for the self-optimizing control structure," *Journal of Process Control*, vol. 10, no. 5, pp. 487–507, 2000.

[22] E. T. C. Vogt and B. M. Weckhuysen, "Fluid catalytic cracking: recent developments on the grand old lady of zeolite catalysis," *Chemical Society Reviews*, vol. 44, no. 20, pp. 7342–7370, 2015.

[23] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 411–444, 2022.

[24] International Electrotechnical Commission, "Industrial communication networks – Network and system security – Part 3-3: System security requirements and security levels," IEC 62443-3-3:2013, Edition 1.0, 2013.

[25] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: A survey," in *Proc. 2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 737–744.

[26] L. Bainbridge, "Ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775–779, 1983.

16.6 Exercises

16.1 Discussion · 8 min

An agent proposes raising the valve target of the preheat override in Chapter 3 from 95 to 99 % of the fuel valve, to win back crude. Use Table 3.8 for the exchanger outage (scenarios 5 and 6) and the harder outage of scenarios 7 and 13. Say what the proposal gains, what it costs, and what you would ask for before accepting it.

16.2 Discussion · 8 min

The override of Chapter 7 tripped the unit on a feed cut when it read the raw bed rise and held it when the inlet was lagged. A new unit is to protect its beds with an override written by an agent. Say what you would test about the signal the override reads, and relate the hydrocracker's lesson to the vacuum maximizer of Chapter 5 and to the cases of Section 16.4.1.

16.3 Discussion · 7 min

On the graded day of Chapter 8 an agent proposes a sulfur setpoint of 6.5 wppm instead of 7.5, because the worst reading then falls to 7.94 wppm. Discuss the proposal with Table 8.5 and the trade-offs of Chapter 9 and Chapter 11, and say how a quality setpoint should be chosen.

16.4 Discussion · 8 min

An agent reports that a predictive controller on the hydrocracker raised the mean conversion through a feed cut from 0.5427 to 0.5501 and held the bed rise at 27.5 °C. Using Table 7.5, the preheat comparison of Table 3.8 and Section 16.4.1, say what you would ask before crediting the gain to prediction.

16.5 Discussion · 8 min

Three chapters held a quality from an instrument that was wrong without saying so: the tail gas analyzer of Chapter 13, the octane analyzer of Chapter 9 and the blend analyzer of Chapter 15. For each, say what the loop showed, what the product did and which second source caught it. Then propose what a monitoring agent should watch on such a unit.

16.6 Discussion · 8 min

A refinery wants a coordinator, proposed by an agent, that raises the crude rate until the first unit reaches a limit. Using Chapter 3, Chapter 5, Chapter 6 and Chapter 14, discuss what the coordinator must know about each limit, how the limits move during a day, and which decisions should stay with the operators.

Answers to Selected Exercises

Short answers are given for the odd-numbered exercises. Numeric answers from an app are for its default seed and settings unless the exercise says otherwise.

Chapter 1. Introduction

1.1 (a) 270.6 °C. (b) 44.9 MW, or 74.8 % of full firing.

1.3 (a) 222, 342 and 384 °C, each 12 K up. (b) Cut the diesel draw by 14.4 t/h.

1.5 (a) 363.35 °C. (b) It makes the diesel heavier, raising the cut by about 2.7 K.

1.7 (a) $L = 0.6$ and $A = 2.5$. (b) $L = 0.8$ and $A = 5$, stable only while $\varepsilon_u < 0.5$.

1.9 (a) About 621 t/h. (b) At 653.6 t/h the preheat is 239.4 °C and the duty 57.0 MW, the target.

Chapter 2. Crude Desalter Mixing Optimum and Rag Layer

2.1 (b). Contact η_c rises with Δp, but the drop size falls as $\Delta p^{-0.6}$, so more contacted water is carried over below the cut size.

2.3 (d). At 30 ppm the least salt with the BS&W at 0.2 % or less is 1.005 PTB at 0.87 bar.

2.5 (a) 750 mm, 1250 mm, 700 mm. (b) 632.5 mm, 1367.5 mm, 582.5 mm. (c) 564 mm of rag, a rag top at 1564 mm with a 386 mm gap, and a setpoint near 718 mm to keep the rag top where it was.

2.7 (a) −0.2 mm after the setpoint step and 55.3 mm after the wash step, with the valve 16.7 points further open. (b) −1.0 mm and an overshoot of 28 mm. (c) 39 mm. The layer needs no lasting change of brine after a setpoint step, but does after the wash step.

2.9 (a) 493.3 min. (b) 0.84 PTB, 0.213 %, missed on the BS&W. (c) Both pass. The fixed 30 ppm averages 30.0 ppm, the suggested settings 40.0 ppm, a third more chemical.

2.11 (a) About 8 m^3, since $26 + 50 - 60 = 16\,m^3/h$ for half an hour. (b) The tight tuning, which keeps the rag top lower and the current under the 2.0 A alarm but sends the slug straight to the wastewater plant at the valve's full flow.

Chapter 3. Crude Preheat Fouling and Throughput Override

3.1 (a). The loop moves the fuel valve to hold the temperature, so the lost duty appears only in the valve position.

3.3 (c). Below the point where the constraint binds a tighter target buys no temperature and costs tonnes.

3.5 (a) 363.05 °C, 0.74 MW, 81.3 %. (b) 1.644 °C per percent. (c) $3.65 \times 1.644 = 6.0$ °C, and the true inlet is about 354 °C.

3.7 (a) 97.5 %, 76.7 TJ, 239.6 h. (b) 215.2 h, and 352.1 °C at the end. (c) The faster crude fouls toward a lower asymptote, 0.514 with a time constant of 400 h, but each tonne more needs more firing, so the valve runs out first.

3.9 (a) 559.1 °C h, 240 t/h, 262.6 °C, 199 056 t. (b) 13.9 °C h, 290.5 t/h, 344.9 °C, 200 315 t. (c) Eight hours, then none: with no overlap the override never reaches its floor.

3.11 (a) $R_f = 0.088$, 21 % of the threshold. (b) About 532 h, 12 % shorter. (c) Each run starts dirtier and ends sooner, so the period shrinks and the phase between two exchangers drifts.

Chapter 4. Crude Atmospheric Tower Degrees of Freedom

4.1 (d). The end point is $T_1 + \sigma_1$ and the flash point is $T_1 - \sigma_1 - 112$ °C, so one reflux fixes both.

4.3 (b). The request is multiplied by $1 - \min(0.6, (\Phi_{min} - \Phi_m)/1.5)$, a proportional cut, and the wash zone is still dry for 23 min.

4.5 (a) 337.6 °C, moved 17.6 °C by the kerosene step. (b) $\phi_v = 0.6315$, $\Phi = 3.28$ vol%, 0.185 vol% per °C and about 6.9 °C. (c) 1.82 °C and 0.75 MW per megawatt, so 25.5 °C and 10.5 MW for the step.

4.7 (a) 184.0 °C, 53.9 °C and 150.3 t/h, then 176.0 °C, 48.8 °C and 170.1 t/h, a ratio of 0.63. (b) 188.0 °C at 170.4 min, off specification from 165.6 to 184.8 min. (c) 184.3 °C, 54.0 °C and 149.6 t/h.

4.9 (a) 3.81 t/h. (b) 6.39 t/h for a switch at 10.0 min, 5.97 to 6.57 t/h between 9.5 and 10.5 min. (c) 4.99 vol% and 12.02 °C against 5.01 vol% and 12.05 °C, within 0.3 percent.

4.11 (a) A real crude switch changes the shape of the curve and its sulfur, acid number and metals, while the model only shifts and tilts one assay, and the properties are correlations on the sliced assay rather than laboratory methods. (b) Kerosene sulfur or smoke point, and diesel cetane number.

Chapter 5. Vacuum Tower Cut Points and the Ejector Break

5.1 (d). The HVGO rises from 130.3 to 155.1 t/h and the 95 % point from 545.8 to 570.5 °C, for 0.7 MW more firing.

5.3 (b). The physical loop gain is far below one. The runaway goes through a controller that does not know the sign of its gain has changed.

5.5 (a) $\kappa = 0.729$ mbar per kg, $\tau = 0.55$ min. (b) 71.0 mbar a, 964 kg/h, recovery below 844 kg/h. (c) About 700 and 580 kg/h. The app's values are lower because more pressure makes more cracked gas.

5.7 (a) 130.3 t/h, 545.8 °C. (b) 155.1 t/h (± 0.4 t/h), 570.5 °C (± 0.2 °C), 0.87 ppm (± 0.02 ppm), a move of 24.7 K, more than the line alone. (c) 12.8 min after the change.

5.9 (a) 1097 t, ejector break at 352 min, fails. (b) 1154 t, no break, passes. (c) 1024 t, no break, fails on the 1140 t target: it gives up lift all day to cover the hour when the water is warm.

5.11 (a) About 709 kg/h at 32 °C and 637 kg/h at 33 °C, so the break comes between, at 32.4 °C in the app. (b) A cooler outlet, which cuts the cracked gas and the HVGO, and less stripping steam, which cools the condensers but biases the inferred 95 % point.

Chapter 6. Cat Cracker Heat Balance and Selector Control

6.1 (c). In Equation (6.11) the circulation term acts at once, and the coke that pays for it arrives after a pass around the loop.

6.3 (a). External reset, Equation (6.26), gives each loser the selected output plus the slack as its upper limit.

6.5 (a) 230.4, 213.4 and 232.8 t/h, a loss of 7.4 %. (b) 1.35 tonnes of feed per tonne of air capacity, so 22.9 t/h lost and 3.3 t/h gained, as measured.

6.7 (a) 2.11 °C and −0.46 °C. (b) 7.05 °C and 5.09 °C, and the carbon peaks at 0.0577 % by weight 13 min after the move. (c) 29 % and 98 %. The slow part is the regenerator.

6.9 (a) 284.5 t/h, 230.4 t/h, \$17,235 per hour; then 261.5, 213.4 and \$15,986; then 287.8, 232.8 and \$17,408, with the air blower binding throughout. (b) 236.7 t/h at 92.0 min, and the blower approach peaks at 143.7 % at 90.1 min. (c) $23.0/17.0 = 1.35$. The undershoot is the proportional kick of the air blower constraint controller.

6.11 (a) $234 \times 8760 \approx$ \$2.05 million a year. (b) 72 % recovered. \$747 per hour, about \$6.5 million a year, remains, and it lies in severity, which only an optimizer or coordinating layer moves.

Chapter 7. Hydrocracker Runaway and Layered Protection

7.1 (b). Conversion per bed goes as the inverse space velocity: the rise goes from 23.9 to 34.1 °C and the index from 0.538 to 0.756.

7.3 (d). Bed 1 has no quench above it, so nothing fast acts on it once its index passes one.

7.5 (a) 3.38 t/h. (b) −6.57 °C per t/h, so about 0.30 t/h. (c) 2.0 °C and 0.74 °C above the reaction rise, while the compensated rise reads the reaction rise.

7.7 (a) 404.86 to 408.45 °C, a gain of 3.59 per degree. (b) The tile reads 3.9. The step gain is a secant across a degree in which the index rises from 0.538 to 0.592, so it lies between the tangent gains at 381 and 382 °C.

7.9 (a) Bed 1 at 379.50 °C and beds 2 to 4 at 381.00 °C, a largest rise of 27.50 °C, no override, 246 t. (b) All inlets at 379.79 °C, 29.92 °C, override from 21.10 min, 243 t. (c) From bed 1's own lower target: only bed 1 gives up temperature.

7.11 (a) Under one degree, 0.87 °C; the steady rise would reach 34.1 °C, past the 33 °C alarm. (b) Lower the inlet setpoint with the feed, to 379 °C or below, and add a feedforward from the charge rate to the inlet setpoint.

Chapter 8. ULSD Hydrotreater With a Late Sulfur Analyzer

8.1 (a). Wrong-way behavior: the outlet climbs from 368.69 to 369.79 °C, peaks at 31.5 min and falls only after 35 min.

8.3 (c). The inlet climbs to 349.6 °C and bed 1's rise to 29.5 K, half a kelvin under its limit.

8.5 (a) $n = 1 + 1/m = 1.65$. With $m = 1$, $dX/d\tau = -\bar{k}X^2$. (b) X is fixed by the reactor, so the product is $8.00 \times 1.5 = 12.0$ wppm. (c) The extra sulfur releases heat in bed 1 and the hotter bed removes more. The absolute form makes a sourer feed come out cleaner.

8.7 (a) 5.27 K at 30 min and 12.90 K at 400 min, 2.39 times the direct 5.39 K. (b) 2.56 K in 2 min, about four-fifths of it in the first minute.

8.9 (a) Fails the readings, the mean and the rise: the fouling shuts the bypass and the fixed heater lets the inlet fall. (b) Fails the readings and the mean. (c) Passes, with 9.11 wppm worst and 1.21 days per day.

8.11 It removes the miss at seed 26 and lowers the default worst reading from 9.70 to 9.11 wppm, for 1.21 instead of 1.17 days of catalyst per day. At 6.5 wppm the life (1.30) and rise (30.05 K) limits fail. The analyzer's delay, the feed switch and the noise set the margin.

Chapter 9. Reformer Interheaters and Inferred Octane

9.1 (a). Bed 1's drop grows from 70.4 to 79.7 K, so it keeps 9.3 K of the 15.7 K rise.

9.3 (c). Above it the regenerated catalyst carries coke into bed 1, and each kelvin buys less.

9.5 (a) 8.0 h. (b) About 238 kg/h, 60 % of the capacity. (c) 4.95 wt%. (d) 1.17 wt%.

9.7 (a) The outlet rises from 500.0 to about 515.7 °C, the drop from 70.39 to 79.65 K: bed 1 keeps about 9.3 K and passes about 6.3 K. (b) The exchanger adds about 6 K, slowly, as the warmer gas works through all four beds.

9.9 (a) Only the octane band is missed, 84.7 % against 95 %. (b) The run passes with 100.0 % in the band, 606 °C, 364 kg/h, H_2/HC 2.90 and 91.16 wt%. The inferred octane rides through the NIR shift at 16 h.

9.11 (a) $0.5 \times 0.75 = 0.375$ points, about 0.56 t/h of reformate turned into gas and hydrogen. (b) Yes, because a giveaway cost is a difference along the trade curve, which depends on its slope and not on its level.

Chapter 10. Sulfuric Acid Alkylation on Three Clocks

10.1 (d). The purity drifts from 89.305 to 86.094 % a day later, and the loop's I/O is computed with 89.3 %.

10.3 (b). The linearization fixes the ASO's slope at its design value of 14.96 kg/h per point, and the slope grows as the strength falls.

10.5 (a) 283 kg/h and 503 kg/h. (b) 3211 kg/h. (c) 0.35 and 0.58 wt% an hour.

10.7 (a) 91.45 to 92.31 wt%, 3205 to 6472 kg/h. (b) 91.92 to 92.03 wt%, 3774 to 5774 kg/h. (c) Each report moves the loop by a whole period of integral action on an error that is hours old.

10.9 (a) 94.02, 93.52, 89.08 wt%, \$18,738: the average RON, the lowest RON and the strength fail. (b) 94.91, 94.65, 91.45 wt%, \$21,995: pass. (c) 94.95, 94.66, 91.37 wt%, \$24,250: a pass \$250 under the cap.

10.11 (a) $0.6(525.53 - 289.09) + 85.46 + 1.00 \approx 228\,\text{kmol/h}$. (b) No. The low-I/O trip stops the run after 2 min below 5. A real setting comes from the unit's safety review.

Chapter 11. Delayed Coker Drum Switch Control

11.1 (c). A shell below its vapor holds back vapor, and the gas made with it, at the rate of Equation (11.2).

11.3 (a). The foam rides on the bed at a steady 9.05 m, and the bed climbs 1.12 m an hour until the foam reaches the vapor line.

11.5 (a) 0.054 t/h per K, 55.6 min. (b) 5.38 t/h and 16.9 t/h, 3.14 times. (c) 15.6 t and 5.0 t, and in four hours 15.4 t and 4.9 t, the values the chapter measures.

11.7 (a) 0.940 bar g at 129.2 min, 73.4 % at 137.4 min. (b) 0.988 to 1.014 bar g, 70.8 % at 122.0 min. (c) The pressure moved first on the loops and the compressor first with the feedforward, which knew the switch time and the new shell's temperature.

11.9 (a) 9.7 °C, 0.186 bar, 18.9 kg, missed. (b) 1.0 °C, 0.024 bar, 6.8 kg, met. (c) The feedforward fixes the pressure and the HCGO, the outage loop the antifoam, and both failures began at the cold switch at 28 h.

11.11 (a) The warm-up gets what the twelve off-line hours leave, and the new shells hold back 6.0, 9.8 and 16.1 t/h. (b) Only the feed rate. The decoking steps set the 12 h floor.

Chapter 12. Steam Methane Reformer and PSA Hydrogen Plant

12.1 (b). The lead-lag acts on the demand: the feed follows it through a 1.5 min lag going up while the steam setpoint takes it at once.

12.3 (d). The swing grows from 0.221 to 0.390 bar: the drum's time constant is short against a pulse that lasts most of a step.

12.5 (a) 0.0968 bar a minute. (b) 5.2 min, and 0.194 bar a minute. (c) 0.52 min and 0.13 min, both short against the step.

12.7 (a) 2.73, and 2.729 at 2.83 min. (b) 2.99. (c) 21 477 Nm^3/h against 19 527 Nm^3/h: the lead-lag holds the feed back while the steam catches up.

12.9 (a) Met: 0.394 bar, 99.929 %, 2.886, 916.7 °C. (b) Missed on S/C, 2.720 at 63.8 min, just after the demand step. The firm header loop moves the rate fast enough to outrun the steam ramp.

12.11 (a) 1.85 %, about 400 Nm^3/h. (b) About 0.82 t/h of carbon dioxide. A fuel report sees the makeup fall and records a saving, while the plant burns its own hydrogen and uses more natural gas in all.

Chapter 13. Claus Sulfur Recovery Air Demand Control

13.1 (b). $D = y_{H_2S} - 2y_{SO_2}$ is positive when the furnace is short of air and negative when it has too much.

13.3 (d). The plant runs 1.4 % rich with the stack at 9.60 t/d, over the permit, and only the stack alarm reports it.

13.5 (a) 17.33 kmol/min. (b) 18.02 kmol/min, up 3.96 %, so the ratio settles at 1.040. (c) $R = 0.9165$ against 1.200 needed, and about 89.76 % recovery.

13.7 (a) 96.64, 10.53, 0.366; 96.64, 10.53, −0.712; 97.62, 7.46, 0.000. (b) About −36 (lean −36.6, rich −35.6). (c) After both off-design steps; it clears near 100.7 min.

13.9 (a) About 24.1 min; 0.9165, 89.76 %, 1226.9. (b) 7.19, 1.032 at 23.70 min, 7.73 t/d, alarm latched. (c) The trim answers an error the feedforward has already removed.

13.11 (a) Equilibrium favors low temperature in every bed, but only bed 1 hydrolyzes carbonyl sulfide and carbon disulfide, and beds 2 and 3 recover what it gives up. (b) Wet catalyst loses activity over hours and recovers in days, while the benefit of cold is immediate. Bed 1 hotter, bed 3 colder.

Chapter 14. Refinery Fuel Gas Header and Flare

14.1 (a). The header is a pure integrator until a heater starves, and the starved heaters close the balance near 1.74 bar g.

14.3 (c). The plan never opens both valves, but the stems follow it with lags of 5 and 3 s.

14.5 (a) 54.69 MJ/Nm^3. (b) HHV $= 40.96\,\mathrm{MJ/Nm^3}$, $SG = 0.5921$, $W = 53.24\,\mathrm{MJ/Nm^3}$. The heating value falls 7.11 % and the Wobbe index 2.66 %. (c) -0.264, about −2.64 % at a tenth.

14.7 (a) 14.38 min, 50 Nm^3 flared and 362 Nm^3 of makeup. (b) 0.00 min, 0 Nm^3 and 310 Nm^3. (c) The overlap bought about 52 Nm^3 more natural gas, close to the 50 Nm^3 it flared, while the header stayed within a few hundredths of a bar of setpoint in both runs.

14.9 (a) Target met: header within 0.154 bar, 6 Nm^3 flared, 0.00 min of import and flare, outlet within 0.34 °C. (b) Only the pressure fails, 0.292 bar from setpoint at about 181 min, in the FCC upset. The So far tile reads “missed: pressure”.

14.11 Many plants route the header’s flare valve into the flare header, where a compressor with spare capacity would take back what an overlapped split range sends, so the overlap would show up as recycled gas and compressor load rather than flaring. Real flares do not burn completely: measured flares destroyed 91.1 % of their methane against the 98 % generally assumed, so the tile understates the emissions.

Chapter 15. Gasoline Blending to a Tank-Average Octane

15.1 (d). $D = \Delta/f$, and $0.25/0.05 = 5$.

15.3 (b). The loop is almost pure delay, and at this tuning Equation (15.16) has roots outside the unit circle.

15.5 (a) 125.9 and 98.4. (b) RON 94.84, MON 85.09, AKI 89.96. (c) 8.31 psi, over the limit by 0.51 psi.

15.7 (a) 9.08 min (±0.1), 63.5 kPa or 9.21 psi (±0.1). (b) 6.67 min (±0.1). (c) Each report describes blend made 3.0 min earlier, and the tank, nearly empty, follows the header at once.

15.9 (a) Tank 86.85 (±0.01), NIR 87.15, header 86.50, NIR error 0.77: off specification. (b) 87.03 (±0.01). (c) The bias follows the NIR, and only the laboratory measures the blend independently.

15.11 (a) 7.84 psi, a lift of 0.48 psi, about half the model’s E10 lift. (b) Above, so the RVP margin is smaller than the optimizer believes. The RVP channel of the analyzer and its bias update would see it.

Chapter 16. Summary and Outlook

16.1 A good answer reads both sides of the table. At 99 % the override leaves 393 instead of 555 t unprocessed on the first outage and 1003 instead of 1584 t on the second, a gain of 162 and 581 t. The cost is control of the column inlet. The IAE rises from 1.27 to 16.8 and from 59.4 to 427.2 °C h, and the valve sits on its stop for 185 and 707 min. The answer asks for the same target with a charge feedforward, which keeps most of the tonnes at a small error (419 and 1525 t, IAE 1.24 and 4.33), and for the temperature result, not only the tonnes.

16.3 A good answer notes that the lower setpoint improves the quality reading and fails the equipment budgets, catalyst use 1.30 days a day and a rise of 30.05 K. At 7.5 wppm the day passes at 1.21 and 29.2 K, and all 62 trials over 31 seeds and two steps pass. At 8 wppm seed 26 put two readings over specification. On the reformer 10 K buys 2.01 RON for 1.51 points of yield, and on the coker the same 10 K trades 1.9 points of coke for coils that coke 2.2 times as fast. The setpoint is the smallest margin that covers the noise and the feeds.

16.5 A good answer notes that each loop held its reading. The sulfur plant’s air demand stayed on target while the stack went to 9.60 t/d over a 9.0 t/d permit, caught only by the stack permit alarm. The reformer’s trim on the shifted analyzer let the true octane fall to 94.24 and stay under 94.7 from 54 to 248 min, and a trim on the inferred octane never fell under 94.97. The blender certified at 87.06 with the laboratory correction and missed at 86.45 without it. A monitor should compare each analyzer with an independent source, such as a model, a laboratory sample or a downstream measurement, and alarm on their disagreement, not on the loop error.

Glossary

Each term is given as this volume uses it, in one plain sentence, followed by the chapter where it is first used in earnest. The two indexes give every page.

acid gas Gas rich in hydrogen sulfide that the treating units strip from fuels and send to sulfur recovery. *Chapter 13*.

acid strength The share of sulfuric acid in the acid of an alkylation unit, which falls as water and oily by-products build up. *Chapter 10*.

adiabatic reactor A reactor with no cooling, so the heat of the reaction stays in the stream and changes its temperature. *Chapter 1*.

afterburn Burning of carbon monoxide above the catalyst bed of a regenerator, where no catalyst is left to take up the heat. *Chapter 6*.

air demand How far the air fed to a sulfur plant is from what the reaction needs, as read by the tail gas analyzer. *Chapter 13*.

alkylation Joining light olefins to isobutane with an acid catalyst to make a high-octane gasoline component. *Chapter 10*.

analyzer An instrument that takes a sample of a stream and reports its composition or a property, usually some minutes later. *Chapter 1*.

analyzer delay The age of an analyzer reading: the time in the sample line plus the time the analysis takes. *Chapter 1*.

anti-knock index (AKI) The octane number posted at the pump, the average of the research and motor octane numbers. *Chapter 15*.

API gravity A scale for the density of crude oil on which a lighter oil has a higher number. *Chapter 2*.

atmospheric equivalent temperature The temperature at which a liquid boiling under vacuum would boil at atmospheric pressure. *Chapter 5*.

back-off The margin kept between a setpoint and a limit so that disturbances do not push the plant over the limit. *Chapter 8*.

basic sediment and water (BS&W) The water and solids left in crude oil, given as a percentage of the oil. *Chapter 2*.

bias update A correction added to a calculated property each time a laboratory or analyzer result shows how far off it was. *Chapter 1*.

bumpless transfer Switching a controller between manual and automatic without a jump in its output. *Chapter 3*.

catalyst activity How well a catalyst still works compared with fresh catalyst, which falls as coke or metals cover it. *Chapter 6*.

Claus process The process that burns a third of the hydrogen sulfide in acid gas and reacts the rest with the product to make sulfur. *Chapter 13*.

coil outlet temperature (COT) The temperature of the oil where it leaves the tubes of a fired heater. *Chapter 1*.

coke burn Burning the coke off a catalyst with air, which restores the catalyst and releases heat. *Chapter 6*.

coke drum The tall vessel of a delayed coker in which hot residue cracks and leaves solid coke behind. *Chapter 11*.

Conradson carbon residue (CCR) A laboratory measure of how much coke an oil tends to form when it is heated. *Chapter 11*.

constraint control Control that holds a unit at a limit, or that takes over a valve when a limit is reached. *Chapter 3*.

control hierarchy The layers of refinery control, from fast loops on flows and temperatures up to slow optimization of targets. *Chapter 1*.

crude oil Petroleum as it comes from the ground, a mixture of many hydrocarbons with some water, salt and sulfur. *Chapter 2*.

cumulative average The average of everything put into a tank so far, on which a blended product is certified. *Chapter 1*.

cut point The boiling temperature at which one product of a tower ends and the next begins. *Chapter 1*.

dead band A range of the controller output in which neither of two valves opens. *Chapter 14*.

dead time The wait between a change and the first sign of its effect at the measurement. *Chapter 13*.

degrees of freedom The number of things in a unit that can be set independently of one another. *Chapter 1*.

delayed coker A unit that heats vacuum residue and lets it crack in a drum, which keeps the coke and sends the vapor to a fractionator. *Chapter 11*.

demulsifier A chemical added in small doses to help the water drops in oil merge and settle. *Chapter 2*.

desalter A vessel that washes crude oil with water

and uses an electric field to settle the salty water out. *Chapter 2*.

distributed control system (DCS) The plant's network of controllers and operator screens that runs the loops of a unit. *Chapter 1*.

electrostatic coalescence Merging small water drops in oil into larger ones that settle, by means of a high-voltage field. *Chapter 2*.

exchanger fouling Deposits on the walls of a heat exchanger that cut the heat it transfers. *Chapter 1*.

feed and effluent exchanger A heat exchanger in which the hot stream leaving a reactor warms the feed that enters it. *Chapter 1*.

feedforward Moving a valve as soon as a measured disturbance changes, before the controlled variable shows the effect. *Chapter 1*.

fired heater A furnace that burns fuel around tubes to heat the process stream flowing inside them. *Chapter 1*.

flare A stack at which surplus gas is burned safely. *Chapter 14*.

flash zone The part of a tower where the heated feed enters and separates into vapor and liquid. *Chapter 1*.

fluid catalytic cracking (FCC) Cracking heavy oil over a hot powdered catalyst that flows like a liquid between a reactor and a regenerator. *Chapter 6*.

fuel gas The mixture of light gases from the units that a refinery burns in its own heaters. *Chapter 14*.

give-away Product quality better than the specification asks for, which is paid for and not needed. *Chapter 15*.

heat integration Using the heat of hot products to warm cold feed, which saves fuel and ties units together. *Chapter 1*.

hydrocracking Cracking heavy oil over a catalyst under hydrogen at high pressure, a reaction that releases heat. *Chapter 7*.

hydrodesulfurization Removing sulfur from a fuel by reacting it with hydrogen over a catalyst. *Chapter 8*.

hydrotreater A reactor that uses hydrogen to remove sulfur and other impurities from a fuel. *Chapter 8*.

import and flare Buying natural gas and burning surplus gas at the flare at the same moment, a loss on both sides. *Chapter 14*.

inferential control Control of a property that is calculated from other measurements, because it cannot be measured directly or soon enough. *Chapter 1*.

integrating process A process whose output keeps rising or falling while inflow and outflow differ, like the level in a tank. *Chapter 2*.

interface level The height at which the oil layer meets the water layer in a vessel. *Chapter 2*.

inverse response A response that first moves the wrong way and then heads for its final value, also called wrong-way behavior. *Chapter 4*.

knock engine A laboratory test engine that measures the octane number of a gasoline sample. *Chapter 9*.

limit cycle A steady oscillation that a loop settles into and does not leave by itself. *Chapter 5*.

linear programming (LP) A calculation that finds the cheapest recipe or plan that still meets a set of linear limits. *Chapter 15*.

mid-ranging control Using a fast handle for quick corrections while a slow handle brings the fast one back to the middle of its range. *Chapter 8*.

model predictive control (MPC) Control that uses a model to predict the plant and plans the moves of several handles against several limits. *Chapter 1*.

near-infrared (NIR) analyzer An analyzer that passes infrared light through a stream and calculates properties such as octane from the spectrum. *Chapter 9*.

octane number A measure of how well a gasoline resists knock in an engine, tested by a research method (RON) and a motor method (MON). *Chapter 9*.

open-loop instability The behavior of a plant that runs away from its operating point unless a controller holds it there. *Chapter 7*.

outage The empty height left above the contents of a coke drum. *Chapter 11*.

overflash The small share of the vaporized feed that condenses just above the flash zone and keeps the wash zone wet. *Chapter 1*.

override control A second controller that takes over a valve when a limit is about to be crossed. *Chapter 1*.

PID control The standard feedback controller, which acts on the error, on its accumulated sum and on its rate of change. *Chapter 1*.

positive feedback A loop in which part of a change returns to add to itself, so the change grows. *Chapter 1*.

pressure swing adsorption (PSA) Purifying hydrogen by trapping the other gases on beds of solid at high pressure and releasing them at low pressure. *Chapter 12*.

pressure-compensated temperature A tray temperature corrected for pressure, so that it follows the composition and not the pressure. *Chapter 1*.

pumparound A loop that draws hot liquid from a tower, cools it against the crude and returns it, which takes heat out of the tower. *Chapter 4*.

quench Cold hydrogen injected between catalyst beds to take up the heat of the reaction. *Chapter 7*.

rag layer A layer of mixed oil and water between the two clear layers of a desalter that will not come apart. *Chapter 2*.

ratio control Holding one flow at a set proportion of another. *Chapter 2*.

real-time optimization (RTO) A calculation that finds the most profitable targets for a unit from a model and resets them as conditions change. *Chapter 2*.

regenerator The vessel of a cat cracker in which the coke is burned off the catalyst with air. *Chapter 6*.

Reid vapor pressure (RVP) A standard measure of how readily a gasoline evaporates. *Chapter 15*.

riser The vertical pipe of a cat cracker in which hot catalyst meets the feed and the cracking takes place. *Chapter 6*.

runaway A reaction that heats itself faster than the heat is carried away, so that the temperature keeps climbing. *Chapter 1*.

selector A device that passes on the lowest or the highest of several signals, so that the most cautious controller wins. *Chapter 3*.

slide valve A sliding gate in a catalyst pipe that sets how much catalyst flows. *Chapter 6*.

Smith predictor A controller that uses a model to see the effect of its moves before a delayed measurement reports it. *Chapter 8*.

soft sensor A calculation that estimates a property from measurements that are available at once. *Chapter 1*.

split-range control One controller output that drives two valves, each over its own part of the range. *Chapter 14*.

steam jet ejector A device with no moving parts that uses a jet of steam to pull gas out of a tower and hold a vacuum. *Chapter 5*.

steam methane reforming Making hydrogen by reacting natural gas with steam over a catalyst in heated tubes. *Chapter 12*.

steam-to-carbon ratio The steam fed to a reformer for each carbon atom in its feed, kept high enough that carbon does not form on the catalyst. *Chapter 12*.

tail gas The gas left at the end of a process, from the adsorber of a hydrogen plant or from the last condenser of a sulfur plant. *Chapter 12*.

throughput The amount of feed that a unit processes in a given time. *Chapter 1*.

trip An automatic shutdown action taken when a measurement passes a safety limit. *Chapter 2*.

true boiling point (TBP) curve A laboratory curve that gives the temperature at which each fraction of a crude oil boils off. *Chapter 1*.

tube metal temperature The temperature of the tube wall in a fired heater, also called the tube skin temperature, which limits the firing. *Chapter 1*.

ultra-low-sulfur diesel (ULSD) Diesel fuel with nearly all of its sulfur removed, at most 15 parts per million on United States highways. *Chapter 8*.

vacuum distillation Distilling the residue of the crude tower at low pressure, where it boils at a lower temperature. *Chapter 5*.

valve position control A slow controller that watches how far open a valve is and moves another handle to keep that valve in range. *Chapter 1*.

wash bed The packing above the flash zone of a vacuum tower, where liquid washes heavy drops out of the rising vapor. *Chapter 5*.

weighted average bed temperature (WABT) One number for the temperature of a hydrotreater's catalyst, averaged over its beds. *Chapter 8*.

weighted average inlet temperature (WAIT) One number for the severity of a reformer, the average of its bed inlet temperatures weighted by catalyst share. *Chapter 9*.

wet gas compressor The compressor that takes the gas made in a cat cracker and so sets the pressure of the reactor. *Chapter 6*.

Wobbe index The heating value of a gas divided by the square root of its specific gravity, which sets the heat a burner releases. *Chapter 1*.

Index of Control Topics

active constraint
    moving with the conditions, 148
actuator rate limit, 292
actuator saturation, 66
adiabatic reactor, 7
AI agents
    engineering practice, 384
    open problems, 389
    reproduction, 387
    truth audit, 386
    verification record, 385
    while studying, 18
air demand control, 308
    history, 311
analyzer, 10
    in closed loop, 384
analyzer delay, 112, 216, 269, 310, 366
analyzer drift, 381
    detection, 389
averaging level control, 39

back-calculation, 143
back-off, 381
batch step in a continuous plant, 262
bias update, 10, 36, 220, 245, 295, 364
bumpless transfer, 16, 60, 113, 194

commissioning order, 91
computer control
    history, 3
constraint control, 51, 88, 144
    approach span, 137
    design procedure, 382
constraint pushing, 104, 113
constraint with hysteresis, 104
continuous catalyst regeneration, 211
continuous emission monitoring, 322
control hierarchy, 11
control structure design
    history, 88
control system
    security, 384
cumulative average, 10, 353
    stopping floor, 364

data export, 17
dead band, 335
dead time, 310
dead-time compensation, 195
degrees of freedom, 6, 81
distillation
    cut point, 6
disturbance model
    output bias, 89, 246
dynamic compensation, 169

endpoint control, 353
equilibrium-limited reactor, 312
external reset, 144
external reset feedback, 294
extremum seeking control, 36

feasibility, 34, 298
feed and effluent exchanger, 8
feedforward
    fuel quality, 340
    heat balance, 61, 294
    measured load, 144, 220, 244, 314, 338
feedforward control, 381
    from a schedule, 262
    model error, 271
    model-based, 271
final control element, 11
fired heater, 4
flare management plan, 347
flow scheme, 4
fouling, 383

heat integration, 8, 184
high select, 113
hysteresis, 110, 339

IMC tuning, 88
inferential control, 10, 84, 112, 295
inferential octane, 208
integral action
    on an integrating process, 37
integrating process, 32, 293, 329
    gain from the volume, 334
interaction
    triangular, 81
interface level control, 23

interheater control, 208, 220
interlock, 243
internal model control, 315
inverse response, 89

Kalman filter, 17

laboratory measurement, 216, 241, 364
limit cycle, 115, 299, 316, 367
linear programming, 356
loop interaction, 299
loop pairing, 82
low select, 62

measurement bias, 31, 68, 318
mid-ranging control, 184
model predictive control, 63, 88, 146, 169, 195, 221, 245, 295, 339
  against constraint control, 381
  history, 147, 196, 382
  in the apps, 16
  maintenance, 383
  model mismatch, 340
  preview of a disturbance, 339
  relinearized, 271
  valve interlock, 339
move blocking, 246
moving horizon estimation, 17
multiple steady states, 128

open-loop instability, 161
optimization
  infeasible start, 386
optimum
  interior, 33, 309
override control, 50, 88, 113, 144, 161, 168, 195, 294, 314, 379
  history, 62
  measurement, 380
  open problems, 122

parametric sensitivity, 163
  history, 163
partial control, 130
PID control
  anti-windup, 16, 61, 113, 168
  velocity form, 16, 60, 87, 113, 142, 168, 194, 220, 244, 270, 293, 313, 338, 363
positive feedback, 7, 117, 127
  loop gain, 128, 193
pressure transmitter, 346
pressure-compensated temperature, 10, 112

ratio control, 35, 234, 284, 314, 362
  lead-lag, 293
  with a floor, 296
real-time optimization, 23, 155, 245, 382
  bias update, 36, 365
  by feedback, 384
  interior optimum, 38
  steady-state wait, 99
recycle, 242
refinery control
  open problems, 387
runaway, 7, 161
  index, 165

scenarios, 17
  completion rules, 17
scorecard, 17
selective control, 127
selectivity, 237
selector, 88
  low, 144, 294
simulation engine, 15
simulator framework, 14
Smith predictor, 184, 314
soft sensor, 68, 220, 245, 383
split-range control, 329, 331
  overlap, 335
  valve sequencing, 347
step test, 140
stoichiometric ratio, 311
successive linear programming, 245, 365

throughput manipulator, 379
time-scale separation, 58, 233
transport delay, 143
trip
  rate of rise, 168
tube metal temperature, 5, 215, 294
  infrared survey, 302

unmeasured disturbance, 68, 318

valve position control, 8, 50, 194
van 't Hoff equation, 289

weighted average inlet temperature, 218
Wobbe index, 5, 332
  history, 332
  limits, 348
wrong-way behavior, 193

Index of Applications

acid activity, 238
acid gas, 310
acid runaway, 243
  in service, 255
acid strength, 236
  dilution, 240
  industrial controller, 235
  titration and density, 255
acid-soluble oil, 234
afterburn, 134
air blower, 137
alkylation, 233
  history, 236
  solid acid and ionic liquid catalysts, 256
analyzer
  near-infrared
    calibration, 226
  sample system, 97
  sampling, 87
  sulfur, 191
    measuring principle, 201
analyzer validation, 221
antifoam, 268
API gravity, 28
aromatic saturation, 187
asphaltenes, 28
ASTM D86 distillation, 83
atmospheric equivalent temperature, 109
atmospheric tower, 81
autorefrigeration, 239

basic sediment and water (BS&W), 24
burner, 336

C5+ yield, 216
capacitance probe, 26
  failure, 43
carbon capture
  steam reforming, 302
carbon formation, 286
carbon monoxide promoter, 134
carbonyl sulfide, 313
catalyst activity, 132
catalyst bed
  adiabatic, 161
  endothermic, 208
catalyst deactivation, 136, 191, 215
catalyst-to-oil ratio, 131
catalytic naphtha reforming, 208
  history, 211
Claus converter, 310
Claus process, 308
  history, 310
cleaning schedule, 53
  online optimization, 75
co-processing, 383
coil outlet temperature, 4, 86
coke burn, 134, 215
coke drum, 262
  fatigue, 279
  sequencer, 267
  shell temperature, 267
  switch practice, 278
  warm-up, 267
coker gas oil, 269
column inlet temperature, 54
combustion air
  main and trim valves, 322
Conradson carbon residue, 264
contactor, 236
continuous catalyst regeneration
  history, 211
control valve
  positioner, 73
  stiction, 201, 301, 372
cooling water, 117
cracked gas, 110
cracking
  history, 3
crude distillation unit, 81
  history, 83
  industrial control structure, 97
crude oil
  desalting, 23
crude preheat train, 50
cut point, 85, 104, 269

deisobutanizer, 236
delayed coker, 262
  drum switch, 336
delayed coking
  history, 264

delta coke, 134
demulsifier, 25
  history, 26
desalter, 23
diesel cloud point, 86

electrical treater, 25
electrostatic coalescence, 25
emergency depressuring, 178
emissions permit, 311
entrainment, 107
ethanol blending, 355
excess oxygen, 134
exchanger bypass, 56
exchanger cleaning, 57
exchanger fouling, 8, 50
  deposition and removal, 53
  history, 53
  monitoring, 73
  threshold, 53
  velocity dependence, 56

feed and effluent exchanger, 193
  tube leak, 202
feed cut, 166
feed preheat, 142
feed–effluent heat exchanger, 214
fired heater, 54, 86, 214
  coil coking, 111, 264
  fuel gas valve, 227
flare, 329
flare gas recovery, 331
flash zone, 105
flow control loop
  as a final element, 97
flow meter
  orifice and vortex, 300
fluid catalytic cracking, 127
  heat balance, 128
  history, 130
  light feeds, 155
foam
  in a coke drum, 268
fouling resistance, 55
four-lump kinetics, 131, 213
fuel gas
  heating value, 56
fuel gas header, 329

gasoline blending, 353
  certification, 373
  in-line blender, 353
  online control, 366
give-away, 359
globe valve, 43
grid trip, 30

heating value, 334
heavy vacuum gas oil, 107
hindered settling, 29
hot spot
  detection, 179
hydrochloric acid corrosion, 26
hydrocracker, 161
  runaway incident, 167
hydrocracking, 163
hydrodesulfurization, 187
hydrogen, 285
  recycle, 215
  refinery demand, 286
hydrogen header, 287
hydrogen on coke, 134
hydrogen plant, 284
hydrogen quench, 165
hydroprocessing reactor, 7
hydrotreater, 184

IAPWS-IF97, 291
import and flare, 337
instrument, 11
intercondenser, 110
internal reflux, 86

kerosene
  flash point, 86
  freeze point, 86
knock engine, 356

load shedding, 341
log-mean temperature difference, 56

main fractionator, 265
Maxwell–Bonnell correlation, 109
metals (nickel and vanadium), 111
methane slip, 290
minimum stable charge rate, 67
mixing valve, 28

naphtha end point, 86
near-infrared analyzer, 356
  calibration model, 373
  principle, 372

octane, 216, 240, 355
  history, 356
oil in brine, 31
outage, 268
overflash, 86, 111
  measurement, 121

petroleum refining, 1
  history, 2
pitch, 267
pneumatic actuator, 43
pounds per thousand barrels (PTB), 24
pressure swing adsorption, 284
  history, 287
  recovery and purity, 291
  step schedule, 291
pressure transmitter
  absolute, 120
pumparound, 84

quench
  interbed, 190
quench valve
  construction, 178
  stuck, 168

Rachford–Rice equation, 109
  history, 110
radiation level detector, 268
  principle, 277
rag layer, 25
  kinetics, 30
  measurement, 44
reaction furnace, 310
  kinetic model, 323
recovery compressor, 332
refinery fuel gas, 329
reformate
  benzene limit, 228
reformer furnace, 290
refrigeration compressor, 239
  speed and surge, 255
regenerator, 135, 215
  coke burn in service, 227
Reid vapor pressure, 355
  blending index, 355
relief load, 342
riser, 131

salt in crude oil
  laboratory test, 44
slide valve, 136
  construction, 153
  erosion, 154
soft sensor
  drift, 203
specific gravity, 334
standpipe aeration, 136
steam jet ejector, 104
  critical load, 110
  maximum discharge pressure, 110
  operating modes, 121
  unstable mode, 106
steam methane reforming, 284
  equilibrium cells, 289
steam-to-carbon ratio, 286
stripper, 133
sulfur dew point, 312
sulfur recovery, 308
sulfuric acid, 233
surge drum, 291

tail gas, 285
  fuel balance, 290
tail gas analyzer, 310
  principle, 321
tail gas treating, 323
thermocouple, 153
  bed array, 177
  reference junction, 71
thermowell, 153
throughput, 51
true boiling point curve, 83
tube skin temperature, 111, 268

ultra-low-sulfur diesel, 184
  regulation, 186

vacuum distillation, 104
  assumptions, 113
  history, 106
vacuum residue, 107, 263

wash bed, 107
  coking, 111
wash zone, 86
waste-heat boiler, 291
water seal, 332
water-gas shift, 286
weighted average bed temperature, 190
wet gas compressor, 135, 269
  speed control, 278

About the Editor

John D. Hedengren is a professor in the Department of Chemical and Biological Engineering at Brigham Young University, where he leads the Process Research and Intelligent Systems Modeling (PRISM) group. He received the B.S. and M.S. degrees in chemical engineering from Brigham Young University and the Ph.D. degree from the University of Texas at Austin in 2005. He developed advanced process control applications for polymer production at ExxonMobil Chemical Company and commissioned nonlinear model predictive controllers with PAS, Inc. before joining the faculty at Brigham Young University in 2011.

His teaching covers process dynamics and control, engineering computing, data science, machine learning, and dynamic optimization, with course material published online for use inside and outside the classroom. He developed the Temperature Control Lab, a portable heat transfer device now used by many universities for process control and data science instruction, and he maintains the APMonitor and GEKKO optimization software used in those courses. He serves as Chair of Control Education for the IEEE Control Systems Society and as a trustee of the CACHE Corporation. He received the AIChE David Himmelblau Award in 2014, the AIChE Computing Practice Award in 2018, and the John R. Ragazzini Award in 2024, and he was inducted into the Process Automation Hall of Fame in 2025.

Volumes in This Series

Interactive Control Engineering: Principles and Applications has 20 volumes. Nine are organized by control technology and eleven by industry or application area. Every chapter in every volume has a simulator at https://apmonitor.com/apps/.

1 **Process Dynamics, Modeling, and Identification.** How plants respond, and how to find the model.
2 **Practical PID Control.** Tuning, anti-windup and implementation that works on real loops.
3 **Feedforward, Cascade, and Ratio Control.** Choosing a structure from where the disturbance enters.
4 **Override, Split-Range, and Constraint Control.** When limits decide what the controller may do.
5 **Delays, Slow Measurements, and State Estimation.** Controlling what arrives late or cannot be measured.
6 **Multivariable and Plantwide Control.** Pairing loops, managing inventories, closing the recycle.
7 **Nonlinear Dynamics, Oscillations, and Stabilization.** Plants that will not stay put without control.
8 **Batch, Run-to-Run, and Hybrid Control.** Recipes, endpoints and learning from the last run.
9 **Model Predictive Control and Economic Optimization.** Planning moves ahead, and choosing where to operate.
10 **Control in Upstream and Midstream Oil and Gas.** From the wellbore to the terminal.
11 **Control in Petroleum Refining.** Throughput, quality and limits from crude unit to blending.
12 **Control in Petrochemicals and Polymers.** Severity, selectivity and grade from cracker to pellet.
13 **Control in Chemical Plants.** Recycle, equilibrium and energy in continuous chemical processes.
14 **Control in Semiconductor Manufacturing.** Run-to-run control where little can be measured in the chamber.
15 **Control in Mining and Mineral Processing.** Variable ore, late assays and equipment at its limit.
16 **Control in Iron, Steel, and Metals Production.** Slow furnaces, fast strip, and measurements that arrive once.
17 **Control of Energy Conversion and Storage.** Fast regulation and slow dispatch on a changing grid.
18 **Control of Nuclear and Fusion Systems.** Reactor physics, protection limits and coordinated control.
19 **Control of Biological Systems and Human Performance.** Variable subjects, sparse samples and limits that cannot be crossed.
20 **Motion Control, Robotics, and Autonomous Systems.** Tracking, navigation, contact and allocation.