INTERACTIVE CONTROL ENGINEERING
Principles and Applications

Volume 1 Process Dynamics, Modeling, and Identification

CHAPTER 12

Multivariable Plant Identification Workbench with Exact Structural Zeros

Pages 284–314

Edited by John D. Hedengren

How to cite this chapter. "Multivariable Plant Identification Workbench with Exact Structural Zeros," in *Process Dynamics, Modeling, and Identification*, John D. Hedengren, Ed., vol. 1 of *Interactive Control Engineering: Principles and Applications*. Advanced Process Solutions, LLC, 2026, ch. 12, pp. 284–314.

The page numbers in this file are those of the printed volume.

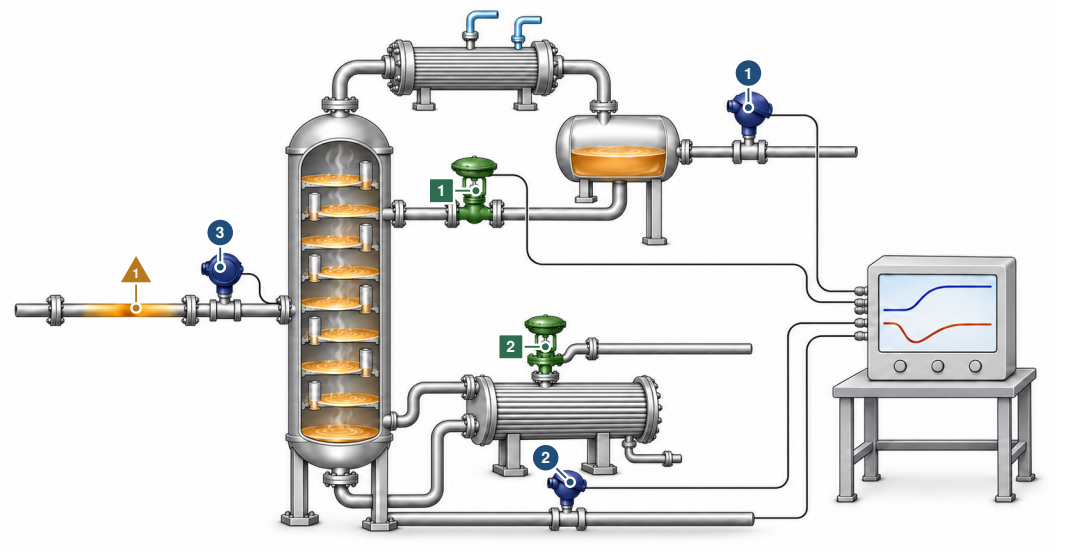
The app for this chapter
https://apmonitor.com/apps/identify/

The volume online: contents, other chapters and citation records
https://apmonitor.com/apps/books/v01-dynamics-modeling/#identify

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

CHAPTER 12

Multivariable Plant Identification Workbench with Exact Structural Zeros



The controller moves 1 reflux flow, stepped in the plant test 2 steam to the reboiler, stepped in the plant test

The controller reads 1 distillate composition analyzer 2 bottoms composition analyzer 3 feed flow meter

What disturbs it 1 feed rate changes, measured but not adjusted

Notice that reflux and steam both move each product composition, so a fit must divide the credit between them from one record; the column pressure transmitter of the chapter is not drawn. Illustration, not to scale.

IN PLAIN TERMS Which Input Moved Which Output?

The main idea. A plant keeps a record of the handles its operators moved and the readings that followed. This chapter fits models to such a record to learn how strongly, and how late, each handle moves each reading.

The challenge. When several handles move at once, a careless fit gives one handle the credit for another's work. Engineers also know that some handles cannot touch some readings. The model must leave those links out completely, and it must be tested on a part of the record it never saw.

Why it matters. A controller that steers a plant from such a model can only be as right as the model.

THREE WAYS TO READ THIS CHAPTER

⊘	**Explore**	12 min	the first page, the scale of predictive control in industry (Section 12.1), the history of Identification and Command (Section 12.2), the try-it on an absent path, the fact on true and false zeros and the try-it on the predictive controller (Section 12.6) and the key points
🕮	**Study**	65 min	the whole chapter with its three pause points and the twelve exercises at its end
⛰	**Practice**	20 min	the first page, the results (Section 12.6), Technology in Practice (Section 12.7), Limitations (Section 12.8) and the open challenge

The control question A plant record holds several inputs and outputs that move together, sometimes under feedback. How is the credit for each output divided among correlated inputs, how is an absent path removed exactly, and how is a model judged on data the fit never saw?

Learning objectives

- Explain why the inputs into a CV are fitted jointly, and why a fitted gain depends on the model structure.
- Compute the sampled form of a path with fractional dead time and the parameter count of a block state-space model.
- Tell a correct structural zero from a false one by free-run validation error, and choose a family for an integrating or ringing record.
- Pair and tune PI loops from a fit, and explain why an output-bias predictive controller loses an unmeasured load.

Prerequisites Sections 1.4.3 to 1.4.6, 1.4.8 and 1.4.9, and Chapter 11. Sections 1.6.2 and 1.6.3 only for Section 12.5.

The app https://apmonitor.com/apps/identify/

See also Volume 6, Chapter 2 (loop pairing) and Volume 6, Chapter 8 (predictive control above a decentralized structure). Volume 9, Chapter 4 (offset-free predictive control with a Kalman estimator).



12.1 Introduction

The models inside industrial multivariable controllers are seldom derived from conservation laws. They are identified from plant tests in which operators step each manipulated variable in turn or together, and the result is a matrix of responses that relates every manipulated and measured disturbance variable to every controlled variable. That arrangement goes back to the first predictive controllers, which predicted from impulse-response models obtained on the plant [1], and it remained the dominant model form across the industrial vendors surveyed a quarter century later, where the identification step and its plant test were reported as the most time-consuming part of a project [2]. A textbook treatment of identification usually stops at one input and one output, and the single-input reaction curve is the model a first course teaches a student to read. Section 1.4.8 of Chapter 1 fits that case, and Chapter 11 chooses among single-input models of a heater.

DID YOU KNOW? More Than 4600 Predictive Control Applications

A survey of industrial predictive control published in 2003 counted more than 4600 applications carried out by the vendors themselves, over twice the number in a survey five years earlier. Refining held the largest single block, and the size of an application ranged from a single variable to several hundred variables [2].

Three things change when there are several inputs. The first is credit. Two inputs that moved during the same hour both explain the same output movement, and a fit of each input alone credits both with it, so the inputs entering a CV must be fitted together and their correlation becomes a property of the data that limits what any method can recover [3]. The second is structure. An engineer usually knows that some paths do not exist, for example that a reflux valve has no mechanism by which to move column pressure, and a model that estimates a small spurious gain on that path is both less interpretable and, when the controller acts on it, less safe than one that holds it at zero. The third is honesty in scoring. A model that predicts one sample ahead from the latest measurement looks accurate on almost any record, and the free-run simulation that a controller actually relies on is a much harder test [4], as Table 1.1 of Chapter 1 shows for a single heater. Distillation trains, fired heaters, reactors with jacket and feed handles, building air handlers and compressor stations all share this shape, which is why multivariable plant testing for predictive control has its own literature of test signals, order selection and error bounds [2, 3].

The misconception this application is built to correct is that a multivariable model is a collection of single-input models. Fitting each path separately and adding the answers double-counts correlated inputs, a small estimated gain is mistaken for an absent path, and a one-step-ahead fit is mistaken for a model. The workbench removes each shortcut. All active inputs into a CV are fitted jointly, a path set to zero is removed from every model family so exactly that its Markov parameters are 0.0 and not 10^{-12}, and candidates are ranked on free-run error over a validation interval that the fit did not see, with a final test interval held back until a model is chosen.

This chapter works with that workbench, its five model families and the measured behavior of each family on five records whose true structure is known. Every estimator and controller in it is an established method. Section 12.2 traces plant testing and identification for multivariable control. Section 12.3 states the identification problem: the signals and their roles, the preparation of a record, the five families, the method settings and seven assumptions. Section 12.4 gives the algorithms as coded, in enough detail to rebuild the application, among them an exact zero-order-hold discretization with fractional dead time shared by fitting and export, a structural block realization that enforces a zero path in a subspace model, and an export that proves its own equivalence to the application in GEKKO. Section 12.5 builds a predictive controller and decentralized PI loops from a chosen fit. Section 12.6 reports validation scores, recovered gains and parity errors read off the running pipeline, including a weak path that is real, whose removal raises validation error by a factor of 3.7, and measures the predictive controller against the PI loops. Section 12.7 describes the devices and platforms behind such a record, and Sections 12.8 to 12.10 give the limitations, a discussion and a conclusion, and the exercises of Section 12.11 follow the thirteen scenarios of the app.

12.2 Plant Testing and Identification for Multivariable Control

Statistical identification of dynamic systems from process data took its modern shape in the 1960s with maximum-likelihood estimation of difference equations from normal operating records, work that was aimed explicitly at industrial processes and that introduced the noise model as part of the estimation problem [5]. By 1971 a survey could already organize the field into correlation, least-squares, instrumental-variable and maximum-likelihood methods, with model structure, input design and validation as recognized topics in their own right [6]. The same decade produced the Box-Jenkins family of input and noise polynomials, and the prediction-error framework that later unified them treats every structure, ARX and output-error included, as a predictor whose one-step error is minimized [4, 7]. Realization theory developed in parallel. A finite-dimensional state-space model can be constructed from its sequence of impulse-response (Markov) parameters [8], which is the result this application relies on when it certifies a zero path by checking those parameters.

Industrial practice took a more direct route. The predictive controllers of the late 1970s stored the plant as an impulse or step response identified from a plant test [1], and finite step-response and impulse models remained the most common representation in commercial packages surveyed in 2003, alongside a growing use of state-space forms [2]. The response matrix was attractive because each cell can be read by a process engineer. A gain has a sign that can be checked against a balance, a dead time can be checked against a pipe length, and an empty cell states a structural belief about the unit. Plant testing for these controllers matured into multivariable test design with simultaneous perturbation signals, model order chosen by an asymptotic criterion and frequency-domain error bounds on each path [3].

> ⧖ HISTORY **Identification and Command, 1976 to 1978**
>
> The first applications of model predictive control were presented at a conference in 1976 and summarized in a journal paper in 1978. Their software was called IDCOM, an acronym for Identification and Command. Its plant model was a finite impulse response identified from plant test data, by an algorithm that minimized the distance between the plant and model impulse responses in parameter space. The same algorithm then solved the control problem, which its authors treated as the mathematical dual of identification. The papers

> also fixed a vocabulary this chapter still uses: manipulated variables that the controller adjusts, disturbance variables that are not available for control, and controlled variables at the outputs. The first papers on the rival step-response method, in 1979 and 1980, did not discuss their identification technology at all [2].

Subspace methods supplied the state-space alternative. MOESP estimates the extended observability matrix of a deterministic system from an LQ factorization of block Hankel matrices of input and output data [9], its past-output variant handles a system in innovations form [10], and N4SID reached the same structure through oblique projections [11]. These methods need no iterative search and deliver a coupled model of every output at once, which is their strength and also the reason a zero path cannot be imposed on them by editing one entry of a matrix. Plant records are also frequently collected while controllers run, and closed-loop identification showed that the direct prediction-error approach remains consistent under feedback only when the noise model is right, while instrumental and two-stage methods trade efficiency for robustness to that error [12].

Two benchmarks recur in teaching this material. The Wood-Berry transfer matrix of a methanol-water pilot column, with its large opposing gains and unequal dead times, has been a standard multivariable test problem since 1973 [13]. The Temperature Control Lab supplies a desktop two-input, two-output thermal plant on which ARX and other empirical models are identified and then used for control [14], and the GEKKO Optimization Suite provides the `arx` and `state_space` objects into which such models are loaded for simulation, estimation and predictive control [15].

12.3 Identification Problem and Data Model

The object identified is a linear time-invariant map from m inputs to p outputs on a uniform sample grid. The inputs are the MVs followed by the measured DVs, and the outputs are the CVs. An unmeasured disturbance and sensor noise enter the record but not the model, and every quantity the algorithms touch is labeled in Figure 12.1, using the distillation column record as the instance.

12.3.1 Signals and Roles

Every column of an imported table receives one of five roles, Time, MV, DV, CV or Ignore, together with a display name, units and a description. An MV is an input an operator or controller can change, a DV is a measured input nobody adjusts, and a CV is an output to predict. MVs and DVs both enter the input vector, MVs first, and keep their identity in every view and export. The application proposes roles from header text, such as `valve`, `Q1` or `level`, and marks each proposal as a guess until the user confirms or changes it. A column cannot be both an input and an output. During identification an MV is a recorded trajectory and never a decision variable. Table 12.1 classifies the signals of the two records used most in what follows.

12.3.2 Preparation of the Record

Raw records are kept as text and never edited, and every transformation is a field of a stored recipe, so a reopened project reproduces the same prepared data. Every model uses one uniform sample period T_s, by default the median recorded interval. MVs are held between samples as an actuator is. CVs and DVs are resampled by default with the mean over the preceding interval $(t - T_s, t]$, a causal anti-alias filter, or by linear interpolation or hold on request. Linear interpolation never reaches forward into a different split, and no method bridges an interval longer than the gap limit of three sample periods. A gap, an excluded stretch, a return to an earlier split or a record boundary starts a new segment with its own initial state, so two campaigns are never joined into one trajectory and no model history runs from held-out data back into training. Adjacent training, validation and test stretches in chronological order stay in one segment and keep their causal input history. Duplicate stamps are averaged, rows out of order are sorted, and both are reported.

Splits are chronological and never shuffled. The default marks the first 60 % of each record as training, the next 20 % as validation and the last 20 % as the final test, and a record shorter than 300

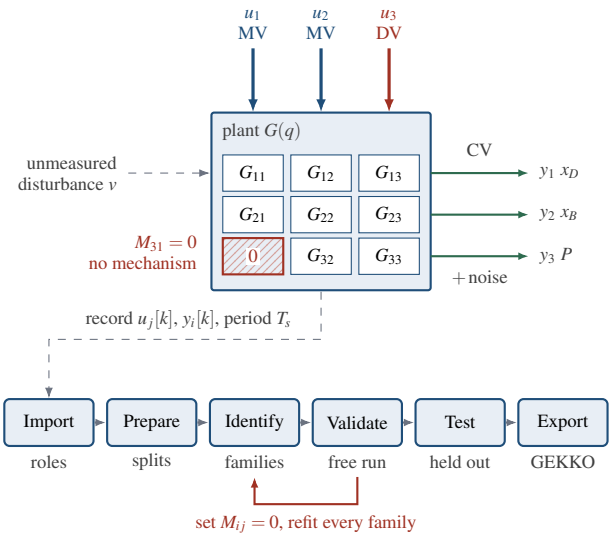


Figure 12.1 Multivariable plant as the identification sees it, and the six identification steps of the workflow, which a seventh step that builds a predictive controller from the chosen fit follows (Section 12.5). Two MVs and one measured DV enter the transfer matrix $G(q)$, whose rows are the CVs, and an unmeasured disturbance and sensor noise reach the record without entering the model. For the column record u_1 is reflux, u_2 steam and u_3 feed rate. The hatched cell is a path fixed at zero from process knowledge, which every family enforces exactly and which the validation step can contradict.

Table 12.1 Classification of Signals in the Two-Heater and Column Records. Ranges Are Those of the Recorded Test Signals.

Signal	Role	Units	Recorded range
TCLab two-heater test, $T_s = 5$ s			
Q_1 heater 1	MV	%	0 to 80, held 180 to 480 s
Q_2 heater 2	MV	%	0 to 80, held 180 to 480 s
T_1, T_2	CV	°C	measured, 0.32 °C steps
distillation column, $T_s = 1$ min			
R reflux	MV	lb/min	1.75 to 2.15, held 25 to 70 min
S steam	MV	lb/min	1.55 to 1.85, held 25 to 70 min
F feed rate	DV	lb/min	2.2 to 2.7, held 60 to 160 min
x_D distillate	CV	mol%	near 96
x_B bottoms	CV	mol%	near 0.5
P pressure	CV	bar	near 1.2
model quantities, both records			
$x[k]$	latent state	—	not a physical inventory
$e_i[k]$	residual	CV units	unmeasured
M_{ij}	structural mask	—	1 estimated, 0 fixed at zero

samples receives a 70 % and 30 % training and validation split with no test. Splits are marked by dragging across a chart or typing exact times. Each mark is a translucent band named along its length in its own color, overlaps blend and stay visible, the later mark decides, and a ribbon along the top of each chart shows the resolved split (Figure 12.2). A zoom strip under the charts starts on the whole record, so a short stretch of a long record can be marked. Final-test scores stay hidden until a separate Test step reveals them for the selected current model, and every inspection is recorded. Baselines $\bar{u}_j$

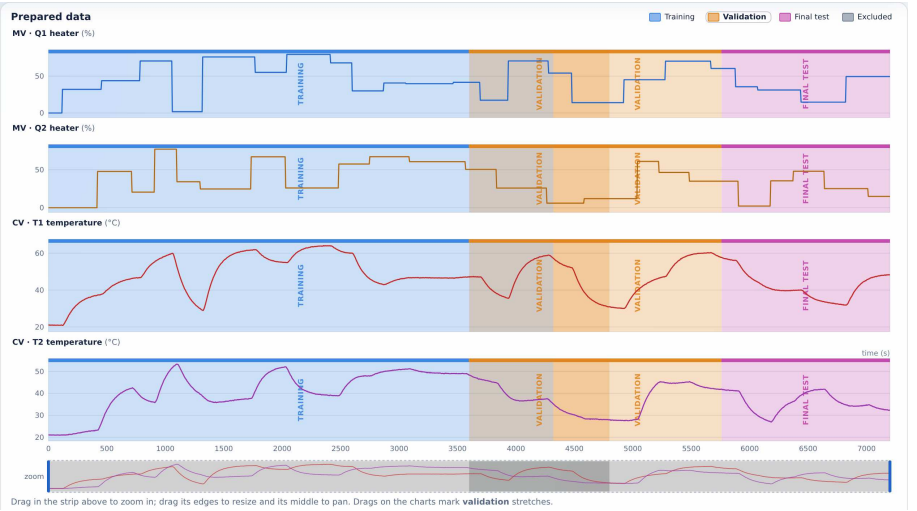


Figure 12.2 The Prepare step on the TCLab record after a validation mark from 3600 to 4800 s was added over the default splits (training to 4320 s, validation to 5760 s, final test to 7200 s). Each mark is a translucent band named along its length. Where the new mark overlaps training the two colors blend, and where it overlaps the default validation band the color deepens. The ribbon along the top of each chart shows the resolved split, validation from 3600 s because the later mark decides, and the zoom strip under the charts shows the whole record.

and $\bar{y}_i$, scales $\sigma_{u,j}$ and $\sigma_{y,i}$, optional drift slopes and optional outlier limits are computed on training samples only. A readiness summary reports, before any fit, inputs that do not move, inputs that move together (level correlation above 0.9 or move correlation above 0.8 for limited confidence, 0.995 and 0.98 for cannot estimate), channels pinned at a limit, outputs that move with no input moving, too many coefficients per training sample, dynamics slower than any training stretch, closed-loop records, and an input that tracks a CV so closely (correlation above 0.995) that it may be derived from it.

12.3.3 Model Families

All five families work in deviation variables about the training baselines, $u'_j[k] = u_j[k] - \bar{u}_j$ for each input and $\hat{y}_i[k] = \bar{y}_i + \delta_i[k]$ for each CV. The two transfer-function families give each active path a gain K_{ij}, one or two lags and a dead time θ_{ij}, as shown in Equation (12.1), the forms of Equations (1.15) and (1.11) in Chapter 1.

$$G_{ij}(s) = \frac{K_{ij}\,\mathrm{e}^{-\theta_{ij}s}}{\tau_{ij}s+1} \quad \text{or} \quad \frac{K_{ij}\,\mathrm{e}^{-\theta_{ij}s}}{\tau_{ij}^2 s^2 + 2\zeta_{ij}\tau_{ij}s+1} \tag{12.1}$$

where τ_{ij} is a time constant in record time units and ζ_{ij} is a damping factor. For $\zeta_{ij} \geq 1$ the second-order path has two real lags $\tau_{ij}(\zeta_{ij} \pm \sqrt{\zeta_{ij}^2 - 1})$, and for $\zeta_{ij} < 1$ it rings. The CV is the sum of its active paths plus an offset c_i. Any path inside the second-order candidate can be held first order, and any path can be given a gain sign, bounds on its lag and dead time, or fixed parameter values.

The ARX family writes each CV as a difference equation in its own history and every active input, as shown in Equation (12.2), of which Equation (1.20) of Chapter 1 is the case of one input with

$n_a = n_b = 1$.

$$\begin{aligned}\delta_i[k] = c_i + \sum_{l=1}^{n_a} a_{i,l}\,\delta_i[k-l] \\ + \sum_j \sum_{m=0}^{n_b-1} b_{i,j,m}\,u'_j[k-d_{ij}-m] + e_i[k]\end{aligned} \tag{12.2}$$

where n_a and n_b are the output and input orders, $d_{ij} \geq 1$ is an explicit dead time in samples (zero only when same-sample feedthrough is enabled) and e_i is the equation error. The output polynomial is diagonal, with no cross-output terms, so the transfer matrix entry is $B_{ij}(q)/A_i(q)$ and removing a path is exactly removing its numerator coefficients.

The FIR family has no assumed shape, as shown in Equation (12.3).

$$\delta_i[k] = c_i + \sum_j \sum_{l=0}^{H-1} h_{i,j,l}\,u'_j[k-d_{ij}-l] \tag{12.3}$$

where H is the horizon in samples and $h_{i,j,l}$ are the impulse-response coefficients. A step response is the running sum of $h_{i,j,l}$, and a horizon long enough to settle carries one coefficient per sample per path.

The state-space family shares a small latent state among all inputs and outputs, as shown in Equation (12.4).

$$\begin{aligned}x[k+1] = Ax[k] + B\tilde{u}[k], \qquad \tilde{u}_j[k] = u'_j[k-d_j] \\ \delta[k] = Cx[k] + D\tilde{u}[k] + c\end{aligned} \tag{12.4}$$

where d_j is a whole-sample delay per input and $D = 0$ unless feedthrough is allowed. A state block too slow for the record also adds a constant drift to its state equation (Section 12.4.5). Every family is scored with the normalized root-mean-square error of Equation (12.5).

$$\mathrm{NRMSE}_i = \frac{1}{\sigma_{y,i}} \sqrt{\frac{1}{N} \sum_k \left(y_i[k] - \hat{y}_i[k]\right)^2} \tag{12.5}$$

where the sum runs over the N samples of one split and $\sigma_{y,i}$ is the training standard deviation, so every split and every candidate is measured on one scale. The aggregate is the mean of NRMSE_i over the CVs. RMSE, MAE and bias in engineering units and a fit $100(1 - \|y_i - \hat{y}_i\|/\|y_i - \bar{y}_i\|)$, which is negative for a model worse than the split mean and undefined for a constant output, are reported beside it.

12.3.4 Method Settings

The application has no hidden plant parameters, because the plant is whatever record is loaded. Its parameters are the method settings in Table 12.2, where Status marks what the Methods panel and the Complexity selector expose, what lies behind Advanced, and what is fixed in the code.

The Complexity selector changes four of these settings together. The simple level uses 2 starting points, ARX orders up to 2 and 2 and state order up to 4, and the detailed level uses 6 starting points, ARX orders up to 5 and 6 and state order up to 12. The standard level is the default of Table 12.2.

12.3.5 Assumptions

The models rest on seven assumptions, and each one fails somewhere.

Linearity and time invariance. Every family is linear about the training baselines. A plant whose gain changes with the operating point receives an average over the training range, and extrapolation outside that range is never validated. The TCLab record is generated from an energy balance with a radiative term, so its identified gains are averages of a gain that changes with temperature.

Table 12.2 Method Settings and Defaults. S Is the Longest Training Stretch in Time Units, at Least a Quarter of the Longest Segment and $10T_s$, N Is the Number of Training Samples and N_s the Longest Segment in Samples.

Setting	Default	Status
preparation		
sample period T_s	median recorded interval	visible
CV and DV resampling	interval mean	visible
gap limit	$3T_s$	visible
splits per record	60/20/20, under 300: 70/30	visible
drift removal, outlier limit	off, 0	visible
transfer functions		
starting points per CV	4 (+ previous fit)	Advanced
iterations, scouting then total	8, then 60	Adv., fixed
bounds on $\ln\tau$	$\ln(T_s/20)$ to $\ln 3S$	per path
bounds on θ, ζ	0 to $S/2$, 0.1 to 28.3	per path, fixed
LM start damping, cost and step tolerances	10^{-3}, 10^{-9}, 10^{-8}	fixed
delays, ARX and FIR		
impulse onset, delay scan	8 % of peak, onset -3 to $+1$	fixed
orders searched n_a, n_b	1 to 3, 1 to 4	Advanced
parsimony in order search	1 %	fixed
ridge, simulation refinement	0, on	Advanced
refinement iterations, tolerance	30, 10^{-7}	fixed
FIR horizon H	$N/(3m)$, 10 to $\min(240, 720/m)$	Advanced
FIR penalty λ	7 values, 10^{-6} to 1	fixed grid
state space		
largest order per block	8	Advanced
Hankel horizon i	$\max(n_{\max}+2, 8)$, $2i(m+l) \le 480$	Advanced
refinement iterations, tolerance	25, 10^{-8}	fixed
feedthrough D	off	Advanced
starting rule, scoring and checks		
measured start when $\lvert z\rvert \ge$	$1 - 5/N_s$	fixed
unstable only when $\lvert z\rvert >$	$1 + 1/N_s$	fixed
warm-up samples not scored	3 per segment	fixed
CV weights	$1/\sigma^2_{y,i}$	Advanced
ranking parsimony	2 %	fixed
zero-path horizon	$\max(1000, 4\times$ response$)$	fixed
bootstrap replicates, block	30, $\max(10, 2N^{1/3})$	20 to 100, fixed
refit pause after an edit	650 ms	fixed

Zero-order hold. Inputs are taken as constant between samples. This is exact for an MV written by a controller and an approximation for a measured DV, which varies within the interval. The interval-mean resampling of a DV partly compensates by averaging over that interval.

Steady state at each segment start, unless the dynamics are too slow. A free run begins with every input held at its first value for all earlier time and the model at the matching steady state, so a record that starts mid-transient puts that transient into the early residual. An ARX output or state block slower than a fifth of the longest segment, an integrator included, starts from its first measurements instead (Section 12.4.7).

No noise model. The transfer-function, FIR and state-space families model no noise, so their predictor is their simulation. ARX starts from the equation-error least-squares estimate, the one-step fit of Section 1.4.8, which is biased when the disturbance is colored, and by default refines it on simulation error, which makes its final estimate an output-error fit with ARX polynomials. Output-error, ARMAX and Box-Jenkins noise models and instrumental variables are not implemented.

Open-loop excitation unless declared otherwise. Under feedback the MVs are correlated with the

disturbance through the controller, and every estimator here can be biased toward the inverse of the controller unless external excitation such as a setpoint step dominates the input [12]. A closed-loop declaration adds a readiness finding and changes how the residual-input correlation is described, but it does not change the estimator.

Integrating paths have no finite gain. The transfer-function families contain no integrator, and a long lag looks like one over a short record (Section 1.4.4). A path whose fitted time constant exceeds the longest training stretch S is reported as integrating with its largest step-response slope rather than with a gain the data never saw settle. ARX and state space can hold the integrator itself, and a pole within T_s/S of the unit circle is reported the same way.

A zero path is a structural claim. Fixing $M_{ij} = 0$ asserts that no mechanism connects input j to CV i over the whole operating range. It is not a statistical result, and the validation score is the only evidence the application can offer for or against it.

12.4 Identification Algorithms as Implemented

12.4.1 Exact Sampling with Fractional Dead Time

Every transfer-function path is sampled exactly for a held input, with no Padé approximation and no rounding of the dead time. The dead time is split as $\theta = dT_s + \phi$ with $0 \le \phi < T_s$, so over one interval the delayed input equals $u[k-d-1]$ for the first ϕ and $u[k-d]$ for the remaining $T_s - \phi$, as shown in Equation (12.6).

$$\begin{aligned} x[k+1] &= \Phi x[k] + \Gamma_1 u[k-d-1] + \Gamma_0 u[k-d] \\ y[k] &= x_1[k] \\ \Phi &= \mathrm{e}^{AT_s}, \qquad \Gamma_0 = \int_0^{T_s-\phi} \mathrm{e}^{As}\,\mathrm{d}s\,B \\ \Gamma_1 &= \mathrm{e}^{A(T_s-\phi)} \int_0^{\phi} \mathrm{e}^{As}\,\mathrm{d}s\,B \end{aligned} \tag{12.6}$$

where the unit-gain continuous realization is $A = -1/\tau$, $B = 1/\tau$ for first order and, on the states y and $\tau\dot{y}$ for second order, is given by Equation (12.7).

$$A = \begin{bmatrix} 0 & 1/\tau \\ -1/\tau & -2\zeta/\tau \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1/\tau \end{bmatrix} \tag{12.7}$$

Each pair e^{Ah} and $\int_0^h \mathrm{e}^{As}\mathrm{d}s\,B$ comes from one exponential of the augmented matrix $\left[\begin{smallmatrix} Ah & Bh \\ 0 & 0 \end{smallmatrix}\right]$, computed by a degree-6 Padé approximant with scaling and squaring. The path output is K times $y[k]$. A step applied at $k = 0$ therefore first appears in the sampled output at $k = d + 1$, the first sample after the dead time, and θ keeps its continuous meaning without rounding. Equation (12.6) is also the realization written into the export, so the fitted model and the exported model are the same object.

12.4.2 Transfer Functions by Variable Projection

All active paths into one CV are fitted together. For fixed lags, damping factors and dead times, the free-run output is linear in the gains and the offset, so those are solved exactly by least squares inside the residual function and the outer search moves only the nonlinear parameters, which is the variable projection idea for separable problems [16]. With $s_{ij}[k]$ the unit-gain response of path j to the recorded u'_j for nonlinear parameters p, the reduced residual is given by Equation (12.8).

$$r_i(p) = \sqrt{w_i}\, W^{1/2} \left(\delta_i - c_i^{\star}(p) - \sum_j K_{ij}^{\star}(p)\, s_{ij}(p) \right) \tag{12.8}$$

where δ_i is the training target, W holds the per-sample weights (zero for a down-weighted outlier), $w_i = 1/\sigma_{y,i}^2$ is the CV weight, and $K_{ij}^{\star}$ and $c_i^{\star}$ are the least-squares gains and offset for that p. The

nonlinear vector holds $\ln \tau_{ij}$, a damping coordinate q_{ij} (second order) and θ_{ij} for every parameter not fixed by the user, with the damping given by Equation (12.9).

$$\zeta_{ij} = \begin{cases} \mathrm{e}^{q_{ij}}, & q_{ij} \le 0 \\ 1 + \sinh q_{ij}, & q_{ij} > 0 \end{cases} \tag{12.9}$$

The coordinate is monotone with unit slope on both sides of $\zeta = 1$, linear where a path rings and nearly logarithmic in the ratio of the two real lags, which grows as e^{2q}. Its bounds $q = \ln 0.1$ and $q = 4$ give ζ from 0.1 to 28.3, a lag ratio of about 3200, so a path whose second lag the data cannot see runs down a straight valley to the upper bound and stops, whereas in (τ, ζ) coordinates the same path crept along a curved valley to its iteration limit. A gain sign constraint is enforced by an active set that pins a gain of the forbidden sign at zero and re-solves the others. Each path's response column is cached by its exact parameters, so a difference step on one path re-simulates only that path.

The gains come from cached weighted inner products of the path columns by an equilibrated Cholesky factorization, and the Jacobian is the variable projection Jacobian of Golub and Pereyra [16] rather than a least-squares re-solve per parameter. For a parameter of path j with column derivative D, one forward difference of that column alone, the change of the fit is given by Equation (12.10).

$$\begin{aligned} \partial f &= K_j D + A\, \partial c \\ \partial c &= M^{-1} \left(e_j \langle D, We \rangle - K_j A^{\mathsf{T}} W D \right) \end{aligned} \tag{12.10}$$

where A holds the free columns and the constant, $M = A^{\mathsf{T}} W A$, c the solved gains, e the misfit and e_j the unit vector of path j. A fixed gain keeps only the second term inside the parentheses, and a pinned gain contributes nothing. Tests check (12.10) against differences of the residual to 10^{-4} relative and check that it reaches the same minimum as a plain route that re-solves the gains by QR for every perturbation.

The outer search is a bounded Levenberg-Marquardt method [17] with (12.10) as its Jacobian and an active set at the bounds. A parameter on a bound, within 10^{-9} of its span, whose gradient would carry it outside is held for the iteration and the step is computed for the others. Clipping a joint step instead, as the first release did, spoils the step of every other parameter and lets a search creep to its limit. Differences use a step of $10^{-6} \max(|p|, 1)$, at most 1 % of the bound span and inward at an upper bound. The damped normal equations add $\lambda \max(d_{aa}, 10^{-12} d_{\max})$ to each diagonal entry of $J^{\mathsf{T}} J$ and are solved by Cholesky, with λ starting at 10^{-3}, divided by 5 after an accepted step and multiplied by 8 after a rejected one (10 if the factorization fails), up to 12 attempts per iteration, and trial points projected onto the bounds. The search stops on a relative cost decrease below 10^{-9}, a relative step $|\Delta p| / (|p| + 1)$ below 10^{-8}, a vanishing gradient, every free parameter held, or no damping giving a decrease. A parameter that ends within 10^{-6} of its bound span from a bound is reported in words, except a dead time of zero, which is simply a path without dead time.

Starting points come from a regularized FIR fit with the automatic horizon. Each path's FIR step response gives a gain from its last 15 % and a lag from its 63.2 % crossing, and the dead time is $\theta_{ij} = (o_{ij} - 1) T_s$, where the onset o_{ij} is the FIR delay plus the first lag whose impulse coefficient reaches 8 % of the largest. The onset is right for an integrating path, whose 5 % step crossing lies far up its ramp, where the first release put its dead time. For second order, an overshoot above 5 % gives ζ from the logarithmic decrement and τ from the time to peak, and otherwise $\zeta = 1$ and τ is the first-order lag divided by 2.15. Four multistarts scale the lag and dead time and shift the damping coordinate by $(1, 1, 0)$, $(0.4, 1, 1.5)$, $(2.5, 1, 0)$ and $(1, 0.5, -0.7)$, and a refit adds the previous valid model as a fifth. Every start receives 8 scouting iterations and the best continues for the remaining 52 of its 60, which is what lets a record of overlapping random steps, rather than isolated clean steps, be fitted at all.

12.4.3 Delays, then ARX by QR Least Squares

Each CV's recurrence (12.2) is solved by streaming the rows into a Givens QR factorization with column equilibration and a rank decision on the diagonal of R at 10^{-11} relative, so a regressor that carries no information is dropped instead of receiving a large coefficient. A row is used only when its target and every lagged output lie in the training split.

The dead times d_{ij}, shared by ARX and state space, start from the onsets o_{ij} and are refined for each CV by a one-step loss. With the other inputs at their current delays, every candidate from $o_{ij}-3$ to $o_{ij}+1$ is scored by the weighted residual sum of squares of a one-step ARX with $n_a = n_b = 2$, in two passes over the inputs. The candidates share one parameter count, so no validation data are needed, a one-step loss is not inflated by an integrator, and one factorization per input covers all of its candidates as column subsets of the triangular factor.

All twelve order pairs with $n_a \le 3$ and $n_b \le 4$ are solved from one factorization of the largest pair, each smaller pair by a small QR of the matching columns of its R plus the few rows that a shorter output history makes usable after an excluded stretch. Each pair is simulated once from the starting rule of Section 12.4.7 and scored on training and validation, a pole more than $1/N_s$ outside the unit circle scores as infinite, and a pair replaces the best so far only if it lowers the validation NRMSE by 1 %, or at all with no more coefficients.

The chosen pair is then refined on simulation error by default, by an unbounded Levenberg-Marquardt search over every a, b and c of the CV with the exact Jacobian of the sensitivity recursion (12.11).

$$s_\vartheta[k] = \sum_{l=1}^{n_a} a_{i,l}\, s_\vartheta[k-l] + r_\vartheta[k] \tag{12.11}$$

where $s_\vartheta = \partial\hat{\delta}_i/\partial\vartheta$, r_ϑ is the regressor of coefficient ϑ in the free run ($\hat{\delta}_i[k-l]$ for $a_{i,l}$, the lagged input for a b coefficient, 1 for c_i), and the history before the record is the derivative of the starting value, zero for a measured start. Tests check it against differences for both starting rules. The search runs for at most 30 iterations at a cost tolerance of 10^{-7}, and its result is kept only if the training free-run error does not rise and no pole moves more than $1/N_s$ outside the circle. On the bundled records it converged in 3 to 24 iterations per CV. A pole within T_s/S of the circle makes a path integrating, with the ramp slope $B_{ij}(1)/(T_s\tilde{A}_i(1))$ where $A_i(q) = (1-q^{-1})\tilde{A}_i(q)$ and $\tilde{A}_i(1) = \sum_l l\, a_{i,l}$. ARX alone also reports one-step and h-step predictions ($h = 10$ by default), each labeled with the measured-output advantage it carries.

12.4.4 FIR with Penalties Chosen on Validation

The FIR regression (12.3) is formed in normalized units, with inputs divided by $\sigma_{u,j}$ and the CV by $\sigma_{y,i}$, accumulated as normal equations, and solved by Cholesky with the penalty of Equation (12.12).

$$\min_h \|\delta_i - Xh\|^2 + \lambda\bar{s}\left(0.1\|h\|^2 + 10\|D_2 h\|^2\right) \tag{12.12}$$

where D_2 takes second differences $h_l - 2h_{l+1} + h_{l+2}$ within each path, $\bar{s}$ is the mean diagonal of $X^\mathsf{T}X$ so that λ is dimensionless, and λ is chosen from $10^{-6}, 10^{-5}, \ldots, 1$ by the lowest validation mean-square error. Every regressor is a shifted copy of one input, so each entry of $X^\mathsf{T}X$ is a lagged cross product. A sum then follows from its predecessor on the same diagonal with two end corrections per run of equally weighted rows, and only the first sum of each diagonal takes a full pass, which turns about $Np^2/2$ multiplications into about Nmp for m inputs. When per-sample weights break the training rows into runs shorter than the horizon on average, the rows are accumulated one by one instead. The horizon is $N/(3m)$ samples, clipped to between 10 and $\min(240, 720/m)$, which holds the coefficient count below about 720 per CV so the normal equations stay interactive. A FIR fit notes a path whose step response still moves by more than 10 % of its largest value over the last fifth of the horizon, because its gain is then a truncated sum. A zero path's coefficients are simply absent from the regressors.

12.4.5 Subspace State Space with Prediction-Error Refinement

The state-space fit proceeds in four steps. Each input first receives one delay d_j, the shortest refined delay d_{ij} over the CVs it drives less the one sample a strictly proper model already has. A past-output MOESP step then stacks future inputs, past inputs, past outputs and future outputs of every training run into block Hankel rows with horizon i, each channel scaled by its training standard deviation, and an LQ factorization (a streamed QR) isolates the block that relates past data to future outputs after the future inputs are projected out [10]. A one-sided Jacobi SVD of that block suggests the order, and $\Gamma = U_1 S_1^{1/2}$ gives C from its first block row and A from shift invariance by least squares.

An integrating mode lands on either side of the unit circle by noise, so when the spectral radius ρ of A exceeds 1, A is scaled by $1/\rho$, which puts that mode on the circle and moves the others by the same small factor. With A and C fixed, B, D and the offsets c are linear in the free-run output and are solved jointly by least squares on every training sample of the block, using the state-sensitivity recursion $S_j \leftarrow AS_j + Iu_j[k]$ from its steady state. A block whose slowest mode has $|z| \geq 1 - 5/N_s$ starts instead from Equation (12.13).

$$x_0 = G\,(y_N - Z_N \vartheta)\,, \qquad G = \left(\Gamma_N^{\mathsf{T}} \Gamma_N\right)^{-1} \Gamma_N^{\mathsf{T}} \tag{12.13}$$

where y_N stacks the block's first $N = 2n$ measured outputs, Γ_N stacks $C, CA, \ldots, CA^{N-1}$, and $Z_N \vartheta$ is the response of the unknowns $\vartheta = (B, D, c)$ over those samples from zero state, so the output stays linear in ϑ. Such a block also receives a constant state drift, solved as the B column of a virtual input fixed at 1, for the integrated mean of its inputs that an output offset cannot carry. Orders 1 through $\min(8, il - 1)$ are scored on validation free run, an order with a pole more than $1/N_s$ outside the circle is rejected, and the 1 % parsimony rule of ARX picks the order, while the largest log gap among the singular values is only displayed. A block of order n with m inputs and l outputs carries $n^2 + nm + ln + l$ parameters, plus n for a drift.

The chosen model is then refined on the weighted free-run error of the training samples over every entry of A, B, C, the offsets, a drift and, when enabled, D, with 25 Levenberg-Marquardt iterations at a tolerance of 10^{-8}. Trial models with spectral radius of at least 1.2 are rejected inside the search, and the result is kept only if the training error does not rise and no pole lies more than $1/N_s$ outside the circle. The refinement is skipped and reported when the parameter count n_p and residual count N_r make $n_p N_r (n^2 + nm + 1)$ exceed 1.5×10^9 or $N_r n_p^2$ exceed 10^9, which is the interactive cost budget.

12.4.6 Exact Zero Paths in Every Family

A zero entry $M_{ij} = 0$ removes the entire deterministic path, delayed effects and feedthrough included, from every candidate family. In the transfer-function families the path and its parameters leave the problem. In ARX and FIR its coefficients are not regressors, and the remaining coefficients of that CV are re-estimated. A state-space model cannot be treated by deleting one entry of B, C or D, because the input can reach the output through shared states, and the constraint that matters is on the whole transfer entry, as shown in Equation (12.14).

$$D_{ij} = 0, \qquad C_i A^k B_j = 0 \quad \text{for } k = 0, 1, \ldots, n-1 \tag{12.14}$$

where C_i is row i of C and B_j is column j of B. By the Cayley-Hamilton theorem the first n Markov parameters vanishing makes every later one vanish in exact arithmetic, which is the realization argument of Ho and Kalman applied to one entry [8]. The application satisfies (12.14) by construction. Outputs are grouped by their set of allowed inputs, each group receives its own block of states driven only by those inputs, and the full model is block diagonal, so a prohibited input has an identically zero column in the block that its CV reads. One group, with no zeros, is the fully coupled model. The price is that dynamics shared across blocks are not modeled and the total order can rise, which the candidate's notes state.

🕮 **Pause.** Try Exercise 12.2 before reading on.

Every zero path is then verified end to end rather than by inspecting a coefficient. The whole fitted model is simulated with every input at its baseline, again with a unit impulse at $k = 1$ on input j, and again with a unit step from $k = 1$, over $\max(1000, 4n_r)$ samples where n_r is the response horizon, and the CV's trajectories must agree exactly, to 0.0. For state space the Markov parameters of (12.14) and D_{ij} are checked as well. A zero that fails is labeled as a violated constraint and the candidate becomes ineligible for ranking and export. Each matrix cell carries one of nine statuses in words and an icon, among them fixed at zero, small estimated gain (an effect at one input standard deviation below 2 % of the CV standard deviation), cannot estimate (an input with fewer than two moves in training), integrating, unstable, failed and out of date, so a small gain is never displayed as a zero. By default each cell draws its unit-step response in its own frame, with the vertical extent and the time span written in the cell and the span running until that response settles, so a path with a small gain or a fast response is not a flat line in a corner. A shared scale per CV, a scale in CV standard deviations and a shared horizon remain as view settings.

12.4.7 Initialization, Ranking and Uncertainty

Every free run starts from the steady state for the first inputs of its segment, held before the record begins, unless the model is too slow to show one. An ARX output or a state block whose slowest pole has $|z| \geq 1 - 5/N_s$, a time constant beyond a fifth of the longest segment, starts from its measurements, an ARX history at its first measurement and a state block at (12.13). The formal steady state of such a model is an extrapolation far beyond the record and was one reason the first release's ARX and state-space models drifted on the surge tank. The flag is set at fit time and stored in the model, so the export applies it too. An ARX with an exact unit root also holds its first measurement, and an unmarked block with $I - A$ singular starts at zero. This one convention makes the five families comparable and is what the GEKKO export reproduces. The first 3 samples of each segment are not scored.

Candidates are ranked only if they are current, meaning computed from the present data, splits and mask, without a pole more than $1/N_s$ outside the unit circle (an integrator counts as stable, and FIR has no poles) and with every zero verified exactly. Among those, the lowest aggregate validation NRMSE of the free run leads, and a candidate within 2 % of the best with fewer parameters is preferred. Two baselines frame every ranking, persistence (the first measurement held, or the previous one for one-step prediction) and a static regression $\hat{y}_i = \bar{y}_i + c_i + \sum_j K_{ij} u'_j$ under the same mask. Residual autocorrelation and residual-input cross-correlation over 25 lags are computed within runs of one split and never across a gap, against the $\pm 1.96/\sqrt{N}$ band.

Uncertainty comes from a moving-block bootstrap, because residuals of a dynamic model are autocorrelated and resampling single samples would treat them as independent draws [18]. The free-run residuals are cut into blocks of $\max(10, \mathrm{round}(2N^{1/3}))$ consecutive samples, which is 19 for the 864 training samples of the TCLab record. Blocks are drawn only from the training stretches of each segment, a source stretch chosen in proportion to its length and a start chosen so the block stays inside it, and they replace the residual on training samples only, so no validation or test residual enters a replicate. Each replicate refits the fitted model's own structure, with each CV's ARX orders and delays, each CV's FIR horizon and chosen penalty, each state block's order, the input delays and the mask held at their fitted values. Transfer-function lags and dead times are re-estimated, starting from the fitted values. The default is 30 replicates with a fixed seed, and the 5th, 50th and 95th percentiles of every gain and the pointwise 5th and 95th percentiles of every step response are reported. The application's test of this rule adds a sinusoid of amplitude 500 to the validation and test outputs of a single-path record with a true gain of 2 and requires the 5th to 95th percentile gain interval of 12 replicates to stay between 1.8 and 2.2, which the ARX, FIR, state-space and first-order families all satisfy.

12.4.8 Worker Protocol and Stale Results

Fitting runs in a browser module worker. The prepared data are posted once with a data key, and each job carries only settings, the mask and a revision key. A result is published only if both keys still match the interface state, so a fit that finishes after a newer edit can never be shown as current, and the last valid result of an older revision is displayed as out of date and is excluded from ranking and export. Starting a job cancels the one running, and cancellation terminates the worker and starts a fresh one with the data re-posted, so nothing from a cancelled job can arrive later. A structural edit saves an undoable revision, marks every candidate out of date and refits the family on screen after a 650 ms pause, which batches rapid edits.

12.4.9 Export Realization and GEKKO Adapters

Every family is written as one discrete realization $x[k+1] = Ax[k] + Bu'[k]$, $y[k] = y_0 + Cx[k] + Du'[k]$ with typed states. Dynamic states carry lags, oscillations and ARX output history, register states form shift registers that realize whole-sample dead times exactly, and a constant state fixed at 1 carries the ARX constant term and the drift of a measured-start state block. Registers start at the first input, the constant at 1, measured-start states as in Section 12.4.7, and the remaining dynamic states at the solution of $(I - A_{dd})x_d = A_{d,o}x_o + B_d u'[0]$, with the fallbacks described there. A measured-start block is solved last, by Equation (12.13) with the response of every other state as the known part, and the realization carries the per-state flags and the per-block sample counts that this needs. The generated `model.py` needs only numpy and GEKKO. It simulates the realization by a numpy recursion and compares it with the application's free-run predictions at a tolerance of $10^{-8}\max(1, \sigma_{y,i})$, builds the same model in GEKKO with `remote=False` in simulation mode (`IMODE` 4, `NODES` 2, `RTOL` and `OTOL` 10^{-10}) with the MVs and DVs supplied as recorded values and never as decisions, compares GEKKO with numpy at $10^{-5}\max(1, \sigma_{y,i})$, recreates every step-response cell, checks that every zero path is exactly zero, and exits with an error status if any check fails [15].

Two GEKKO conventions were measured rather than assumed. The discrete `state_space` object takes the value assigned to each state at the first time point as $x[0]$ and does not evaluate the output equation at that point, so the script computes $y[0] = y_0 + Cx[0] + Du'[0]$ itself. The `arx` object implements Equation (12.15) with zero history before the first point.

$$y_i[k] = \sum_{l=1}^{n_a} a_{l,i}\, y_i[k-l] + \sum_g \sum_j b^{\mathrm{G}}_{i,g,j}\, u_j[k-g-1] + c^{\mathrm{G}}_i \tag{12.15}$$

The adapter maps the canonical coefficients by $b^{\mathrm{G}}_{i,g,j} = b_{i,j,g+1-d_{ij}}$, so row g of the GEKKO input matrix is lag $g+1$, and it works in deviation-from-start form, with inputs $u'_j[k] - u'_j[0]$, outputs measured from the output history y^s_i that the application starts from, steady or measured, and the constant $c^{\mathrm{G}}_i = c_i - (1 - \sum_l a_{i,l})(y^s_i - \bar{y}_i) + \sum_g \sum_j b^{\mathrm{G}}_{i,g,j} u'_j[0]$. Zero history in that form is the start of the application. An ARX with same-sample feedthrough uses the state-space route instead.

12.5 Predictive Control from the Fitted Model

A seventh step uses a chosen fit twice, as the internal model of a model predictive controller (MPC) and as the plant of a closed-loop simulation, and tunes a decentralized PI from the same fit for comparison. The dynamic-matrix form (DMC) predicts with the fitted step responses and a constant output bias, the form of most industrial packages [2], posed with input bounds as a quadratic program [19]. The state-space form predicts with the exported realization and a Kalman filter. Both minimize one objective under one set of bounds, so a difference between them is one of prediction and estimation. The objective has the form of Equation (1.29) in Chapter 1 without slack terms, and the state-space form's estimator is a steady-state Kalman filter with added disturbance states, as in Section 1.6.3.

12.5.1 Controller Model, Objective and Constraints

The controller works in the deviation variables u'_j and $y'_i = y_i - \bar{y}_i$ with the fit's record constants (ARX constant, state drift, offset) dropped, so rest is an equilibrium of every family. The first CVs in role order, as many as the MVs, receive setpoints y_i^{sp}, and the realization is pruned to the states that reach them, which keeps the 270 min dead time of the spurious pressure path (Table 12.5) out of the column's controller. The controller reads the fit, $\sigma_{y,i}$, the range each MV covered in the records and, at each sample, the measured CVs, setpoints and DVs, never a record's generator. In the nominal cases the simulated plant is the fit itself, so only the cases with the MV gains scaled by 0.7 and 1.5 or with another family's fit as the plant measure model error. At every sample k both forms choose input levels $v_{j,\nu}$ on move blocks $\nu = 1,\dots,N_b$ over a horizon of N_p samples, as shown in Equation (12.16).

$$\begin{aligned} \min_{v} \; & \sum_{\ell=1}^{N_p} \sum_{i\in\mathscr{C}} w_i \big(\hat{y}'_i[k+\ell] - y_i^{\mathrm{sp}}\big)^2 \\ & + \sum_{\nu=1}^{N_b} \sum_{j} R_j \big(v_{j,\nu} - v_{j,\nu-1}\big)^2 \\ \text{s.t. } & u_j^{\min} \le v_{j,\nu} \le u_j^{\max}, \qquad R_j = \lambda_u \textstyle\sum_{i\in\mathscr{C}} w_i K_{ij}^2 \end{aligned} \tag{12.16}$$

where $\mathscr{C}$ holds the controlled CVs, $w_i = 1/\sigma_{y,i}^2$ is the weight of Equation (12.8), $v_{j,0}$ is the input applied now, K_{ij} is the fitted gain and $\lambda_u = 1$ makes a move cost the weighted steady error it could remove. The twelve blocks last $1,1,1,1,2,2,3,4,6,9,15$ samples and then the remainder, and only the first is applied before a dense dual active-set solver repeats the problem at $k+1$. The default horizon $N_p = T_{\mathrm{set}}/T_s$, with T_{set} the slowest 2 % settling time of an MV-to-CV step response, was fixed before any loop was closed. The application declares no MV limits, rate limits or CV bands, so each MV moves inside the range it covered in the records, the evidence the fit rests on, and the PI is clipped to the same box. The TCLab FOPDT fit ($T_s = 5\,\mathrm{s}$) gives $T_{\mathrm{set}} = 790\,\mathrm{s}$, $N_p = 158$ and boxes of 0 to 79.4 and 0 to 77.2 % for Q_1 and Q_2, and the column fit ($T_s = 1\,\mathrm{min}$) gives $N_p = 85$ and 1.77 to 2.13 and 1.55 to 1.84 lb/min for reflux and steam. The present input is held whenever the solver status is not optimal or a planned level is not finite or leaves its box by more than 10^{-6} of the width.

12.5.2 Dynamic-Matrix and State-Space Predictions

The DMC predicts from the step-response coefficients $g_{ij}[\ell]$ of the fit, as shown in Equation (12.17).

$$\begin{aligned} \hat{y}'_i[k+\ell] &= \beta_i[k] + \sum_j \sum_{t<k} g_{ij}[k+\ell-t]\,\Delta u'_j[t] \\ &\quad + \sum_j \sum_{t=k}^{k+\ell-1} g_{ij}[k+\ell-t]\,\Delta u'_j[t] \\ \beta_i[k] &= y'_i[k] - \sum_j \sum_{t<k} g_{ij}[k-t]\,\Delta u'_j[t] \end{aligned} \tag{12.17}$$

where $\Delta u'_j[t]$ is the move of input j at sample t, planned only at block starts, g_{ij} is held at its last value beyond the model length N_m, 316 on the TCLab fit and 170 on the column, and $\beta_i[k]$ is the output bias. Measured DVs enter the first sum with their present move, and their future moves are zero. The prediction matrix is fixed, so the Hessian is formed once. The bias is the whole estimator of this form and assumes that the discrepancy seen now persists over the horizon. The DMC is refused for an integrating path, as on the surge tank.

The state-space form runs the pruned realization, with dead times as exact shift registers, augmented by one constant disturbance d_j per MV that enters where a held input of that MV reaches the dynamic states, as shown in Equation (12.18).

$$\begin{aligned} x[k+1] &= Ax[k] + B_u u'[k] + B_w w'[k] + B_d d[k] \\ d[k+1] &= d[k], \qquad y'_{\mathscr{C}}[k] = Cx[k] \end{aligned} \tag{12.18}$$

where w' holds the measured DVs, held over the horizon. A steady-state Kalman filter estimates the dynamic states and d with process variances of 10^{-6} of the mean CV variance per state and $10/\sum_i w_i K_{ij}^2$ per disturbance and measurement variances $\sigma_{y,i}^2$, converging in 253 Riccati iterations on the TCLab fit and 182 on the column. The disturbance estimate removes the steady offset of a model error or an unmeasured load [20], the mechanism Volume 9, Chapter 4 uses for a fuel bias. The realization has 27 states on the TCLab fit, 23 of them registers, and 26 on the column, and the form is refused for FIR fits and above 90 states. With the output bias selected in place of the filter both forms solve the same program, and they agree to 1.0×10^{-5} of an MV's box width on the TCLab fit and 4.3×10^{-5} on the column, the difference being the truncation of the step model.

12.5.3 Decentralized PI and Controller Export

The comparison is a decentralized PI without a decoupler, paired by the relative gain array [21] on the diagonal of both records ($\Lambda_{11} = 1.039$ and 2.019), the one-to-one pairing with positive relative gains closest to 1. Volume 6, Chapter 2 treats this pairing choice on a plant whose relative gain hides a one-way interaction. Each loop is tuned by SIMC [22] on the paired path's FOPDT equivalent, with $\theta_e = \theta + T_s/2$, $\tau_c = \theta_e$, $K_c = \tau/(K(\tau_c + \theta_e))$ and $\tau_I = \min(\tau, 4(\tau_c + \theta_e))$, which gives $K_c = 7.61\,\%/^\circ\mathrm{C}$ and $\tau_I = 104\,\mathrm{s}$ on T_1, $K_c = 12.0$ and $\tau_I = 94.9\,\mathrm{s}$ on T_2, and $K_c = 0.441$ and $-0.104\,(\mathrm{lb/min})/\mathrm{mol}\%$ with $\tau_I = 12.0$ and 14.3 min on x_D and x_B. Each loop runs in velocity form with the integral alone absorbing the clipping (Section 1.6.2), and one column row adds the static feedforward $-(K_{iF}/K_{ij})w'_F$ of the measured feed that the MPC predicts with.

A second export, `mpc.py`, carries the controller to Python. Its plain-Python checks confirm that the realization reproduces the DMC's coefficients, that the first move equals the optimum of the same quadratic program rebuilt in the script, with an active set on the bounds when one binds, and the SIMC gains. On the bundled fits these agreed to 2.0×10^{-14} relative (21 fits with a DMC) and 6.5×10^{-16} of the box width (all 25). With numpy and GEKKO the script also builds the controller on the `state_space` object in control mode (`IMODE` 6, output bias by `FSTATUS` 1) and compares its closed loop with the application's within 15 % of IAE [15], since GEKKO moves at every sample where the application holds blocks.

12.6 Measured Results

12.6.1 Five Records with Known Structure

Every bundled record is generated from a stated model with seeded noise, so every visitor sees the same data and every answer can be checked against the generator. The five records and the behavior each one tests are listed in Table 12.3. All results below use default settings, the default chronological splits and the full mask unless a zero is stated, and were produced by the application's own fitting pipeline in a measurement test that calls the same job functions as the worker.

Validation scores for all five families on all five records are collected in Table 12.4. No family wins everywhere. The first-order matrix ranks first on the two-heater and column records, the state-space model on the tank and the stage, and the first-order matrix again on the exchanger, where three candidates fall within 2 % of the best (0.0939 and 0.0940 twice) and parsimony selects 7 parameters over 9 and 19. Every candidate beats both baselines on every record. The one-step persistence baseline reaches 0.0260 on the TCLab record, close to the best free-run model, which is why a one-step score is never used for ranking. Without its simulation-error refinement ARX scores 0.0643, 0.1208, 1.546, 0.0355 and 0.0971 in table order, so the refinement helps on four records and costs 0.008 on the exchanger, which was recorded under feedback.

🕮 **Pause.** Try Exercise 12.1 before reading on.

Fit times on the measurement host ranged from 4 ms (first order, exchanger) to 771 ms (state space, column), and the first-order matrix of the column took 59 ms. The FIR candidate reuses the preliminary FIR fit that supplies every family's delays and starting points. Every first-order CV fit and eight of the nine second-order ones stop before the iteration limit. The exception is the column's

Table 12.3 Bundled Records. Samples Are Training, Validation and Final Test on the Uniform Grid.

Record	Inputs, CVs	T_s, samples	Generator and purpose
TCLab	2 MV, 2 CV	5 s, 864/288/289	radiative two-heater balance, 12 s sensor lag, 0.32 °C quantization, weak real cross paths
Column	2 MV, 1 DV, 3 CV	1 min, 900/300/300	Wood-Berry matrix plus pressure with no reflux path and a weak feed path, two campaigns
Surge tank	1 MV, 1 DV, 1 CV	2 s, 1080/360/360	integrating level of a 12 m^3 tank, no steady state
Stage	1 MV, 1 DV, 2 CV	0.02 s, 1800/600/600	table at 1.6 Hz with $\zeta = 0.15$, 40 ms transport
Feedback	1 MV, 1 DV, 1 CV	0.5 min, 864/288/288	exchanger under PI control with three setpoint steps and a drifting unmeasured disturbance

Table 12.4 Free-Run Validation NRMSE by Family and Record, With Estimated Parameter Counts in Parentheses. Bold Marks the Candidate Ranked First.

	TCLab	Column	Tank	Stage	Feedback
FOPDT	**0.0232** (14)	**0.1093** (30)	0.396 (7)	0.1203 (14)	**0.0940** (7)
SOPDT	0.0275 (18)	0.1096 (39)	0.125 (9)	0.0254 (18)	0.0940 (9)
ARX	0.0523 (20)	0.1146 (36)	0.190 (4)	0.0057 (16)	0.1053 (6)
FIR	0.0283 (574)	0.1166 (894)	4.386 (359)	0.0060 (962)	0.1216 (287)
SS	0.0295 (14)	0.1112 (75)	**0.062** (46)	**0.0056** (23)	0.0939 (19)
static	0.622	0.660	5.960	0.405	0.316
persistence	2.251	1.057	6.258	1.039	0.784

pressure CV, whose absent reflux path runs to the lower damping bound and whose search stops at 60 iterations, 84 with scouting. The final-test NRMSE of each first-ranked candidate is 0.0267 (TCLab), 0.1166 (column), 0.196 (tank), 0.0056 (stage) and 0.0651 (exchanger), the lowest final-test score on four records. On the exchanger, where the first three candidates lie within 0.0001 on validation, ARX, ranked fourth, has the lowest at 0.0619. On the tank the final-test NRMSE of every family that follows the level is 2.8 to 3.5 times its validation score, because an integrating level drifts further the longer the free run continues.

12.6.2 Recovering the Wood-Berry Matrix

The column record is the most direct check of credit assignment, because three inputs move independently but overlap in time and two of the paths into each composition have opposite signs. Table 12.5 compares the first-order matrix with the generator. All six Wood-Berry gains are recovered within 0.9 %, the dead times of six of the seven strong paths within 0.10 min, and every strong-path time constant within 8 %. The weak feed path into pressure is where the record runs out of information. Its gain of 0.0349 against 0.035 is close, but its dead time of 3.51 min against 2 min and lag of 5.69 min against 9 min trade against each other, because the full 0.5 lb/min range of feed moves pressure by at most 0.0175 bar against 0.004 bar of noise.

The feed dead time into distillate composition, 8.87 min against 8.1 min, is the only strong-path dead time off by more than 0.1 min, and its lag is 1.14 min short, so the two errors partly cancel in the delayed rise. The absent reflux path into pressure is estimated at 0.0141 bar/(lb/min) with a lag of 114.6 min and a dead time on its upper bound of 270 min, half the longest training stretch. That is a physically meaningless shape, and its steady effect over the full 0.4 lb/min reflux range would be 0.006 bar against 0.004 bar of noise. That entry is what a structural zero replaces.

Table 12.5 First-Order Matrix of the Column Record Against the Generator. Gains in mol%/(lb/min) for x_D and x_B and bar/(lb/min) for P, Times in Minutes.

Path	K true	K fit	τ true	τ fit	θ true	θ fit
$x_D \leftarrow R$	12.8	12.69	16.7	16.80	1	1.00
$x_D \leftarrow S$	−18.9	−18.74	21.0	20.81	3	3.10
$x_D \leftarrow F$	3.8	3.78	14.9	13.76	8.1	8.87
$x_B \leftarrow R$	6.6	6.64	10.9	10.93	7	7.00
$x_B \leftarrow S$	−19.4	−19.43	14.4	14.31	3	3.06
$x_B \leftarrow F$	4.9	4.91	13.2	12.96	3.4	3.36
$P \leftarrow R$	0	0.0141	—	114.6	—	270.0
$P \leftarrow S$	0.35	0.348	6.0	5.89	1	1.00
$P \leftarrow F$	0.035	0.0349	9.0	5.69	2	3.51

Table 12.6 Validation NRMSE Before and After a Structural Zero, With Parameter Counts. Every Zero Verified Exactly in Every Family.

	Column, $P \leftarrow R$ absent path	Column, $x_D \leftarrow S$ strong path	TCLab, $T_1 \leftarrow Q_2$ weak real path
FOPDT	0.1093→0.1085 (27)	0.1093→0.3892 (27)	0.0232→0.0853 (11)
SOPDT	0.1096→0.1091 (35)	0.1096→0.3913 (35)	0.0275→0.0867 (14)
ARX	0.1146→0.1145 (35)	0.1146→0.3989 (27)	0.0523→0.1081 (16)
FIR	0.1166→0.1154 (795)	0.1166→0.3914 (795)	0.0283→0.1075 (431)
SS	0.1112→0.1128 (57)	0.1112→0.3994 (31)	0.0295→0.0838 (33)

TRY IT IN THE APP 3 · A Path That Is Not There

On 1 · Import load the Distillation column example, and on 3 · Identify fit First order + dead time with the defaults. Open the Pressure ← Reflux cell and watch its dead time sit on the upper bound of 270 min. Press Set response to zero: the cell reads Fixed at zero, verified exactly, and the validation NRMSE falls slightly, from 0.1093 to 0.1085.

12.6.3 Structural Zeros, True and False

Three structural edits were applied to every family, and the results are collected in Table 12.6. Setting the absent pressure path to zero removes 3 parameters from the first-order matrix and moves its validation NRMSE from 0.1093 to 0.1085, and three other families improve slightly, by 0.00002 (ARX) to 0.0013 (FIR). The rows of distillate and bottoms composition are identical before and after in the transfer-function, ARX and FIR families, because those families fit each CV separately, and only the pressure row is re-estimated, with the steam gain moving from 0.3477 to 0.3482. Figure 12.5 captures the application immediately after this first-order refit. The state-space model is the exception that the method predicts. The zero splits its single 6-state block into a 5-state block for the two compositions and a 1-state block for pressure, so the parameter count falls from 75 to 57, but the composition block no longer shares its states with pressure, its distillate row rises from 0.1007 to 0.1090, and the correct zero costs this family 0.0016 of validation NRMSE. Figure 12.3 draws the scores of Table 12.6 as bars on one scale.

Removing the strong steam path into distillate composition is a false structural claim, and every family reports it. The first-order validation NRMSE rises from 0.1093 to 0.3892, and the distillate row alone rises from 0.0996 to 0.9394, close to the CV's own training standard deviation. The refit also shows what credit reassignment looks like. With steam gone, the reflux gain into distillate climbs from 12.69 to 19.47 and its lag from 16.8 min to 42.7 min, and the feed gain from 3.78 to 5.23, because the remaining paths are refitted to absorb as much of the missing steam response as their own movement in this record permits. A fitted gain is therefore a property of the model structure as

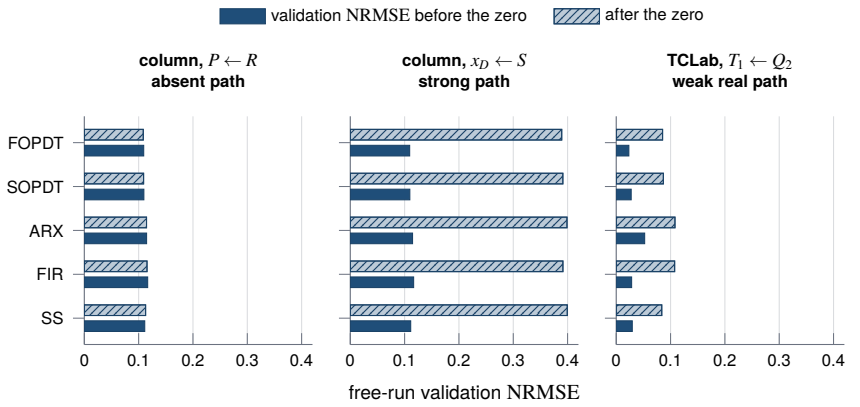


Figure 12.3 Free-run validation NRMSE of every family before (solid) and after (hatched) a structural zero, from Table 12.6, on one scale. Notice that the absent path changes almost nothing, while removing the strong steam path raises the error of every family to between 0.3892 and 0.3994, and removing the weak but real heater path raises it in every family, by a factor of 3.7 for the first-order matrix.

well as of the plant.

🕮 **Pause.** Try Exercise 12.3 before reading on.

The weak heat path from heater 2 to sensor 1 on the TCLab board is real, with a fitted gain of 0.079 °C/% against 0.512 °C/% for the direct path, a dead time of 52.6 s against 10.9 s and a lag of 181 s against 104 s. Setting it to zero raises the first-order validation NRMSE from 0.0232 to 0.0853, a factor of 3.7, and the T_1 row from 0.0299 to 0.1541, while the T_2 row is unchanged at 0.0165. The direct gain rises from 0.512 to 0.542 °C/% and its dead time falls from 10.9 to 4.8 s as the model tries to explain heat that arrives late through the board with heat that arrives early. The generator contains no transport delay at all, so both identified dead times are the apparent delay of a 12 s sensor lag and the transistor's own thermal capacity seen through a first-order model. In the state-space family this zero raises the parameter count from 14 to 33, by splitting the 2-state model into blocks of 1 and 4 states.

✦ **DID YOU KNOW? Removing a Path Costs Little or a Lot**

From the simulator: fixing the absent reflux path into column pressure at zero moves the first-order validation NRMSE from 0.1093 to 0.1085. Fixing the weak but real path from heater 2 to sensor 1 at zero raises it from 0.0232 to 0.0853, a factor of 3.7 (Table 12.6).

12.6.4 Oscillation, Integration and Feedback

The stage record separates the families by model shape. Its table position rings at 1.6 Hz with $\zeta = 0.15$, and the first-order matrix cannot represent that, scoring 0.1203 on validation against 0.0254 for second order. The second-order position path returns $\tau = 0.100$ s against the 0.0995 s of the generator's natural frequency and $\zeta = 0.196$ against 0.15, with a gain of 0.589 mm/V against 0.6, while both motor-speed paths run to the upper damping bound and are reported as effectively first order. The state-space model, at order 3 with 23 parameters, ARX with 16 coefficients and FIR with 962 all reach 0.0056 to 0.0060, likely because the drive enters the table through a motor lag and a transport delay that neither transfer-function shape describes exactly. Without its simulation-error refinement ARX scores 0.0355 on this record.

The surge tank has no steady state. The first-order fit drives both lags to 6049 and 6480 s, the second at its upper bound of three training spans, and the application reports both paths as integrating

Table 12.7 Maximum Absolute Errors Printed by Each Exported `model.py`, Outputs Outlet and Jacket. Step Columns Are the Worst Matrix Cell.

Family	numpy vs app	GEKKO vs numpy	step, numpy	step, GEKKO
FOPDT	7.1e-15, 7.1e-15	4.7e-10, 5.0e-11	5.6e-16	5.0e-12
SOPDT	2.6e-13, 4.3e-13	4.9e-10, 5.0e-11	7.8e-16	5.0e-12
ARX	2.1e-14, 7.1e-15	4.9e-10, 5.0e-11	6.7e-16	5.0e-12
FIR	1.4e-14, 7.1e-15	4.9e-10, 5.0e-11	3.3e-16	5.0e-12
SS	1.4e-14, 1.4e-14	4.8e-10, 5.0e-11	7.8e-16	9.3e-12

with initial slopes of 0.0013861 and −0.0022985 %/s per unit input against the generator's 0.0013889 and −0.0023148, errors of 0.20 % and 0.70 %. The second-order fit reaches 0.125 with slopes of 0.0013904 and −0.0023083. ARX and state space now hold the integrator itself. The chosen ARX, with $n_a = n_b = 1$ and 4 coefficients, has its pole at $|z| = 1.0003$, within $1/N_s$ of the circle, starts from the first measured level and scores 0.190, with slopes 8 to 10 % low. The subspace step placed the slowest mode at $|z| = 1.0018$, the scaling moved it onto the circle, and the 5-state block, with its measured start and drift and 46 parameters, follows the level at 0.062, the best score of any family, with slopes of 0.0013884 and −0.0022376. The FIR model alone cannot represent an integrator within its horizon and scores 4.386, and even the best model reaches 0.196 on the final test as its slope error accumulates.

The exchanger record was taken under PI control with $K_c = 0.9\,\%/^\circ\mathrm{C}$ and $\tau_I = 8\,\mathrm{min}$, three setpoint changes and a drifting unmeasured disturbance, and the readiness summary flags it as closed loop. The setpoint steps supply external excitation, and the valve gains agree across families at 0.911 (first order), 0.896 (ARX) and 0.911 (state space) against the generator's 0.9 °C/%, with a first-order lag of 5.97 min and dead time of 1.46 min against 6 and 1.5 min. The inlet-temperature path is recovered at 0.446 against 0.45. ARX is the one family that the refinement hurts here, at 0.1053 against 0.0971 for its equation-error estimate with a valve gain of 0.907, which is why the app's guide advises switching the refinement off for a record taken under feedback. The FIR model is the outlier, at 0.856 and 0.246, and it also has the worst validation score of 0.1216, so the agreement among the parametric families is the check that the record supports.

12.6.5 Export Parity

The export was verified on a two-input, two-output test system with a zero path, a negative second-order path and fractional dead times, with every family exported and its `model.py` run under numpy 2.5.4 and GEKKO 1.3.2. Its state-space model has a 6-state block for the jacket with a measured start and a drift, so the starting rule and the constant state are tested as well. Before export, the realization of every family reproduces the application's free run and step responses to below 10^{-9}. The errors reported by the scripts themselves are listed in Table 12.7. The numpy recursion matches the application to at most 4.3×10^{-13}, which is double-precision round-off on outputs of order 20 to 40, GEKKO matches numpy to at most 4.9×10^{-10}, consistent with its 10^{-10} relative tolerance, and the zero path is exactly 0.0 in both numpy and GEKKO for every family.

12.6.6 Scale

The specification called for an interactive benchmark of 50 000 samples with 8 inputs (5 MVs and 3 DVs) and 4 CVs. A synthetic record of that size, with first-order paths, dead times of 0 to 4 samples, six true zero paths left estimated, and noise of 0.05, was fitted family by family with default settings in Node.js 26.12 on the development host. The results are given in Table 12.8. Every family finishes in between 2.5 s and 15.0 s, and a second run on the same host agreed to within 0.4 s per family. Every family but FIR includes the preliminary FIR fit, which alone takes the 2.5 s of the FIR row, and the delay scan. The resident memory of the whole test process stayed between 323 and 370 MB. Most of the ARX time is its simulation-error refinement, which the sensitivity recursion (12.11) keeps to

Table 12.8 Fit Time and Validation NRMSE on 50 000 Samples, 8 Inputs and 4 CVs

Family	Time (s)	Parameters	Validation NRMSE
FOPDT	8.9	100	0.0379
SOPDT	15.0	132	0.0379
ARX	7.7	139	0.0615
FIR	2.5	2884	0.0390
SS	7.0	164	0.0539

Table 12.9 Closed-Loop Comparison on the FOPDT Fits at Default Settings. IAE in °C s and mol% min, Interaction in °C and mol%, Overshoot of the Stepped CV in Percent, Input Travel $\sum|\Delta u|$ in Percent and lb/min. The Mismatch Rows Repeat the First Setpoint Step, and DMC, SS Marks Rows Where the Two Forms Agree to Three Digits. Interaction Is the Largest Absolute Deviation of the CV Whose Setpoint Holds.

		TCLab two-heater fit, $T_s = 5$ s						Column fit, $T_s = 1$ min					
Case	Controller	IAE T_1	IAE T_2	Inter.	Ovs.	Q_1	Q_2	IAE x_D	IAE x_B	Inter.	Ovs.	R	S
T_1 or x_D step	PI	347	29.6	0.344	1.5	49.2	8.36	4.10	5.19	0.475	8.0	0.557	0.139
	DMC, SS	392	7.51	0.137	2.1	47.8	5.73	5.49	0.978	0.0739	0.17	0.156	0.0633
T_2 or x_B step	PI	38.9	335	0.387	3.7	5.94	58.4	2.03	7.00	0.142	0	0.166	0.134
	DMC, SS	6.10	366	0.121	1.9	4.54	56.6	0.573	4.92	0.0413	0.20	0.0828	0.0645
Load or feed	PI	217	9.33	—	—	17.4	4.03	0.667	4.59	—	—	0.0682	0.0496
	PI + FF	—	—	—	—	—	—	1.46	1.37	—	—	0.218	0.0788
	DMC	372	72.3	—	—	16.7	1.18	0.414	0.890	—	—	0.0169	0.0428
	SS	132	14.4	—	—	28.0	3.01	0.414	0.890	—	—	0.0169	0.0428
Residual	PI	138	114	—	—	182	256	22.5	21.3	—	—	11.5	2.17
	DMC	156	129	—	—	37.6	43.6	20.8	16.1	—	—	1.33	0.594
	SS	161	134	—	—	64.6	70.1	22.1	17.3	—	—	1.54	0.850
Gain ×0.7	PI	494	42.3	0.293	0	40.5	7.00	4.98	7.20	0.399	3.7	0.425	0.0819
	DMC	548	4.61	0.0961	0	40.3	5.99	7.78	1.18	0.0546	0.03	0.171	0.0673
	SS	548	6.59	0.0734	4.0	41.8	6.64	7.00	1.34	0.0706	2.1	0.203	0.0835
Gain ×1.5	PI	272	24.3	0.434	9.8	65.6	14.3	6.05	10.9	0.593	20	1.12	0.436
	DMC	337	15.0	0.198	14	60.0	5.58	4.29	1.24	0.110	7.2	0.164	0.0562
	SS	337	23.7	0.261	2.4	75.7	9.85	4.66	0.926	0.0734	0.47	0.183	0.0566
SS fit as plant	PI	318	28.2	0.181	0.37	47.0	3.71	3.81	5.82	0.409	9.4	0.516	0.120
	DMC	383	17.7	0.135	0.94	47.0	6.96	5.98	1.21	0.0677	2.0	0.161	0.0679
	SS	390	16.5	0.158	3.1	47.8	8.36	6.01	1.13	0.0743	3.0	0.176	0.0730

one filtered pass per coefficient. ARX still trails the transfer functions here, at 0.0615, likely because each CV shares one output polynomial of order at most 3 among eight inputs whose true lags range from 8 to 26 samples. The state-space refinement was skipped for all four CVs by the cost budget and reported, which explains its validation score of 0.0539 against 0.0379 for the first-order matrix. The FIR horizon was capped at 90 samples by the 720-coefficient rule, giving 2884 coefficients.

12.6.7 Predictive Control Against Decentralized PI

The controllers were compared on the FOPDT fit of the two records with two MVs, ranked first on both, at the defaults ($N_p = T_{\text{set}}/T_s$, $\lambda_u = 1$, Kalman filter). Each case starts at rest and runs for two settling times, 316 samples on the TCLab fit and 170 on the column, or for a replayed stretch, with one plant, box and case for all three controllers. A setpoint step is the smaller of one training standard deviation and the change whose steady inputs use half of the MVs' room in the box, 7.66 °C on T_1, 7.04 °C on T_2, 0.673 on x_D and 0.488 mol% on x_B. This rule replaced one standard deviation after the first column run held both MVs at a bound for 89 to 99 % of the samples, and the interaction finding below also holds under the original rule (0.587 under PI against 0.242 mol%). The disturbances are an unmeasured step of 15.9 %, a fifth of Q_1's recorded range, on the TCLab fit and the measured feed record of the validation stretch on the column. The validation residual is replayed on every CV, and the first step is repeated on plants with the MV gains at 0.7 and 1.5 times the fit's and on the state-space fit, as collected in Table 12.9.

Both MPC forms hold the CV whose setpoint stays far closer than the PI, which has no decoupler,

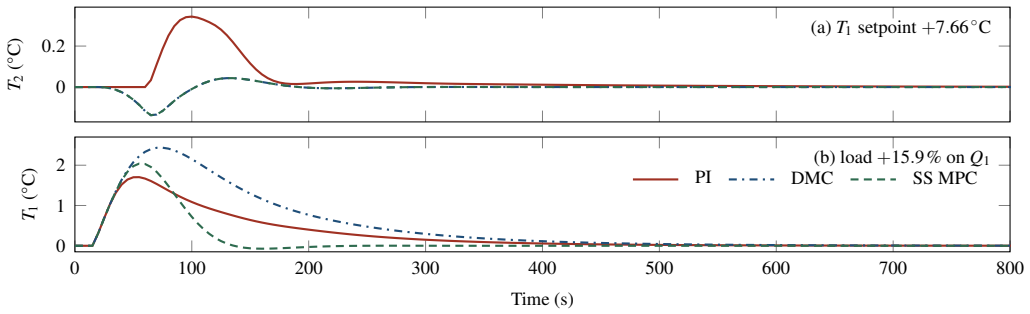


Figure 12.4 Deviation of the CVs of the TCLab FOPDT fit under the decentralized PI (solid), the DMC (dash-dotted) and the state-space MPC with its Kalman filter (densely dashed). (a) T_2 while the T_1 setpoint steps, with a largest deviation of 0.344 °C under PI and –0.137 °C under both MPC forms, whose traces coincide. (b) T_1 after an unmeasured heater load, with an IAE of 217, 372 and 132 °C s over the whole 1575 s case, of which the first 800 s are drawn.

at 0.137 °C against 0.344 °C on T_2 when T_1 steps and 0.0739 against 0.475 mol% on x_B when x_D steps, as shown in Figure 12.4(a), while moving the other MV less (5.73 % of Q_2 against 8.36 %). On the stepped CV the PI is faster in three of four cases, 347 against 392 °C s on T_1, 335 against 366 on T_2 and 4.10 against 5.49 mol% min on x_D, while the MPC is faster on x_B, 4.92 against 7.00. The move suppression prices moves the SIMC loops make freely, and at $\lambda_u = 0.1$, a counterfactual outside the table, both forms beat the PI on T_1 (342) and x_D (3.24). At $\lambda_u = 10$ the T_1 IAE rises to 598 °C s, and at each of the three settings 0.1, 1 and 10 the interaction stays below the PI's.

The DMC loses the unmeasured load, with an IAE on T_1 of 372 °C s against 217 for the PI and 132 for the Kalman form and a peak of 2.43 °C against 1.70 and 2.04, as shown in Figure 12.4(b). The load reaches T_1 through a 104 s lag, at 0.682, 2.55 and 4.67 °C after 20, 50 and 100 s of a final 8.13 °C in open loop, while the bias of Equation (12.17) predicts that each discrepancy stays as it is. Two counterfactual runs place the loss in the estimator. The state-space form with the output bias gives the same 372, and when the same disturbance arrives as an output step, for which a constant bias is the correct model, the DMC reaches 396 against 373 and returns within 10 % of its peak first, at 80 against 115 s. On the column both forms predict the measured feed and reach 0.414 and 0.890 mol% min against 0.667 and 4.59 for the PI. Static feedforward helps the PI on x_B (1.37) but harms x_D (1.46), likely because feed reaches x_D 7.9 min after reflux does (Table 12.5).

> **TRY IT IN THE APP** **13 · MPC From the Fit**
>
> With the TCLab two-heater test fitted by First order + dead time, open 7 · MPC at the defaults and wait for the table. In Setpoint step on T1 temperature, read the Interaction column: 0.344 °C under PI and 0.137 °C under both predictive forms. Then compare the T_1 IAE in Unmeasured load on Q1 heater: 217 for the PI, 372 for the DMC and 132 °C s for the state-space form.

On the replayed residual the PI has the lowest IAE on the TCLab fit, 138 and 114 against 156 and 129 for the DMC, with five to six times the DMC's input travel, and the DMC the lowest on the column. At 1.5 times the gain the Kalman form overshoots least, 2.4 % against 9.8 for the PI and 14 for the DMC on T_1 and 0.47 % against 20 and 7.2 on x_D, and the MPC interaction stays below the PI's in all six mismatch rows. All 444 runs over the 25 bundled fits stayed finite and inside the box with no plan held. In Node.js on a shared host a DMC decision took 18 µs on average and a state-space decision 2.0 ms on the TCLab fit and 1.0 ms on the column. At half the sample period these findings hold (interaction 0.121 against 0.309 °C, load IAE 297 for the DMC against 184 and 79), and only the order of the two MPC forms where they agree to three digits reverses.

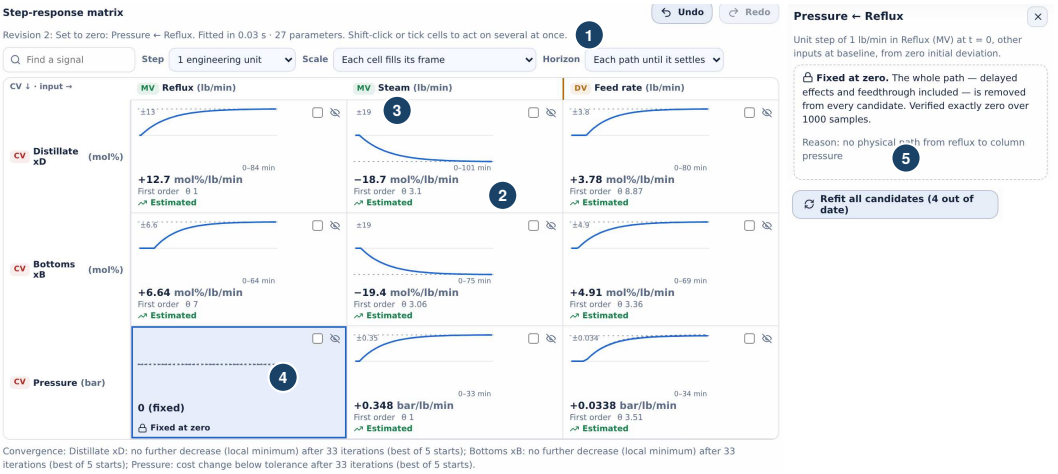


Figure 12.5 The Identify step on the distillation column record with the first order plus dead time candidate after Pressure ← Reflux was set to zero with the reason "no physical path from reflux to column pressure", shown as a composite of three parts of the screenshot: the step-response matrix at left, and the path inspector's report and the refit button at right. (1) The matrix shows revision 2, refitted in 0.03 s with 27 parameters. (2) Each cell prints its gain, 12.7, −18.7 and 3.78 into distillate, 6.64, −19.4 and 4.91 into bottoms, and 0, 0.348 and 0.0338 into pressure. (3) Each cell has its own vertical extent and time span, written in its corners. (4) The zeroed cell reads Fixed at zero. (5) The inspector reports the zero verified exactly over 1000 samples, and the button below it shows the other four candidates out of date until refitted.

12.7 Technology in Practice

The workbench has no instruments of its own. It reads a record, and the devices that made the record decide what any family can recover. Figure 12.6 draws the column record as it would be taken on a plant.

What Measures It

The bundled records carry their sensors in the generator: seeded noise on every record, and a 12 s lag and 0.32 °C steps on the TCLab temperatures (Table 12.3). On a real column the two compositions would come from an online analyzer that separates a sample over time [23], so each reading arrives late and only once per analysis cycle, a delay that the fit cannot tell from the process, as the TCLab sensor lag shows in Section 12.6.

> **TECHNOLOGY** **Sensor: An Online Composition Analyzer**
>
> Product purity at the top and bottom of a column is often measured by a process gas chromatograph. A small sample is injected at the head of a long, narrow column, and a carrier gas washes its compounds along so that they separate over time and reach the detector one by one. One analyzer reports several species, which suits the products of a distillation column. The separation takes time: cycles in the range of minutes are not uncommon, so a certain measurement dead time is inevitable. The sample line needs a continuous flow to waste, and without it the readings become grossly delayed [23]. An identified model cannot tell that delay from the process, so it would appear in every fitted dead time into x_D and x_B.

What Moves It

During identification an MV is a recorded trajectory, held between samples as an actuator is. In a plant it is often the setpoint of a flow, pressure, temperature or level loop of the regulatory layer, so

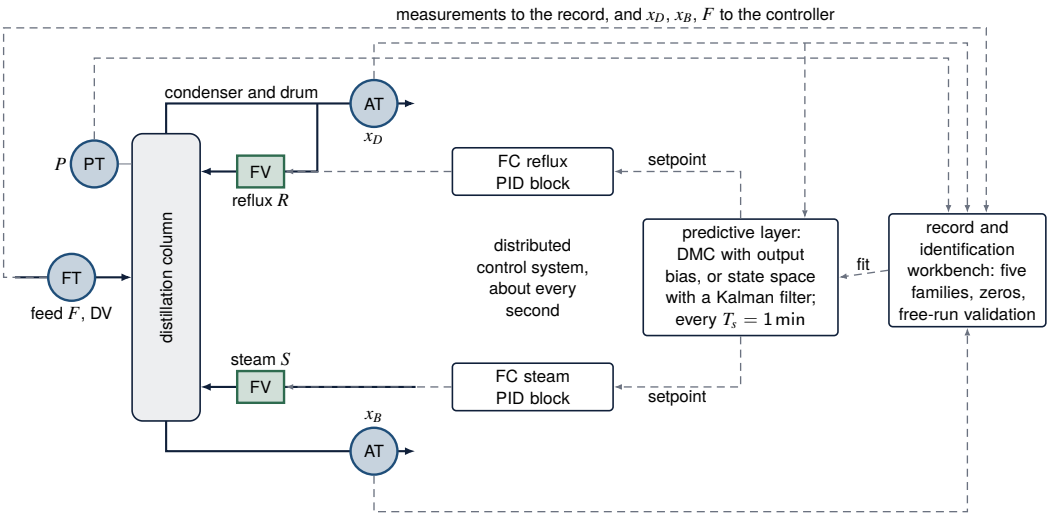


Figure 12.6 The column record of the chapter as it would be taken on a plant. An analyzer (AT) reads each product composition and a pressure transmitter (PT) the column top, a flow meter (FT) measures the feed, and the reflux and steam valves (FV) are positioned by flow loops of the distributed control system. The predictive layer above them writes the two flow setpoints once a minute from a model that the workbench identified on the record. Notice that the identified paths start at the setpoints, so they include the flow loops, and that every composition reading carries the analyzer's own delay.

the identified path runs through that loop and its valve [2]. Industrial packages excite the plant with step-like binary signals or steps of random size, because the test must reveal both the steady state and the dynamics [2].

> **IN PRACTICE** **Cut, Merge and Constrain a Plant Test**
>
> Most industrial identification packages step one, or at most a few, MVs at a time and hold the others at steady state, which is valid as long as the process is linear and superposition holds. A few move several MVs together with uncorrelated signals. Records from several test periods may be merged, which lets the engineer cut out a stretch corrupted by disturbances. The fit mostly minimizes a least-squares criterion, in an equation-error form that feeds past measured outputs back into the model or in an output-error form, and empirical models are often adjusted to known physics, for example a gain known to have a certain sign or value [2].

Where the Controller Runs

The workbench fits in a browser module worker, from 4 ms to 771 ms per family on the bundled records. The controller built from a fit then decides once per sample, every 5 s on the TCLab fit and every minute on the column, in 18 µs for a DMC decision and 2.0 ms or less for a state-space one. Plant control loops run with cycle times from milliseconds to minutes [24], and the predictive layer of a plant, which holds the structure of Section 12.5, sits above the PID function blocks of the distributed control system [2]. The decentralized PI of the comparison corresponds to those function blocks with no predictive layer above them.

TECHNOLOGY Final Element: The Setpoint of a Regulatory Loop

In the control hierarchy of a processing plant, a predictive controller sits above the PID loops of the distributed control system, and those loops position the valves. A survey of industrial practice draws the predictive layer acting about every minute and the PID layer about every second. The moves of the predictive controller are therefore setpoints of lower loops, and its identified paths include them. If the tuning of a lower PID loop changes significantly, a new process model may be needed, built by combining the identified model with a model from the new setpoint to the input [2]. The reflux and steam of the column record are flows in lb/min, the kind of variable such a loop holds.

What Is Modernizing It

Two developments reach this workbench. The first is regularization. Tikhonov proposed it for inverse problems in the 1940s, it became known as ridge regression in linear regression, and around 2000 kernel methods turned it into a kernel approach to identification whose core idea is to accept some bias for a smaller variance and a smaller mean square error. The penalty can also encode structural properties such as stability [25]. The FIR penalty of Equation (12.12), a ridge term plus a smoothness term with its weight chosen on validation, is a simple member of that family. The second is deep learning. Feedforward, convolutional and recurrent networks widen the range of models to choose from, a survey connects them to the kernel methods, and models with very many parameters show counterintuitive behavior such as benign overfitting and double descent [26].

12.8 Limitations

The workbench identifies linear, time-invariant models only. Hammerstein, Wiener, nonlinear ARX and grey-box structures are absent, so a plant whose gain changes with operating point is represented by an average gain and its validity range is the training range. No noise model is estimated. ARMAX and Box-Jenkins structures, instrumental variables and feedback-aware subspace methods are not in this release, so the equation-error ARX start is biased under colored noise, its simulation-error refinement removes that bias only in open loop, and every estimate can be biased by feedback without external excitation [12]. The transfer-function families contain no integrator and no process zero, so integrating paths are recognized and reported by slope rather than modeled in that family, and an inverse response is left to ARX, FIR and state space.

The structural block realization enforces a zero exactly but cannot share dynamics across blocks, so a zero can raise the state-space parameter count, from 14 to 33 on the TCLab record, and a correct zero can cost validation accuracy, 0.0016 on the column. The bootstrap is conditional on the chosen structure, with orders, delays, penalties and mask held fixed, so it understates uncertainty about the structure itself. The step-response dead time is read differently by family, as the fitted θ for transfer functions, the integer d_{ij} for ARX, the first sample above 5 % of the final value for FIR and above 1 % of the peak for state space, so dead times are comparable across families only approximately. A revealed final test is recorded with its revision but cannot be unrevealed. The measured start makes an integrating ARX or state-space model depend on its first measurements, one for ARX and $2n$ for a state block, so a noisy start shifts the whole free run. The 50 000-sample benchmark took 2.5 to 15.0 s per family, which is interactive only in the sense that the page remains responsive while the worker computes, and the state-space refinement is still beyond its budget at that size.

The predictive controller's model is the fit, which is also the nominal plant, so only the scaled-gain and other-fit cases measure model error. There are no rate limits, CV bands or user-set MV limits, and one tuning rule serves every fit. The GEKKO part of `mpc.py` was checked for syntax only, because the development host had neither numpy nor GEKKO, so its closed loop remains unverified.

The records are stylized, and each is defended on its own terms. A pilot column would not be accompanied by a pressure transmitter with exactly no reflux dependence, but the absent path serves to illustrate what a correct structural zero costs and buys. A surge tank whose level is never controlled would overflow in practice, but its record serves to illustrate why an integrator has a slope and not a

gain. The TCLab record is simulated rather than recorded, so that the weak path is known to be real, and the same procedure applies unchanged to a record taken on a physical board (Chapter 4).

12.9 Discussion

A structural zero is the place where process knowledge enters a black-box identification, and the measurements show both sides of that bargain. On the column, a correct zero improved validation NRMSE by at most 0.0013 in four families and cost the state-space model 0.0016. Those changes are small because the absent path was estimated near zero anyway, and the more important effect is that the controller built on the model can no longer act through a gain of 0.014 bar/(lb/min) at a dead time of 270 min that exists only in the noise. A false zero on a strong path raised validation NRMSE by a factor of 3.6 and silently moved the reflux gain from 12.69 to 19.47, which is a model that would mislead a predictive controller into using reflux to do steam's work. The asymmetry is the practical lesson. Removing an absent path costs little and protects the controller, while removing a real one is expensive and visible, but only if the model is judged on data it did not see.

The same discipline bears on model maintenance. A predictive controller's model is identified once and then trusted for years while the unit changes, and the identification step was already the costliest part of a commissioning when surveyed [2]. A record that keeps its roles, recipe, splits, mask with reasons, revision history and an export that checks itself can be re-identified and compared when the plant changes, which is the practical value of making every one of those choices explicit and reproducible.

> ⚠ **OPEN CHALLENGE** **When a Linear Model Is Not Enough**
>
> A road map of nonlinear identification notes that the real world is nonlinear and time-varying, and names distillation columns among the systems with nonlinear dynamic behavior. Compared with linear identification, the experiment design is more tedious, model selection much more involved and estimation more difficult, and very many nonlinear structures have been proposed, so the step raises the cost of the whole effort. Where a periodic test signal can be applied, a nonparametric analysis can detect the nonlinear distortions and estimate how much a nonlinear model could gain [27]. Progress would be a workbench that measures this gap on a plant record before a nonlinear family is chosen.

12.10 Conclusion

This chapter studied a browser workbench for multivariable identification, with first- and second-order transfer-function matrices, ARX, FIR and subspace state-space models fitted to the same prepared record and ranked on free-run validation error. The transfer functions are sampled exactly with fractional dead time and fitted by variable projection with a bounded multistart Levenberg-Marquardt search, and on the column record they recover all six Wood-Berry gains within 0.9 %. A path fixed at zero is removed from every family and verified to 0.0, with state-space zeros enforced by structural blocks whose Markov parameters vanish identically. A correct zero on the column changed the first-order validation error from 0.1093 to 0.1085, a false zero raised it to 0.3892, and a weak but real path on the TCLab record raised it by a factor of 3.7 when removed. An active-set search with a damping coordinate stops eight of the nine second-order CV fits on the bundled records before their iteration limit, and a measured start with a state drift lets ARX and state space follow an integrating level at 0.190 and 0.062, where the first release's models drifted. Oscillatory and closed-loop records show where each family fails, and every exported model reproduces the application to 4.3×10^{-13} in numpy and 4.9×10^{-10} in GEKKO. As a predictive controller the first-order fit cut the interaction to 0.137 from 0.344 °C on the TCLab record and to 0.0739 from 0.475 mol% on the column, while the dynamic-matrix output bias lost a heater load to the PI, 372 against 217 °C s. The models are linear and noise-free, the state-space zero cannot share dynamics across blocks, and the bootstrap is conditional on the chosen structure. A fuller workbench would add ARMAX and Box-Jenkins noise models, instrumental-variable and closed-loop subspace methods, integrating and inverse-response

transfer functions, constrained prediction-error refinement on a coupled state, and recorded rather than generated example data. The workbench is an accessible platform for identifying multivariable models from plant records and for testing what a structural assumption is worth.

KEY POINTS

- A multivariable model is not a collection of single-input models: the inputs into one output are fitted together.
- A fitted gain is a property of the model structure as well as of the plant: the reflux gain climbs from 12.69 to 19.47 when the steam path is falsely removed.
- Removing an absent path costs little and protects the controller, while removing a real one is expensive and visible, but only on data the fit did not see.
- The dynamic-matrix controller, whose output bias assumes that the present discrepancy persists, lost an unmeasured heater load to the PI, 372 against 217 °C s, while the Kalman form reached 132.

Where else this applies. A predictive controller built on an identified response matrix also runs above a decentralized structure in Volume 6, Chapter 8, and Volume 9, Chapter 4 removes an unmeasured fuel bias with a Kalman disturbance estimate, the remedy for the output bias measured here.

References

[1] J. Richalet, A. Rault, J. L. Testud, and J. Papon, "Model predictive heuristic control," *Automatica*, vol. 14, no. 5, pp. 413–428, 1978.

[2] 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.

[3] Y. Zhu, "Multivariable process identification for MPC: the asymptotic method and its applications," *Journal of Process Control*, vol. 8, no. 2, pp. 101–115, 1998.

[4] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, 2010.

[5] K.-J. Åström and T. Bohlin, "Numerical identification of linear dynamic systems from normal operating records," in *Theory of Self-Adaptive Control Systems*. Boston, MA: Springer US, 1966, pp. 96–111.

[6] K. J. Åström and P. Eykhoff, "System identification—a survey," *Automatica*, vol. 7, no. 2, pp. 123–162, 1971.

[7] M. R. Gevers, "A personal view of the development of system identification: A 30-year journey through an exciting field," *IEEE Control Systems*, vol. 26, no. 6, pp. 93–105, 2006.

[8] B. L. Ho and R. E. Kalman, "Effective construction of linear state-variable models from input/output functions," *Regelungstechnik (now at – Automatisierungstechnik)*, vol. 14, pp. 545–548, 1966.

[9] M. Verhaegen and P. Dewilde, "Subspace model identification Part 1. The output-error state-space model identification class of algorithms," *International Journal of Control*, vol. 56, no. 5, pp. 1187–1210, 1992.

[10] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61–74, 1994.

[11] P. Van Overschee and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.

[12] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, no. 7, pp. 1215–1241, 1999.

[13] R. K. Wood and M. W. Berry, "Terminal composition control of a binary distillation column," *Chemical Engineering Science*, vol. 28, no. 9, pp. 1707–1717, 1973.

[14] J. Park, R. A. Martin, J. D. Kelly, and J. D. Hedengren, "Benchmark temperature microcontroller for process dynamics and control," *Computers & Chemical Engineering*, vol. 135, p. 106736, 2020.

[15] 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.

[16] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Problems*, vol. 19, no. 2, pp. R1–R26, 2003.

[17] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[18] H. R. Künsch, "The jackknife and the bootstrap for general stationary observations," *The Annals of Statistics*, vol. 17, no. 3, pp. 1217–1241, 1989.

[19] C. E. Garcia and A. M. Morshedi, "Quadratic programming solution of dynamic matrix control (QDMC)," *Chemical Engineering Communications*, vol. 46, no. 1–3, pp. 73–87, 1986.

[20] 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.

[21] E. Bristol, "On a new measure of interaction for multivariable process control," *IEEE Transactions on Automatic Control*, vol. 11, no. 1, pp. 133–134, 1966.

[22] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *Journal of Process Control*, vol. 13, no. 4, pp. 291–309, 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] 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] G. Pillonetto, T. Chen, A. Chiuso, G. De Nicolao, and L. Ljung, *Regularized System Identification: Learning Dynamic Models from Data*, ser. Communications and Control Engineering. Cham: Springer, 2022.

[26] G. Pillonetto, A. Aravkin, D. Gedon, L. Ljung, A. H. Ribeiro, and T. B. Schön, "Deep networks for system identification: A survey," arXiv preprint 2301.12832, 2023, published in revised form in *Automatica*, vol. 171, art. 111907, 2025, doi:10.1016/j.automatica.2024.111907.

[27] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented roadmap," arXiv preprint 1902.00683, 2019, published in revised form as "Nonlinear system identification: A user-oriented road map," *IEEE Control Systems*, vol. 39, no. 6, pp. 28–99, 2019, doi:10.1109/MCS.2019.2938121.

12.11 Exercises

12.1 ○ Concept check · 2 min

On the TCLab record the one-step persistence baseline, which predicts each sample as the last measured one, scores a validation NRMSE of 0.0260. The best free-run candidate scores 0.0232. Why does the workbench rank candidates on free-run error and never on one-step error?

(a) A one-step prediction carries the last measurement forward, so it scores well with no model, while a controller relies on the free run from the inputs

(b) A one-step prediction draws on final-test samples, which must stay hidden until a model has been chosen and the Test step reveals them

(c) A one-step error is always larger than a free-run error on the same split, so ranking on it would reject every candidate on every record

(d) A one-step prediction cannot be computed for a record with two inputs and two outputs, so it does not cover every record

12.2 ○ Concept check · 2 min

The reflux path into column pressure is fixed at zero and the state-space family is refitted. How does the workbench make that path exactly zero?

(a) It sets the reflux column of B to zero and keeps the single coupled block of six states

(b) It gives pressure its own block of states driven only by steam and feed, so the Markov parameters of the path vanish identically

(c) It adds a large penalty on the reflux gain into pressure, which drives that gain below 10^{-12} in the refit

(d) It sets the pressure row of C to zero, so that no state of the model reaches the pressure output

12.3 Concept check · 1 min

When the strong steam path into distillate composition is falsely fixed at zero, the refitted reflux gain into x_D rises from 12.69 to 19.47 mol%/(lb/min). What does the rise show?

(a) The reflux valve has become more effective since the record was taken

(b) The search stopped early and would return 12.69 given more iterations

(c) The remaining paths absorb what they can of the missing steam response

(d) The zero removed noise that had hidden the true reflux gain all along

12.4 Theory and calculation · 7 min

The first-order path from heater 1 to sensor 1 of the TCLab record has $K = 0.512\,^{\circ}\mathrm{C}/\%$, $\tau = 104\,\mathrm{s}$ and $\theta = 10.9\,\mathrm{s}$, sampled at $T_s = 5\,\mathrm{s}$.

(a) Split the dead time as $\theta = dT_s + \phi$. *Checkpoint:* $d = 2$, $\phi = 0.9\,\mathrm{s}$.

(b) With $A = -1/\tau$ and $B = 1/\tau$, evaluate Φ, Γ_0 and Γ_1 of Equation (12.6), and show that $\Gamma_0 + \Gamma_1 = 1 - \Phi$. *Checkpoint:* $\Phi = 0.95306$, $\Gamma_0 = 0.038656$, $\Gamma_1 = 0.0082834$, sum 0.046940.

(c) Heater 1 steps by 10 % at $k = 0$. Find the first sample at which T_1 moves, the deviation there and one sample later, and compare both with the continuous response $K\Delta u(1 - \mathrm{e}^{-(t-\theta)/\tau})$. What would rounding the dead time to 10 s give at $k = 3$? *Checkpoint:* $k = 3$: 0.198 °C; $k = 4$: 0.429 °C, equal to the continuous values at 15 and 20 s. Rounded: 0.240 °C.

12.5 Theory and calculation · 6 min

(a) Evaluate the damping factor of Equation (12.9) at the lower bound $q = \ln 0.1$, at $q = 0$ and at the upper bound $q = 4$. At the upper bound, find the ratio of the two real lags $\tau(\zeta \pm \sqrt{\zeta^2 - 1})$ and compare it with e^{2q}. *Checkpoint:* $\zeta = 0.1$, 1 and 28.3. Ratio 3199, against $\mathrm{e}^8 = 2981$.

(b) The stage generator rings at a natural frequency of 1.6 Hz with $\zeta = 0.15$, and the second-order fit returns $\tau = 0.100\,\mathrm{s}$ and $\zeta = 0.196$. Find the generator's τ, and from Equation (1.12) the overshoot and period of the generator and of the fit. *Checkpoint:* $\tau = 0.0995\,\mathrm{s}$. Overshoot 62.1 % against 53.4 %, period 0.632 against 0.641 s.

12.6 Theory and calculation · 6 min

Use the fitted gains, lags and dead times of Table 12.5 with $T_s = 1\,\mathrm{min}$.

(a) Compute the relative gain Λ_{11} of the reflux and steam pair on x_D and x_B from the fitted gains, and again from the generator's. *Checkpoint:* 2.019 from the fit, 2.009 from the generator.

(b) Tune the $x_D \leftarrow R$ loop by the SIMC rule of Section 12.5. *Checkpoint:* $\theta_e = 1.5\,\mathrm{min}$, $K_c = 0.441$ (lb/min)/mol%, $\tau_I = 12.0\,\mathrm{min}$.

(c) Tune the $x_B \leftarrow S$ loop the same way, and say what the sign of its gain means. *Checkpoint:* $\theta_e = 3.56\,\mathrm{min}$, $K_c = -0.103$ (lb/min)/mol%, $\tau_I = 14.3\,\mathrm{min}$.

12.7 App challenge · 8 min

Use scenarios 1, Read a two-heater matrix, and 2, A weak path that is real. The page opens on the TCLab two-heater test with the default splits. On 3 · Identify select First order + dead time and press Fit selected methods. The bundled records carry fixed noise, so the fit repeats, and every value below holds within one unit in its last digit.

(a) Open the cells $T_1 \leftarrow Q_1$ and $T_1 \leftarrow Q_2$ and read K, τ and θ of each. *Checkpoint:* 0.512, 104 s, 10.9 s; and 0.0789, 181 s, 52.6 s.

(b) On 4 · Validate read the candidate's validation NRMSE, and its T_1 and T_2 rows in the per-CV table. *Checkpoint:* 0.0232; T_1 0.0299, T_2 0.0165.

(c) Back on Identify, open $T_1 \leftarrow Q_2$, press Set response to zero and let the model refit. Read the validation NRMSE and its rows again, and K and θ of $T_1 \leftarrow Q_1$. *Checkpoint:* 0.0853; T_1 0.154, T_2 0.0165; $K = 0.542$, $\theta = 4.79$ s.

Explain why the direct path's dead time shrinks when the cross path goes.

12.8 App challenge · 8 min

Use scenario 3, A path that is not there. On 1 · Import load the Distillation column (2 MVs, feed DV) example, and on 3 · Identify fit First order + dead time with the defaults. Values hold within one unit in the last digit shown.

(a) Open Pressure ← Reflux and read K, τ and θ. From the reflux range of Table 12.1, find the largest steady pressure change this path predicts. *Checkpoint:* $K = 0.0141$ bar/(lb/min), $\tau = 114.6$ min, $\theta = 270$ min on a bound; 0.0056 bar.

(b) Press Set response to zero on that path and let the model refit. Read the zero check in the details, and on 4 · Validate the parameter count and validation NRMSE. *Checkpoint:* Verified exactly zero over 1000 samples; 27 parameters, 0.1085 (from 30 and 0.1093).

(c) Press Restore response, then set Distillate xD ← Steam to zero instead. Read the validation NRMSE, its x_D row, and K and τ of Distillate xD ← Reflux. *Checkpoint:* 0.389; x_D 0.939; $K = 19.47$, $\tau = 42.7$ min.

12.9 App challenge · 9 min

Use scenario 13, MPC from the fit, on the TCLab two-heater test with its First order + dead time fit. Open 7 · MPC at the defaults (horizon 1, $\lambda = 1$, Kalman filter) and wait for the table, about 8 s. Values hold within one unit in the last digit shown.

(a) In the rows of Setpoint step on T1 temperature, read the IAE of T_1 and the Interaction for PI, DMC and State space. *Checkpoint:* PI 347 and 0.344 °C; DMC and State space 392 and 0.137 °C.

(b) In the rows of Unmeasured load on Q1 heater, read the IAE of T_1 for the three controllers. *Checkpoint:* PI 217, DMC 372, State space 132 °C s.

(c) Set the move suppression to $\lambda = 0.1$, wait for the table, and read the first row pair of (a) again. Repeat at $\lambda = 10$. *Checkpoint:* $\lambda = 0.1$: MPC 342 and 0.0610. $\lambda = 10$: MPC 598 and 0.133. PI unchanged.

Explain why the PI wins on the stepped CV at $\lambda = 1$ and why the DMC loses the load.

12.10 Application area · 3 min

The feed path into column pressure has a true gain of 0.035 bar/(lb/min), a lag of 9 min and a dead time of 2 min. The record moves feed over 2.2 to 2.7 lb/min with 0.004 bar of pressure noise, and the first-order fit returns 0.0349, 5.69 min and 3.51 min.

(a) Find the largest steady pressure change that feed causes in the record, as a multiple of the noise. *Checkpoint:* 0.0175 bar, about 4.4 times the noise.

(b) Name two things the chapter offers an engineer who needs this path's dynamics better than the record gives them.

12.11 Application area · 4 min

A plant engineer has a month of historian data from the exchanger's outlet temperature loop, always on automatic at one setpoint, with the steam valve moving only to reject an unmeasured drift. The engineer wants the valve gain for a predictive controller. Using the chapter's assumptions and its exchanger record, say what can go wrong and what to do instead.

12.12 Review the agent · 5 min

An AI assistant is asked to simplify the TCLab models. It writes: "The cross path $T_1 \leftarrow Q_2$ is weak. Fixing it at zero removes its three parameters from the first-order matrix, from 14 to 11. In the 2-state state-space model it removes the path's entry of B and its effect through C, so that model falls from 14 parameters to 12 or fewer." Find the flaw and say how you would check it.