INTERACTIVE CONTROL ENGINEERING
Principles and Applications

Volume 1 Process Dynamics, Modeling, and Identification

CHAPTER 13

# Process Monitoring Workbench That Diagnoses Records From Declared Simulated Plants

Pages 315–337

Edited by John D. Hedengren

**How to cite this chapter.** "Process Monitoring Workbench That Diagnoses Records From Declared Simulated Plants," 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. 13, pp. 315–337.

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

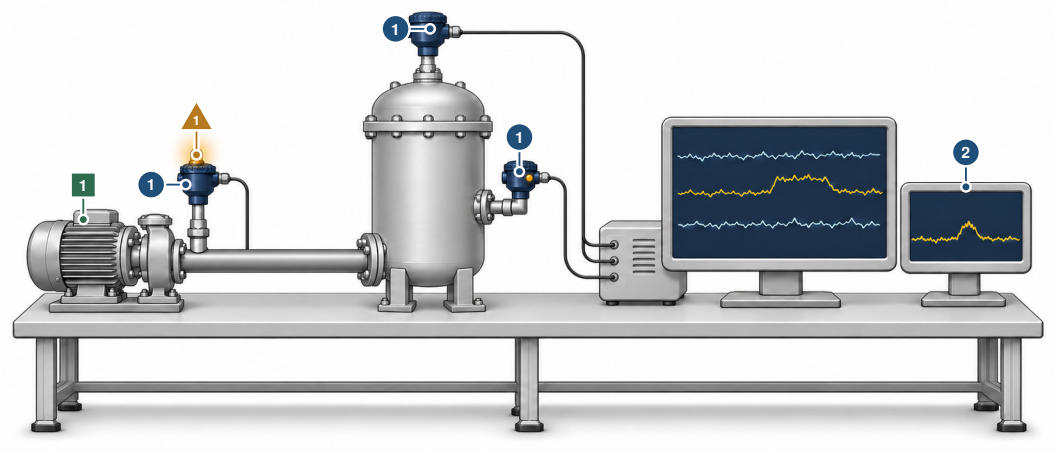
**The app for this chapter**
https://apmonitor.com/apps/monitor/

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

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

CHAPTER 13

# Process Monitoring Workbench That Diagnoses Records From Declared Simulated Plants




**The controller moves** 1 pumps whose schedule sets the flow through the network

**The controller reads** 1 flow, pressure and level transmitters whose readings the record keeps

2 a residual statistic on its own screen

**What disturbs it** 1 one meter that reads a few percent high

*Notice that the meter with the amber light lifts one trend and the residual, yet neither screen says which meter is wrong; the chapter's records come from simulators, not from a bench like this one. Illustration, not to scale.*

### IN PLAIN TERMS Can the Plant's Own Record Be Trusted?

**The main idea.** A plant stores every reading it takes, often for years. This chapter checks such stored records for a meter that reads wrong, a fault that is starting, a valve that sticks and a burst of warnings.

**The challenge.** A warning does not name its cause. One bad meter can make every number that is added up with it look wrong. Readings taken one after another are also linked to each other, so false warnings come far more often than the usual rule of thumb promises.

**Why it matters.** An operator who trusts the wrong warning, or who gets too many, can miss the one that matters.

### THREE WAYS TO READ THIS CHAPTER

| | | |
|---|---|---|
| **Explore** | 12 min | the first page, the fact on reviewing alarms (Section 13.1), the histories of three-sigma limits and of an alarm flood (Section 13.2), the try-it on the water network, the fact on false alarms and the try-it on the alarm flood (Section 13.5) and the key points |
| **Study** | 45 min | the whole chapter with its three pause points and the twelve exercises at its end |
| **Practice** | 15 min | the first page, the results (Section 13.5), Technology in Practice (Section 13.8), Limitations (Section 13.9) and the open challenge |

**The control question** A plant record holds the evidence of a biased flowmeter, a developing fault, a sticking valve and an alarm flood. What does each monitoring statistic actually say about such a record, when one gross error spreads into every balance it enters and the samples of a closed-loop plant are not independent?

**Learning objectives**

- Reconcile a linear flow network, apply the global and measurement tests, and explain why one biased meter fails every test.
- Train a principal-component model and tell what the $T^2$ and $Q$ statistics each detect.
- Update EWMA and CUSUM charts by hand and convert a run length to time.
- Estimate the false detections per day of a run-of-three rule, and explain why a closed-loop record raises far more.
- Tell a sticking valve from an aggressive loop, and rationalize an alarm flood with deadbands and delays.

**Prerequisites** Sections 1.4.8 (least squares, ARX) and 1.4.9 (residuals and their autocorrelation). Section 1.6.3 is needed only for the MPC health tool. The chapter is best read after Chapter 12.

**The app** https://apmonitor.com/apps/monitor/

**See also** Volume 5, Chapter 4, whose drifting kappa analyzer is the digester record here, and Volume 5, Chapter 3 for control through a slow analyzer. Volume 2, Chapter 4, whose sticking steam valve produces loop A, and Volume 2, Chapter 6 for a gain-driven limit cycle that a sticking valve would mimic.

## 13.1 Introduction

Every measurement a plant makes is stored, usually for years, and the questions asked of that store have a long history of their own. Flow and inventory data are adjusted to close their balances before they are used for accounting [1, 2], many correlated signals are charted together against a model of normal operation [3], a loop's routine record is compared with the best variance any controller could reach [4], and the alarm log is judged against the rates an operator can handle [5]. None of these tasks needs a plant model. Each needs a record, a declared structure and a statistic with a limit.

**DID YOU KNOW? Fifty Alarms a Shift, or One Shift an Alarm**

Fixing an alarm system takes time. Guidance on alarm handling from the Health and Safety Executive estimates that a quick first pass may cover perhaps 50 alarms per shift, while a thorough review and redesign may take more than one shift per alarm [6].

The misconception this workbench is built to correct is that a statistic above its limit names a cause, and that a 99 % limit fires on one sample in a hundred. On the shipped water-network record a single 3 % calibration shift on one flowmeter puts all seven measurement tests above their limit, from 11.68 to 17.71 against 2.68, because a gross error spreads into every stream that shares a balance with it. On normal days of the port of the Tennessee Eastman plant [7] studied in Volume 6, Chapter 8, the residual statistic sits above its 99 % limit on 4.3 % of samples, because samples of a closed-loop plant are not independent. The same two facts govern any unit with redundant meters and feedback loops, from a refinery's material balance to a utility's distribution network.

However, a student usually meets these methods one at a time, on small static examples, and almost never on a record whose cause is known and declared. Benchmark plant data exist, but the cause of an anomaly in a real historian is rarely known with certainty, and a synthetic example built for one method teaches nothing about what the others would have said. This chapter works with a workbench in which five tools act on records generated by other simulators of the collection that accompanies this series, with every injected change written in the file. It reports what each tool measures on each record. The workbench simulates no plant itself: it reads records, as the identification workbench of Chapter 12 does, and asks whether the evidence in them can be trusted.

The estimators and tests are the standard ones. The results read off the shipped records include a negative result (one Tennessee Eastman fault is not detectable at the statistics' false-alarm rate, over 61 seeds) and two discrepancies in published sources, each checked by independent routes. Section 13.2 follows the methods' history. Section 13.3 sets out the records and the scripts that generate them, which run other simulators' engines headlessly and declare every change. Section 13.4 states the five tools as coded, in enough detail to rebuild them. Section 13.5 measures the ten cases, each of which has a completion rule in the app, and Section 13.6 adds a sixth tool that reads an exported predictive-control run. Section 13.7 reports the verification, Section 13.8 sets the records beside the devices and systems of a plant, and Sections 13.9 and 13.10 state the limitations and the conclusions.

## 13.2 Balances, Charts and Principal Components on Plant Records

Reconciliation of process data was set out for plant networks in 1976, when flow and inventory measurements of a plant were adjusted by weighted least squares so that the material balances closed, each adjustment scaled by its meter's variance [1]. A test on each adjustment followed, so that a single gross error could be detected and attributed to a measurement [8], and matrix projection removed unmeasured flows from the balances so that the measured ones could be reconciled alone and the unmeasured ones recovered afterward where they were observable [9]. By the end of the century reconciliation and gross-error detection were collected in a textbook that also covered dynamic systems and industrial applications [2]. The link between reconciliation and principal-component analysis was later made explicit, with a six-stream recycle network whose constraints principal components can recover [10].

The univariate charts are older. The cumulative sum was proposed in 1954 as a continuous inspection scheme [11], the geometric (now exponentially weighted) moving average in 1959 [12], and the run length of a CUSUM was computed by a Markov chain on the discretized statistic in 1972 [13]. The EWMA's properties and enhancements were worked out in detail later [14], and the NIST/SEMATECH e-Handbook prints the CUSUM run lengths that practitioners consult [15]. Multivariate monitoring added a limit for the residual of a principal-component model [16], was carried into chemical processes with charts of scores and residuals from correlated plant data [3], and gained lagged variables for autocorrelated records [17]. The Tennessee Eastman problem [7], published as a plantwide control challenge with a list of disturbances, became a common test bed for such monitors.

> ⧖ HISTORY **Three-Sigma Limits From the 1920s**
>
> During the 1920s Walter A. Shewhart proposed a general model for control charts. A sample statistic is plotted in time order with a center line at its mean and an upper and a lower control limit a chosen number of standard deviations away. Three became the accepted standard in industry. For a normal variable the two three-sigma limits are crossed by chance with a probability of 0.0027, and the handbook that sets this out calls such a risk purely arbitrary. Points inside the limits that form a nonrandom pattern also mean that something is wrong [15]. Shewhart rule 1 of the workbench is this limit, and the other Western Electric rules are tests for such patterns.

Loop assessment began with the observation that routine closed-loop data already contain the minimum variance any controller could reach, given the process delay [4]. Oscillation detection followed from the autocorrelation of the control error and the regularity of its zero crossings [18], and valve stiction was characterized by its stick band and slip jump in the plot of valve position against controller output [19]. Alarm management, finally, became a lifecycle with its own standard, ANSI/ISA-18.2-2016, issued by the International Society of Automation [5]. The workbench names that standard as the class of source of its alarm metrics, and every threshold it applies is a teaching value, not a number read from the standard.

**Table 13.1** Shipped Records. Every Record's Header Names Its Generating Application, Scenario, Seed and Script and Declares Every Change Listed Here.

| Record | Generator, case | Rows, period | Declared change |
|---|---|---|---|
| Tennessee Eastman, six records | tennessee-eastman, case 4 settings, seed 20261014 | 1441, 3 min | case event removed, one IDV switched on at 48 h (none in the normal record) |
| Water network | water-network-pumping, case 7, seed 20261030 | 119, 15 min means | station meter $\times 1.03$ from 4 h, three derived streams |
| Digester | kraft-digester, case 5, seed 20261030 | 1441, 1 min | none |
| Loop A | heat-exchanger, case 15, seed 20260825 | 1201, 0.1 min | none (the app's own 3 % stiction) |
| Loop B | continuous-caster, case 5, seed 20260923 | 6001, 0.1 s | none (gain 5, clean gate) |
| Alarm log | turbine-trip case 4 and smr-steam-header case 10, seed 20260906 | 3601, 1 s | trips moved to 1800 and 1500 s, colored noise $\tau$ 2 s, $\sigma$ 0.2 % of span |
| Cabin nitrogen | space-station-eclss, case 7, seed 20261030 | 577, 5 min | none (extra record, unscored) |

> ⧖ HISTORY **Too Many Alarms at a Refinery, 1994**
>
> In 1994 an explosion and fires at a refinery in Wales injured twenty-six people and caused damage of around £48 million. The investigation by the Health and Safety Executive found three key factors. There were too many alarms and they were poorly prioritized, the control room displays did not help the operators understand what was happening, and training for a stressful and sustained upset had been inadequate. In the last 11 minutes before the explosion the two operators had to recognize, acknowledge and act on 275 alarms. The regulator then issued guidance on better alarm handling to prevent similar accidents, and it recommends an industry guide to alarm systems, published in 1999, to which it contributed [6]. The alarm log of case 8 asks the same question in one simulated hour.

## 13.3 The Records and Their Declarations

The workbench reads a comma-separated file with a header row, a time column and any number of signals. Rows need not be evenly spaced, blank, `NaN` and text cells are missing values, and date-times become hours from the first row. A line `# monitor:` carries a JSON declaration of balances, alarm limits or loop roles, and other lines starting with `#` are the header. The shipped records, listed in Table 13.1, are generated by scripts that import another simulator of the collection read-only, run its engine at its own step, keep the measured columns a historian would hold and write the header with the generating application, scenario, seed, sampling, every change the script made and the script's name. No record carries a truth column. When the records were built, regenerating all thirteen files (twelve records and the answer file the tests use) reproduced every byte. Section 13.6 notes the one record that no longer does.

### 13.3.1 Six Tennessee Eastman Records

The Tennessee Eastman records come from the port of the plant in Volume 6, Chapter 8, under its decentralized structure with the pressure override, run for 72 h at the engine's 1 s step and sampled every 3 min from its 1 min historian. Each row holds the 41 measurements (XMEAS 1 to 41, noise and analyzer holds included) and the 11 valve commands (XMV 1 to 11), the 52 signals of the monitoring literature. The script removes the case's own disturbance and switches one IDV on at 48 h, so the first two days of every record are identical normal operation and serve as the training window. The faults are IDV(1), an A/C feed ratio step, IDV(5), a condenser cooling water inlet temperature step, IDV(13), a slow drift of the reaction kinetics, IDV(3), a D feed temperature step,

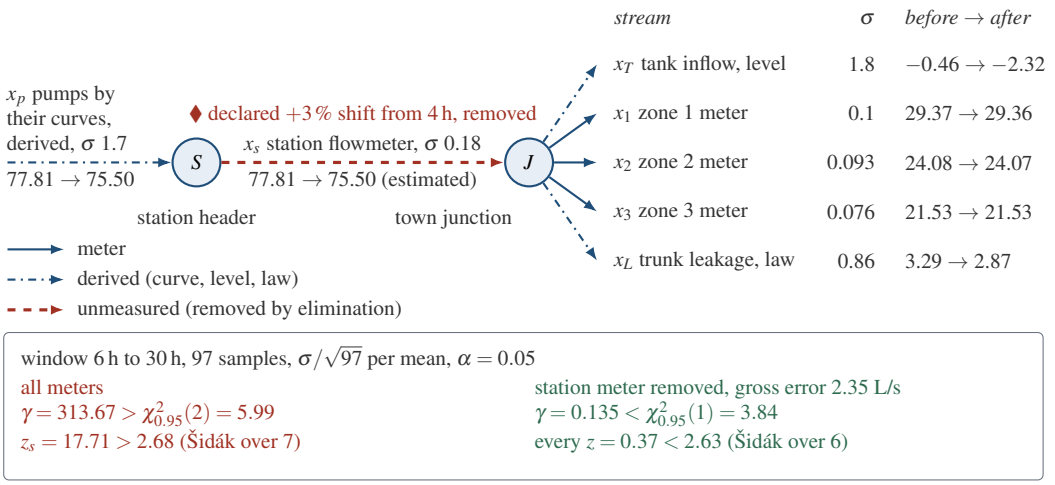



**Figure 13.1** The water-network record as two balances, the station header $S$ and the town junction $J$, with seven streams in L/s. Line style gives each value's source, and $\sigma$ is the declared standard deviation of one 15 min value. Each stream carries its reconciled value before and after serial elimination on the window from 6 to 30 h. The station meter carries the declared 3 % shift and is the stream elimination removes. The clean window from 0 to 4 h gives $\gamma = 0.07$.

and, for the unlabeled challenge, IDV(4), a reactor cooling water inlet temperature step [7]. The challenge record's header says only that its scenario is withheld.

### 13.3.2 A Water Network With a Declared Meter Fault

The water-network record is case 7 of the distribution simulator of Volume 9, Chapter 11: scheduled pumping and a main break in zone 3 at 10 h, run 30 h. Every 10 s step the script reads the simulator's measurement packet and averages it into 15 min blocks stamped at each block's end. Four streams are meters (the station flowmeter and three zone meters). Three more are derived, as a utility engineer would write them down, from measured heads and the level, as shown in Equation (13.1).

$$
\begin{aligned}
x_p &= \sum_{i\,\text{running}} \sqrt{\left(a_i \omega_i^2 - (H_S - z_{cw})\right)/b_i} \\
x_T &= A_T \,\frac{\bar{h}_{k+1} - \bar{h}_{k-1}}{2\Delta t}, \qquad x_L = C_T\, p_J^{N_1}
\end{aligned}
\tag{13.1}
$$

where $x_p$ is the pumps' flow from their declared curves ($a_i$, $b_i$ and speed $\omega_i$) at the measured station head $H_S$ above the clearwell level $z_{cw}$, $x_T$ is the tank inflow from a centered difference of the block-mean levels $\bar{h}$ with tank area $A_T = 600\,\text{m}^2$ and block $\Delta t = 15\,\text{min}$, and $x_L$ is the trunk-main leakage from its declared law at the measured junction pressure head $p_J$ with $N_1 = 1.18$, carried as an allowance. A one-sided difference lagged the block means by half a block and failed the clean 0 h to 4 h window with $\gamma = 9.12$, which is why the difference is centered. Each stream's standard deviation was measured against the engine's truth, then the truth was dropped, with the pump curves floored at 2 % and the leakage law at 30 % of their means as teaching values. The simulator's meters have no bias option and the simulator is not edited, so the script multiplies the station flowmeter by 1.03 from 4 h and says so in the header. The two balances, the seven streams and their declared $\sigma$ are drawn in Figure 13.1.

### 13.3.3 The Other Records

The digester record is case 5 of the kraft digester simulator of Volume 5, Chapter 4, in which the kappa loop runs on an analyzer whose bias drifts from 1 h. The analyzer samples every 15 min and reports 20 min later, the laboratory samples every 4 h and reports 1 h later, and each column holds its last report beside the time it was sampled. The two loop records keep setpoint, measurement, controller output and valve (or gate) position. Loop A is a steam heater's outlet-temperature loop whose valve has the simulator's own 3 % stiction deadband, and loop B is a caster's mold-level loop with a proportional gain of 5 and a clean gate. The alarm log puts two simulators on one clock at 1 s, a steam-generating unit that trips at 1500 s and a turbine at 1800 s, with thirteen alarm limits set close to normal operation. Both simulators' measurements are exact at steady state, which left 26 raw annunciations and one flood that no deadband or delay could reduce, so the script adds declared colored noise with a 2 s time constant and a standard deviation of 0.2 % of each tag's span (seed 20261030). The cabin record, the nitrogen ledger from the sealed-cabin simulator of Volume 6, Chapter 10, is an extra, unscored record for the charts.

### 13.3.4 Truth and the Challenge's Answer

Truth stays in the generating simulators. The answers the completion rules need are in a separate file that the shipped code does not import. The labeled cases' answers are copied into a small key, and the challenge's fault class and three responsible variables are kept only as one-way digests. With seven classes and 52 variables a learner could recover them by hashing every option, so the digest guards against a casual look and is not a security measure. A test re-derives every key entry from the answer file and fails if any bundled source imports it.

## 13.4 Five Tools as Implemented

Each tool reads only the record's measured columns. The settings each exposes and their defaults are listed in Table 13.2. Every statistic is computed in the browser in double precision by code written for the workbench, with no numerical library.

### 13.4.1 Reconciliation and Gross-Error Tests

The balances of a record are linear in its stream flows $x$, with one row of the incidence matrix per node, $+1$ for a stream entering and $-1$ for one leaving. For the water record they are given by Equation (13.2).

$$\begin{aligned} S:&\quad x_p - x_s = 0 \\ J:&\quad x_s - x_T - x_1 - x_2 - x_3 - x_L = 0 \end{aligned} \tag{13.2}$$

where the symbols are those of Figure 13.1. Splitting the incidence matrix into its measured and unmeasured columns, $A = [A_k\ A_u]$, a basis $P$ of the left null space of $A_u$ removes the unmeasured streams, and the reduced balances $B$ are an orthonormal basis of the rows of $PA_k$. With $y$ the measured values (window means) and $\Sigma$ the diagonal matrix of their variances, the reconciled values, the global test and the measurement test on each adjustment $a_i$ are given by Equation (13.3).

$$\begin{aligned} \hat{x}_k &= y - \Sigma B^{\mathsf{T}} \left( B \Sigma B^{\mathsf{T}} \right)^{-1} B y \\ \gamma &= (By)^{\mathsf{T}} \left( B \Sigma B^{\mathsf{T}} \right)^{-1} B y \sim \chi^2(r) \\ z_i &= \frac{|a_i|}{\sqrt{\left( \Sigma B^{\mathsf{T}} (B \Sigma B^{\mathsf{T}})^{-1} B \Sigma \right)_{ii}}} \end{aligned} \tag{13.3}$$

where $r$ is the number of reduced balances and $a = y - \hat{x}_k$. The global test fails when $\gamma$ exceeds the $1-\alpha$ quantile of $\chi^2(r)$. Each $z_i$ is compared with a two-sided normal limit Šidák-corrected over the $m$ redundant measurements, $z_c = \Phi^{-1}\left(1 - [1-(1-\alpha)^{1/m}]/2\right)$, which is 2.68 for $m = 7$ and 2.63 for $m = 6$ (a Bonferroni correction gives 2.69 for $m = 7$). A window mean of $n$ samples is given

**Table 13.2** Tool Settings and Defaults. Visible Settings Are on the Tool's Panel, Declared Ones Come From the Record, Fixed Ones Are in the Code.

| Setting | Default | Status |
|---|---|---|
| *reconciliation* | | |
| balances, streams, kinds | from the record | visible |
| $\sigma$ per sample, window | declared, 6 to 30 h (water) | visible |
| test level $\alpha$ (both tests) | 0.05, range 0.001 to 0.2 | visible |
| *principal components* | | |
| training window | 0 to 48 h | visible |
| components $k$ | 50 % cumulative variance | visible |
| lags $l$ | 0, range 0 to 4 | visible |
| confidence of both limits | 99 %, range 90 to 99.9 % | visible |
| detection run | 3 samples | visible |
| *charts* | | |
| EWMA $\lambda$, $L$ | 0.2, 2.962 | visible |
| CUSUM $k$, $h$ (in $\sigma$) | 0.5, 5 | visible |
| Shewhart rules | rule 1 only | visible |
| chain states, CUSUM and EWMA | 120 and 151 | fixed |
| *loops* | | |
| delay $d$, AR order $p$ | declared (10 samples), 10 | visible |
| window | last two thirds of record | visible |
| moved threshold $\varepsilon$ | 0.02 % of output range | fixed |
| stiction screen | held $>$ 25 %, band $>$ 0.5 % | fixed |
| *alarms* | | |
| deadband, on- and off-delay, filter | 0, 0, 0, 0 | visible |
| flood | $>$ 10 in a 10 min period | teaching |
| chattering | $\geq$ 3 within 60 s | teaching |
| rate target | 1 per 10 min, 144 per day | teaching |

$\sigma/\sqrt{n}$. An unmeasured stream is recovered from $A_u x_u = -A_k \hat{x}_k$ by the pseudo-inverse and reported as observable only when no direction of the null space of $A_u$ moves it. A measured stream absent from every reduced balance is non-redundant and keeps its measurement. An allowance, such as the leakage law, is carried as a measurement with its own large $\sigma$.

Serial elimination reconciles, and while the global test fails it moves the redundant measurement with the largest $z_i$ to the unmeasured set and reconciles again. Each step reports the removed stream, its estimated gross error (its measurement less its estimate once removed) and any stream whose test equals the largest to within $10^{-7}$ relative, an equivalent set that no test can split. The loop stops when the test passes or no redundancy remains. A per-sample global test evaluates $\gamma$ on every row with the per-sample $\sigma$ and $(B\Sigma B^{\mathsf{T}})^{-1}$ formed once. For a record with a fast and a slow instrument of one quantity, a second mode pairs each laboratory result with the analyzer reading whose sample time is nearest its own and fits a least-squares line to the paired differences, whose slope is the analyzer's drift rate.

### 13.4.2 Principal-Component Monitoring

Each variable that moves in the training window is standardized by its training mean and standard deviation, a variable that does not move is left out and listed, and $l$ lagged copies may be appended for dynamic PCA [17]. The eigenvalues $\lambda_i$ and eigenvectors of the training correlation matrix come from the cyclic Jacobi method. The model keeps the smallest $k$ whose eigenvalues reach the chosen share of the total, or a fixed $k$. For a standardized row $z$ with scores $t = P_k^{\mathsf{T}} z$ and residual $e = z - P_k t$,

the two statistics and their limits are given by Equation (13.4).

$$\begin{aligned} T^2 &= \sum_{i=1}^{k} \frac{t_i^2}{\lambda_i}, \qquad Q = \|e\|^2 \\ T_\alpha^2 &= \frac{k(n^2-1)}{n(n-k)} F_\alpha(k, n-k) \\ Q_\alpha &= \theta_1 \Big[ \frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \Big]^{1/h_0} \end{aligned} \tag{13.4}$$

where $n$ is the number of training rows, $\theta_j = \sum_{i>k} \lambda_i^j$, $h_0 = 1 - 2\theta_1\theta_3/(3\theta_2^2)$ and $c_\alpha$ is the normal quantile, the residual limit of Jackson and Mudholkar [16]. A detection is the first sample after training that begins a run of three consecutive samples above a limit, a declared persistence rule that keeps isolated crossings from counting. Contributions to a chosen sample are the squares of the elements of $D^{1/2}z$, with $D = P_k \Lambda_k^{-1} P_k^{\mathsf{T}}$, for $T^2$ and the squares $e_j^2$ for $Q$. Both sets are non-negative, sum to their statistic, are shown as shares, and sum a variable's lagged copies into the variable.

### 13.4.3 Control Charts and Their Run Lengths

A chart is trained on a window, from which the mean $\mu$ and standard deviation $\sigma$ of one signal set $z_t = (x_t - \mu)/\sigma$, and its statistic starts from zero at the window's end. The Shewhart chart signals when $|z_t| > 3$ (rule 1), and the other three Western Electric rules (two of three beyond 2, four of five beyond 1, eight in a row on one side) are a switch. The EWMA and the two-sided CUSUM are given by Equation (13.5).

$$\begin{aligned} w_t &= \lambda z_t + (1-\lambda) w_{t-1}, \qquad |w_t| > L\sqrt{\lambda/(2-\lambda)} \\ C_t^{\pm} &= \max\big(0, C_{t-1}^{\pm} \pm z_t - k\big), \qquad C_t^{\pm} > h \end{aligned} \tag{13.5}$$

where the EWMA uses its asymptotic limit, the one run-length tables assume. The average run length (ARL) of each chart comes from a Markov chain on its discretized statistic [13], with transition matrix $R$ and run lengths $a$ from $(I - R)a = \mathbf{1}$, on 120 states for a one-sided CUSUM and 151 for the EWMA. The two sides of a CUSUM combine as $1/\mathrm{ARL} = 1/\mathrm{ARL}^{+} + 1/\mathrm{ARL}^{-}$, and a side whose chain is singular to working precision, because the shift lies far on the other side, is given an infinite run length.

### 13.4.4 Loop-Performance Indices

On the last two thirds of a loop record the workbench forms the control error $e = \mathrm{SP} - \mathrm{PV}$ and its autocorrelation up to a third of the window, at most 400 lags, of which the lag-one value of Equation (1.22) is the first. The decay ratio is the height of the first autocorrelation peak above the line joining the minima on either side, divided by the depth of the first minimum below the line from lag 0 to that peak, so 1 is an undamped cycle and 0 a damped one. The regularity of the zero crossings is $\bar{T}_p/(3 s_{T_p})$, with $T_p$ twice each interval between crossings, after Thornhill, Huang and Zhang [18]. A sample counts as held when the controller output moved by more than $\varepsilon$ while the valve moved by no more than $\varepsilon$. The held share of all samples on which the output moved, the median output travel during a hold (the stick band) and the number of holds ending in a jump describe the stick-slip pattern [19]. The minimum-variance benchmark fits an autoregressive model of order $p$ to the mean-removed error by linear least squares, the same fit as the ARX model of Section 1.4.8 without an input, as shown in Equation (13.6).

$$e_t = \frac{a_t}{A(q^{-1})}, \quad \sigma_{mv}^2 = \sigma_a^2 \sum_{j=0}^{d-1} \psi_j^2, \quad \eta = \frac{\sigma_{mv}^2}{\sigma_e^2} \tag{13.6}$$

where $\psi_j$ is the impulse response of $1/A$, $d$ is the declared process delay in samples, $\sigma_a^2$ the innovation variance and $\eta$ the Harris index [4], 1 at minimum variance and near 0 far from it.

### 13.4.5 Alarm Metrics

An alarm watches one signal, optionally filtered by $y_k = y_{k-1} + \Delta t\,(x_k - y_{k-1})/(\tau_f + \Delta t)$. A high alarm's condition sets when $y > L$ and clears only when $y < L - \mathrm{DB}$, with the deadband DB in percent of the tag's span, and a low alarm mirrors it. The alarm annunciates once its condition has held for the on-delay and returns once the clear condition has held for the off-delay. Metrics are counted in fixed ten-minute periods from the record's start: annunciations, the rate per ten minutes and per day, the busiest period, flood periods, chattering tags and the annunciations in their bursts, the alarms standing at the end and each tag's share. "Standing at the end" is the set active when an hour-long record stops, not the standard's stale alarm, which an hour cannot show.

### 13.4.6 Assumptions

Reconciliation assumes linear balances at steady state over the window and independent errors with the declared $\sigma$, so component, energy and bilinear balances are outside it, and a window that contains a transient reads as a gross error. Principal-component limits and chart limits assume independent samples from a stationary normal operation. A closed-loop plant wanders between days and its samples are correlated, unlike the independent residuals of a correct model in Section 1.4.9, so both kinds of limit are crossed more often than their nominal rate, which Section 13.5.3 measures. A contribution plot ranks the variables that carry the statistic at one sample, which is where a fault shows and not necessarily where it starts. The loop indices assume uniform sampling (a non-uniform record is resampled to its median interval) and a declared delay, and the stiction screen needs the valve position. Alarms are limit crossings on single signals, with no state-based suppression, shelving or priority.

## 13.5 Ten Cases, Measured

The ten cases and what each measures at the default settings are listed in Table 13.3. Each number below comes from the shipped records, measured with the workbench's own functions in a test that imports them, and the reconciliation, PCA, chart, loop and alarm numbers were also re-derived by independent code (Section 13.7).

### 13.5.1 The Meter That Lies

On the window from 6 to 30 h, 97 samples per stream, the two balances have rank 2 and the global test gives $\gamma = 313.67$ against the $\chi^2(2)$ limit of 5.99 at $\alpha = 0.05$. All seven measurement tests exceed 2.68, the station meter at 17.71, the pump curves at 13.20 and the five junction streams at 11.68 each, because the five form an equivalent set within one balance and one error spreads equally into them. Serial elimination removes the station meter first, with no tie, and estimates its gross error at 2.35 L/s, close to 3 % of its 77.85 L/s mean. The balances then close with $\gamma = 0.135$ against the $\chi^2(1)$ limit of 3.84, and every remaining test reads 0.37 against 2.63. Removal moves the reconciled station and pump flows from 77.81 to 75.50 L/s, the tank inflow from −0.46 to −2.32 L/s and the trunk leakage from 3.29 to 2.87 L/s, as drawn in Figure 13.1. The window from 0 to 4 h, before the shift, passes with $\gamma = 0.07$. The same state of the workbench is captured in Figure 13.2.

🕮 **Pause.** Try Exercise 13.1 before reading on.

Two further runs show what the structure allows. With the trunk leakage declared unmeasured, the junction balance projects away, one reduced balance remains ($\gamma = 177.21$), and the first elimination step reports the pump curves and the station meter as tied, because a meter in series with the only other measurement of its flow cannot be told from it. Around the main break, from 9 to 12 h with the station meter removed, $\gamma = 0.021$ against 3.84, because the broken main's flow passes through the zone 3 meter (22.33 L/s on that window) and the balance still closes.

**Table 13.3** The Ten Cases, What the Default Settings Measure on the Shipped Record, and What Completes Each Case

| # | Case and record | Measured with the defaults | Required to complete |
|---|---|---|---|
| 1 | False-alarm rate, normal days | $k = 11$, limits 25.22 and 42.59, no detection, $Q$ above its limit on 5.0 % of the third day, one-day training detects at 29.70 h | Train on 0 to 48 h, then on one day or less, then confirm the rate |
| 2 | IDV(1) feed ratio step | $Q$ 48.05 h, $T^2$ 48.15 h, a feed variable first in $Q$'s top three at 49.80 h | Find $Q$'s detection, then pick a sample from about 50 h whose top three include a feed variable |
| 3 | IDV(5) cooling water step | $Q$ 48.05 h, $T^2$ never ($Q$ 12.5 %, $T^2$ 0.8 % above) | Train with $Q$ detecting within an hour and $T^2$ not at all |
| 4 | IDV(13) slow drift, product flow | EWMA and CUSUM 52.00 h, Shewhart rule 1 52.45 h, all four rules 49.15 h | Chart one signal with Shewhart and EWMA, the EWMA first |
| 5 | Water network | $\gamma = 313.67$, elimination removes the station meter, $\gamma = 0.135$ | Fail the global test, then let elimination name the meter |
| 6 | Digester, analyzer and laboratory | six pairs, drift 0.2944 kappa/h | Name the drifting instrument and its rate within 50 % |
| 7 | Loops A and B | held 93.9 % against 0.4 %, $\eta$ 0.009 and 0.341 | Diagnose both loops |
| 8 | Alarm flood | 494 annunciations and 6 floods fall to 14 and 0 | No flood and at most a tenth of the raw count |
| 9 | Unlabeled record (IDV(4)) | $Q$ and $T^2$ 48.05 h | Class, detection within 15 min, one of three variables |
| 10 | IDV(3) D feed temperature | $Q$ detects at 54.75 h, above its limit on 8.3 % (normal 5.0 %) | Train on the fault record and on the normal record, then confirm |

**TRY IT IN THE APP 5 · The Water Network: Find the Meter That Lies**

Load the case and press Reconcile. The global test reads 313.67 against its limit of 5.99, and every measurement test is above its limit. Then press Serial elimination and watch the station flowmeter turn red: it is removed with a gross error of 2.35, and every remaining stream passes its test.

The per-sample global test is a warning about the derived tank inflow. Before the shift, 3 of 15 samples exceed 5.99, each within one block of a step of the pump flow, at 1.25, 2.25 and 3.25 h. With the declared shift removed over the whole record, 11 of the 21 samples within one block of the seven steps exceed it, against 1 of the 98 samples away from them, whose mean $\gamma$ is 0.84 against the $\chi^2(2)$ mean of 2. A centered difference spreads each level step over two blocks, and away from the steps the declared $\sigma$ are conservative rather than small.

### 13.5.2 Two Statistics and a Contribution

On every Tennessee Eastman record 50 of the 52 signals vary in training (the compressor recycle and stripper steam valves are constant), the training window holds 961 rows, and 11 components carry 50.3 % of the variance. The limits are 25.22 for $T^2$, from $F(11, 950)$, and 42.59 for $Q$, from the 39 discarded eigenvalues. On the normal record nothing is detected on the third day, while $Q$ lies above its limit on 5.0 % of samples and $T^2$ on 0.2 % (0.8 % for $Q$ inside training). Two departures from a first design are measured choices. A model trained on one day detects falsely at 29.70 h ($T^2$ at 29.35 h), because the closed-loop plant wanders between days, and a model at 90 % variance keeps 33 components and detects falsely at 65.05 h.

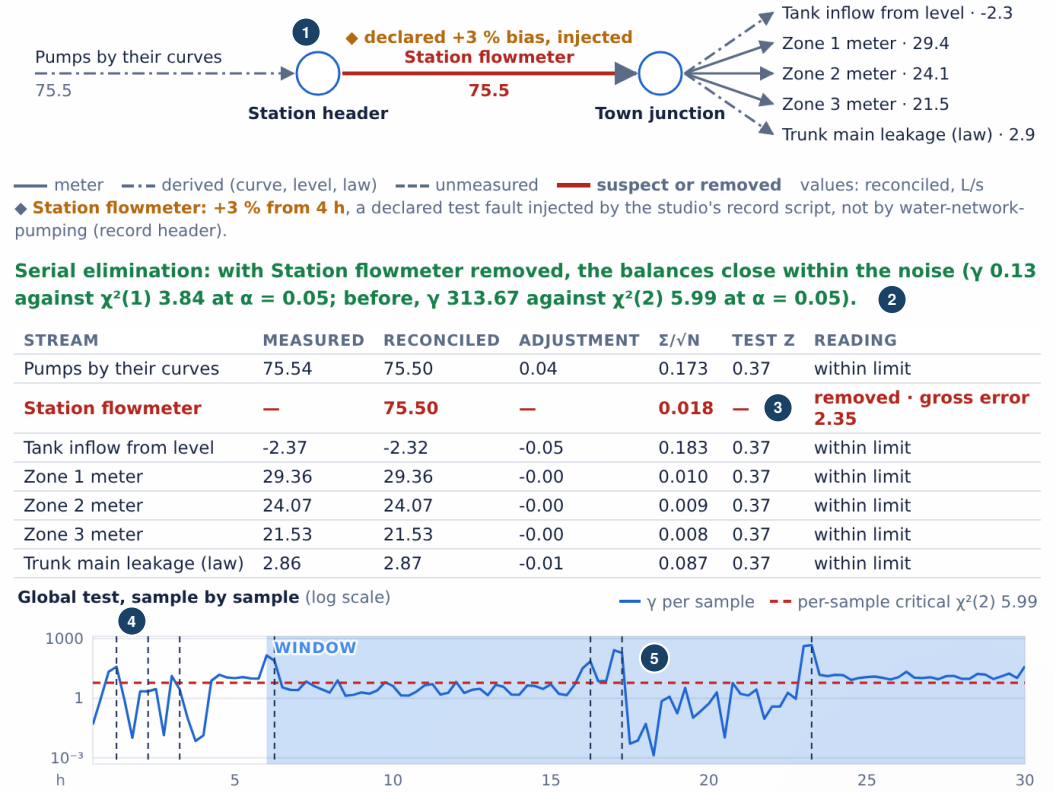



**Figure 13.2** The Reconcile tool on case 5 after serial elimination, window 6 to 30 h at $\alpha = 0.05$, cropped to its result panel. (1) The diagram tags the station flowmeter with the declared 3 % shift once elimination has run and shows reconciled values in L/s. (2) The verdict line reads $\gamma = 0.13$ against $\chi^2(1)$ 3.84 after removal and 313.67 against $\chi^2(2)$ 5.99 before. (3) The table reads the station meter removed with a gross error 2.35 and every other test 0.37 within its limit. (4) The per-sample global test, on a log scale, marks the steps of the pump flow with dotted lines and (5) shades the window.

> ✦ **DID YOU KNOW?** **Five Percent Above a 99 Percent Limit**
>
> From the simulator. A 99 % limit on independent samples is crossed on about one sample in a hundred. On the normal Tennessee Eastman record, $Q$ lies above its limit on 5.0 % of the third day (case 1 in Table 13.3), and over 61 normal days on 4.3 % of samples (Section 13.5.3).

The cooling water step of case 3 separates the two statistics, as shown in Figure 13.3. $Q$ detects at 48.05 h, the first sample after the step, and lies above its limit on 12.5 % of the following day, while $T^2$ never detects and lies above on 0.8 %. The likely mechanism is that the temperature loops absorb the step in directions the 11 retained components did not span in training, which moves the residual and leaves the scores inside their limit. With two lags the model has 150 columns and 17 components, and $T^2$ then detects at 48.10 h. A 99.9 % limit would remove most false detections on normal days but would detect this fault with neither statistic.

🕮 **Pause.** Try Exercise 13.2 before reading on.

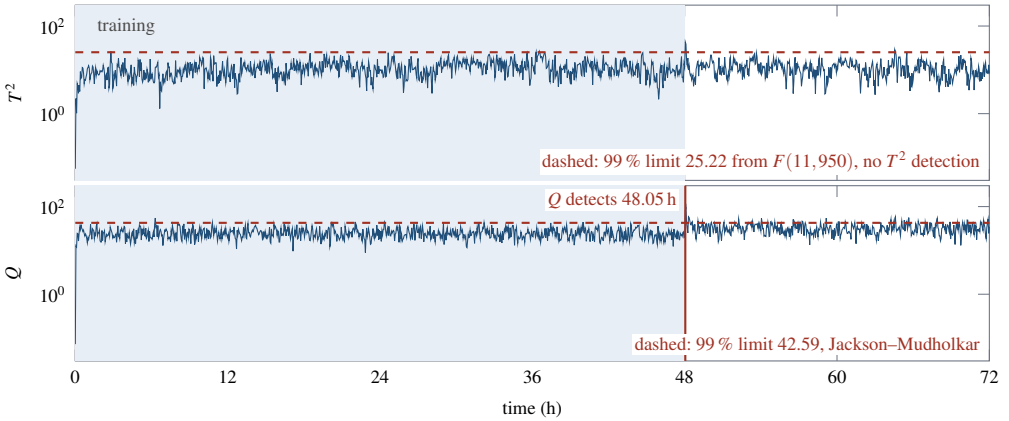



**Figure 13.3** $T^2$ and $Q$ on the cooling water step record (case 3), trained on 0 to 48 h (shaded) with 11 components, against their 99 % limits. $Q$ detects at 48.05 h, the first sample after the step, and $T^2$ never detects.

On the feed ratio step of case 2, $Q$ detects at 48.05 h and $T^2$ at 48.15 h, but the three largest $Q$ contributions at detection are the reactor temperature, the reactor cooling water valve and its outlet temperature. None of the 35 samples between the step and 49.80 h has a feed variable among its three largest, and from 49.80 h on every sample has one, the reactor feed analysis of A beside the compressor work and the separator liquid valve. The variable that moves first is where the temperature loops answer within a sample or two, while the reactor feed analyzer samples every 0.1 h and reports 0.1 h later. On the unlabeled challenge record $Q$ and $T^2$ both detect at 48.05 h, and the largest mean $Q$ contributions over the following hour are the reactor cooling water valve, the reactor temperature and the reactor cooling water outlet temperature, which is the answer the challenge accepts within 15 min and one of three variables.

### 13.5.3 A Fault the Statistics Cannot See

The D feed temperature step of case 10 is absorbed by the control system. On the shipped seed $Q$ detects at 54.75 h and lies above its limit on 8.3 % of samples after 48 h, against 5.0 % on the normal record, and $T^2$ never detects. One seed cannot decide whether that is a detection, so both records were regenerated for 61 seeds (1 to 60 and the shipped one) with the workbench's own generator and judged with its own monitor at the defaults. Normal days raise 76 detection episodes in 61 days, $Q$ or $T^2$ runs of three merged when they start within three samples, which is 1.25 per day (95 % interval 0.7 to 1.8) on 28 of the 61 days, with $Q$ above its limit on $4.3 \pm 2.4$ % of samples against the nominal 1 %. Independent samples above the limit that often would give about 0.04 runs of three per day. The fault days raise 2.02 episodes per day on 36 of 61 days, with $Q$ above on $7.0 \pm 3.5$ % of samples. The paired difference in $Q$'s share is positive on all 61 seeds, by 2.7 points on average, but a monitor never has the paired normal day. A whole-day test on $Q$'s share, set at the normal days' 95th percentile of 8.1 %, flags 20 of the 61 fault days, and only once the day is over. This fault is therefore not detectable at the statistics' false-alarm rate, which is the result the case asks the learner to state.

🕮 **Pause.** Try Exercise 13.3 before reading on.

### 13.5.4 Drift in a Chart and in Time

The product flow of the slow-drift record (XMEAS 17), trained on 0 to 48 h with a mean of 22.9494 and a standard deviation of 0.1169, gives the EWMA and the CUSUM their first signal at 52.00 h and Shewhart rule 1 at 52.45 h. The two accumulating charts signal on the same sample because of the record, where one sample at $z = -2.46$ takes $C^-$ from 3.25 to 5.21 and the EWMA from $-0.80$ to

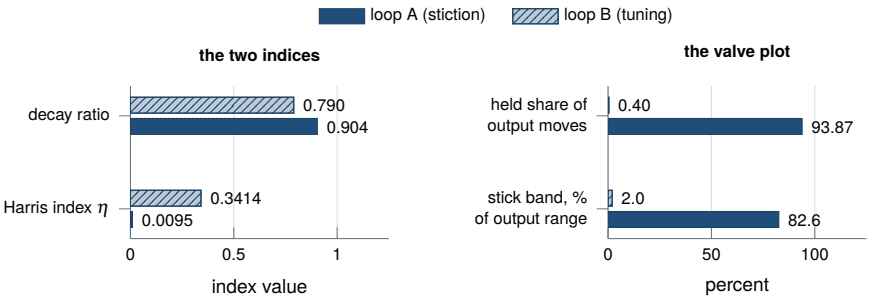



**Figure 13.4** Loop indices of the two loop records, from Table 13.4. Notice that both loops show a high decay ratio and a Harris index far from 1, so the indices do not tell them apart, while the valve plot does: loop A's valve held still on 93.87 % of the samples on which the output moved, loop B's on 0.40 %.

**Table 13.4** Loop Indices on the Last Two Thirds of Each Record, Delay 10 Samples, AR Order 10

| Index | Loop A (stiction) | Loop B (tuning) |
|---|---|---|
| samples, period | 801, 0.1 min | 4001, 0.1 s |
| decay ratio | 0.904 | 0.790 |
| regularity (crossings) | 81.99 (4) | 12.88 (33) |
| oscillation period | 13.82 min | 2.42 s |
| held share of output moves | 93.87 % | 0.40 % |
| stick band, % of output range | 82.6 | 2.0 |
| Harris index $\eta$ | 0.0095 | 0.3414 |
| stiction screen | stiction | none |

$-1.13$ against a limit of 0.987. On the normal record neither the EWMA nor rule 1 signals, and the CUSUM signals at 64.85 h. The four Western Electric rules together fire at 49.15 h on both records, which is why rule 1 alone is the default. The EWMA's chain gives an in-control run length of 499.3 samples, about 25 h at 3 min, against 370.4 for rule 1. On the extra cabin record, a CUSUM on the nitrogen balance residual trained on the first day signals at 25.25 h, 1.25 h after the leak step, while the cabin simulator's own leak flag appears at 27.08 h in the five-minute record.

In the digester's in-time mode the six laboratory samples at 0, 4, 8, 12, 16 and 20 h pair with analyzer readings whose differences grow 0.05, 0.67, 2.60, 3.37, 4.35 and 5.93, a slope of 0.2944 kappa/h against the 0.3 kappa/h drift the simulator declares. The laboratory reads 29.81 at the first pair and 26.90 at the last, while the loop holds the drifting analyzer at its setpoint.

### 13.5.5 Loops and Alarms

The two loop records give the indices in Table 13.4. Both cycle with a high decay ratio and both read far from minimum variance, so neither index tells them apart. The valve plot does. On loop A the valve held still on 93.9 % of the samples on which the output moved, with a stick band of 3.56 output points (83 % of the output's range) over 11 holds ending in a jump. On loop B it held on 0.4 %, and the gate follows its command. Two caveats apply. Loop A's regularity rests on four zero crossings, three intervals, and its Harris index depends on how the autoregressive model is fitted, 0.0095 by least squares against 0.118 by Yule-Walker equations, although both read far from 1. Figure 13.4 draws the indices and the valve plot's shares of Table 13.4 side by side.

The raw alarm log annunciates 494 times in the hour, 73, 77, 30, 56, 137 and 121 per ten-minute period, so all six periods are floods, and five chattering tags carry 473 of the annunciations. A deadband of 2 %, an on-delay of 15 s, an off-delay of 60 s and a 5 s filter leave 14 (1, 0, 3, 10, 0 and

0), no flood and no chattering tag, while 1 %, 10 s and 30 s without the filter leave 16 with one flood of 11 between 30 and 40 min. The limits sit close to normal operation on purpose, as unrationalized limits often do, and every threshold is a teaching value of the kind the standard's class of documents defines [5].

> **TRY IT IN THE APP 8 · An Alarm Flood: Rationalize It**
>
> Load the case and read the raw count: 494 annunciations in the hour, and every ten-minute period is a flood. Set Deadband to 2 % of span, On-delay to 15 s, Off-delay to 60 s and Filter $\tau$ to 5 s. The count falls to 14, with no flood and no chattering tag.

## 13.6 MPC Health From an Exported Run

A sixth tool reads an MPC run without simulating its plant. The Health tab maps measured output, set point and applied input to the record's own columns. Where present, `ctrl_mpc`, `mpc_fallback`, `mpc_status`, `mpc_slack_*`, `mpc_J` and `mpc_iters` supply active time, fallback episodes, soft-band violations, objective and solver work. The time of a flag is the interval from its timestamp to the next, and the final row adds no duration. Consecutive fallback flags count as one episode, not as several failed solves. Positive slack is a violation, while zero slack does not identify a binding constraint. Measured time at or beyond a bound requires an explicitly mapped bound column. Missing quantities are reported unavailable.

A forecast column is scored against a measured output only at its declared future timestamp. With measured $y$ and forecast $\hat{y}_{k+H|k}$, the residual is $y_{k+H} - \hat{y}_{k+H|k}$. RMSE and mean residual use exact matching future timestamps, not interpolation across data gaps. A held endpoint whose lead time changes between solves cannot be scored with one constant $H$. Input travel sums successive absolute changes and breaks at missing input samples. Categorical solver status stays aligned when rows are sorted into time order.

A CSV file downloaded from an app with the MPC option (Sections 1.6.3 and 1.6.4) can carry the accepted controller plan in a comment header. This is a copy of the plan in charge, not a retrospective simulation or an exported optimization model. Output predictions and held future moves share the measured run's time scale. A rejected or unavailable plan contributes no snapshot. The tab marks its origin and extends the set point by an explicitly labeled hold of the last recorded value, without claiming recipe preview. A saved snapshot may also be opened beside a longer matching record to assess its future predictions, and input changes after the snapshot contribute to that residual too. The tab links to a GEKKO resource for further study and does not produce a GEKKO model.

The workbench carries this tool as an extra case, "MPC health", shown in Figure 13.5, which asks the learner to inspect the difference between a prediction and a measured outcome and to record an analysis. Numerical checks exercise unequal sampling intervals, missing values, duplicate timestamps, a fixed-horizon forecast, grouped fallback flags, malformed snapshots and held moves. The case's record is regenerated from the microgrid's own engine with wall-clock costs removed. The original ten records and their completion judgments are unchanged. The alarm record no longer regenerates identically from its upstream plants. It is retained rather than silently replacing the evidence behind the alarm results of Section 13.5.

## 13.7 Verification

### 13.7.1 Code Verification

When the records were built, the workbench's own suite passed 79 tests (one large-data test is skipped without its environment), and regenerating the records reproduced all 13 files byte for byte. The reconciliation was re-derived independently by solving the Karush-Kuhn-Tucker system of the constrained least squares and taking the adjustment covariance from the solution operator. On 138 random networks it agrees with the projection route to $3.7 \times 10^{-15}$ in $\gamma$ and $8.2 \times 10^{-14}$ in $z$, and on

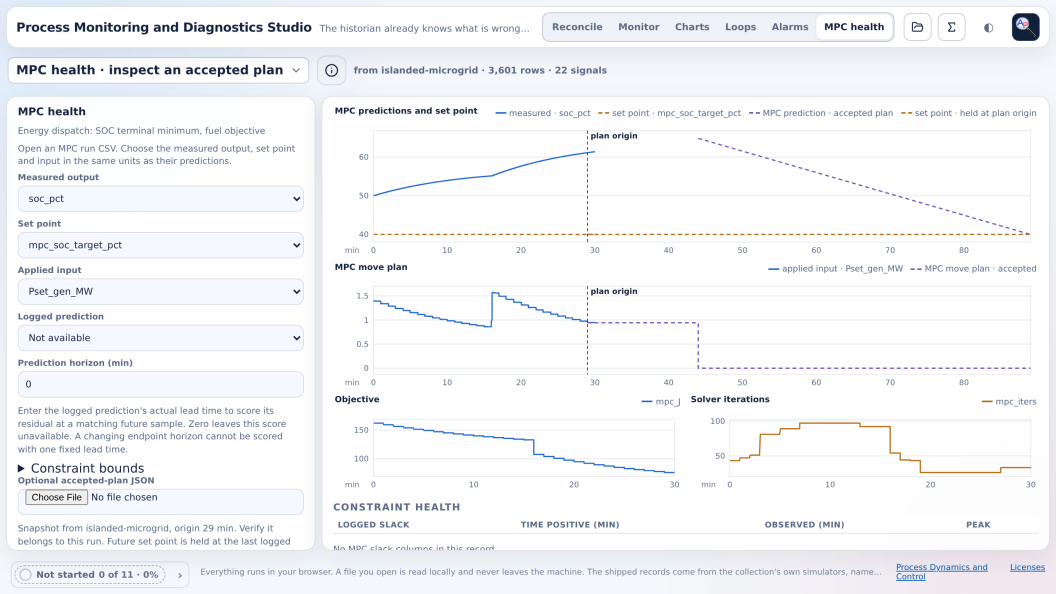



**Figure 13.5** The MPC health case: a measured islanded-microgrid MPC run and its accepted future snapshot. The upper plot compares battery SOC with its terminal SOC target and the accepted prediction. The lower plot compares the genset command with its held move plan. The origin marks when this plan was accepted. Its future is a controller prediction, not a measured outcome. Objective and solver work are read from the exported record.

the water record it gives the same $\gamma$, tests, gross error and reconciled values. Under simulated noise on a four-balance network the global test's 20 000 draws have a mean of 3.992 and a variance of 7.99 against 4 and 8, with 5.13 % false alarms at $\alpha = 0.05$. On a synthetic rank-2 set of eight variables the PCA keeps two components and $Q$ exceeds its 99 % limit on 0.920 % of 20 000 new normal rows ($T^2$ on 0.995 %). An independent Householder and QL solver reproduces the $T^2$ and $Q$ limits on the plant records to the quoted digits, and the Jacobi solver reconstructs 27 symmetric matrices of order 2 to 96, random, clustered within $10^{-9}$ and rank-deficient, to $6 \times 10^{-12}$ of their largest eigenvalue. The Harris index of a first-order moving-average error with coefficient 0.5 is 0.800 at $d = 1$ and 1.000 at $d = 2$, the exact values, and a synthetic alarm log reproduces its hand counts.

The run lengths were checked three ways. For the one-sided CUSUM with $k = 0.5$ the chain, a Nyström solution of the run-length integral equation on Gauss-Legendre nodes, and a Monte Carlo of 400 000 runs give 77.07, 77.08 and $77.06 \pm 0.11$ at $h = 4$ and a shift of 0.25, and 38.01, 38.01 and $38.00 \pm 0.05$ at $h = 5$ and a shift of 0.5. Doubling the chain from 120 to 240 states moves the first value from 77.073 to 77.077. The EWMA with $\lambda = 0.2$ and $L = 2.962$ has an in-control run length of 499.3 by the chain and 499.7 by the integral equation, and the two-sided CUSUM with $h = 5$ has 465.2 by the chain.

### 13.7.2 Agreement With Published Sources

The CUSUM chain agrees with the run-length table printed in the NIST/SEMATECH e-Handbook for $k = 0.5$ within 1.2 % in 18 of its 20 cells [15]. Two printed cells disagree, 74.2 at $h = 4$ and a shift of 0.25 and 30.0 at $h = 5$ and a shift of 0.5, where the three independent routes above agree on 77.07 and 38.01.

The reconciliation reproduces every structural statement of the published example of Narasimhan and Bhatt (Example 1 and Example 2 of the preprint) [10]. With all six streams measured the reconciled errors of F1 and F5, F2 and F6, and F3 and F4 are equal, and with only streams 1, 2 and 5

measured the projection leaves the single balance F1 = F5 with F2 non-redundant. The example's Table 2 prints, for 1000 simulated samples, root sums of squared errors before and after reconciliation (labeled RMSE). Taking the printed "before" values as the error scales and the printed standard deviations as the weights, independent errors predict the four "after" values of F1, F3, F4 and F5 within 1.6 % but F2 and F6 5.9 % high. The printed "before" value of F6 is itself 0.862 times what its stated standard deviation of 0.1 gives, 6.2 standard errors low. A single error correlation of −0.14 between F2 and F6, which is 4.4 standard errors from zero for 1000 independent pairs, reproduces all six printed "after" values within 1 %. The gap therefore comes from the example's simulated errors rather than from the method, and the workbench reproduces the method.

### 13.7.3 Runtime

Every wall time below is a bound over repeated runs on the development host. A record of one million rows and six signals (55.8 MB of text) parses in under 0.5 s, one decimated chart draws in under 10 ms, a PCA fit and score takes under 0.2 s, an EWMA under 10 ms and an alarm evaluation under 15 ms. In a desktop browser at 1366 × 768 a shipped case opens in under 0.4 s on a fresh page and in under 30 ms once its record is fetched, and training, serial elimination and an alarm setting each take under 0.1 s.

## 13.8 Technology in Practice

The workbench has no instruments of its own. It reads records, and the devices that made each record decide what its tools can find. Figure 13.6 draws two of the records as a plant would take them.

#### What Measures It

The water record's station flowmeter is a simulated meter that the record script multiplies by 1.03 from 4 h, because the simulator's meters have no bias option. A magnetic flowmeter suits most water-based liquids, and its reading is linear in flow with one proportionality constant [20]. An error in that constant would scale every reading by one factor, which is the form of the shift the record declares. The alarm log's thirteen limits are soft alarms, raised when a continuous measurement crosses a programmed limit, as distinct from hard alarms raised by process switches. A distributed control system time-stamps and archives its alarms for later analysis [20].

> **TECHNOLOGY** **Sensor: A Magnetic Flowmeter**
>
> A conductive liquid moving through a magnetic field induces a voltage across the pipe, and two electrodes opposite each other in the pipe wall pick it up. The voltage is proportional to the mean velocity, so the meter is linear in volumetric flow with one proportionality constant, and its open bore leaves very little pressure loss. It needs a reasonably conductive liquid and a full pipe, with the electrodes horizontal so that a passing bubble does not break contact. The induced voltage is often a millivolt or less and is easily buried in noise from stray currents unless the flowtube is grounded. The coils are driven with alternating polarity, because a steady field would polarize the ions at the electrodes and corrupt the reading [20].

#### What Moves It

The workbench moves nothing. The final elements are in the records: loop A's steam valve carries its simulator's own 3 % stiction deadband, and the stiction screen of Section 13.4 needs the valve position beside the controller output. On a plant that position comes from the valve's positioner. A process switch has a deadband for a related reason: without one, a variable that sits at the trip point cycles the switch between trip and reset and raises repeated nuisance alarms [20]. The deadband of the alarm tool does the same in software.

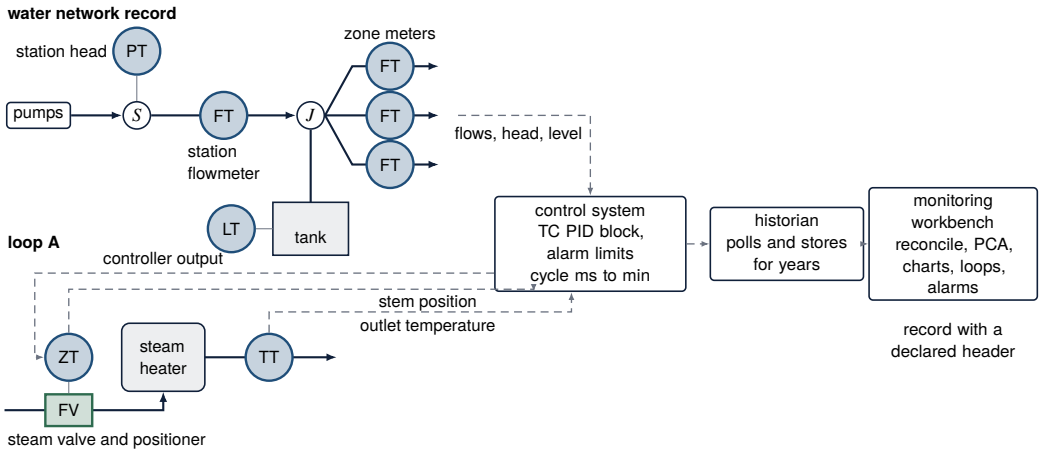



**Figure 13.6** Two of the chapter's records as a plant would take them. Top: the water network's station flowmeter between the station header *S* and the town junction *J*, the station head, the three zone meters and the tank level of Figure 13.1. Bottom: loop A, a steam heater's outlet-temperature loop whose valve positioner reports the stem position that the stiction screen needs. The control system runs the loop and its alarm limits, the historian keeps what it polls, and the workbench reads a record. Notice that the workbench sits at the end of the chain and acts on no device.

**TECHNOLOGY** **Final Element: A Valve Whose Positioner Reports Its Stem**

Packing friction and the force of the process on the plug keep a pneumatic valve's stem from following its actuator pressure exactly. A positioner compares the stem position with the command and adjusts the actuator pressure until they agree, a small control loop of its own. An electronic positioner also reports diagnostic data such as supply and actuator pressure, position error and total stem travel. Its plot of actuator pressure against stem position, the valve signature, shows packing friction as the offset between the opening and the closing traces. A signature recorded after a rebuild can be compared with a later one to judge wear without taking the valve apart [20].

### Where the Controller Runs

The workbench runs in a browser on a stored record: on a record of one million rows a PCA fit and score take under 0.2 s. The loops that made the records run in plant controllers with cycle times from milliseconds to minutes [21], and a data historian polls the control system at regular intervals and keeps the data, typically for years [20]. The shipped records are sampled from every 0.1 s to every 15 min, and a record can be far coarser than the engine that made it: the Tennessee Eastman engine steps every 1 s and its records keep one row in 3 min.

**IN PRACTICE Rate Targets and Quick Fixes for Alarms**

The guidance of the Health and Safety Executive starts by measuring: how many alarms there are, how many occur in normal operation and in an upset, and how many stand. Its targets are a long-term average of no more than one alarm every ten minutes in normal operation, and no more than ten displayed in the first ten minutes after a major upset. Priorities rest on the consequence of a missed response, with about three levels in proportions such as 5, 15 and 80 percent. Quick fixes include removing alarms that have no defined operator response, tuning the settings of nuisance alarms and adjusting the deadbands of repeating ones [6].

### What Is Modernizing It

The first development is the electronic valve positioner, which measures actuator pressure as well as stem position, keeps a count of total stem travel and can predict when the packing will wear out and raise a maintenance alarm [20]. Where such data reach the historian, excess packing friction can be read from the valve's own signature instead of being inferred from the loop record. The second is machine learning. Fault diagnosis can be posed as classification, with sensor readings as the inputs to a classifier that predicts the fault or its absence, and an open benchmark on simulations of the Tennessee Eastman plant now tests eleven strategies for carrying such a classifier from one operating mode to another [22]. The principal-component monitor of Section 13.4 is trained on normal operation alone, while a classifier needs labeled examples of every fault it is to name.

## 13.9 Limitations

Reconciliation is linear and steady, so component, energy and dynamic balances are absent, and the water record needed three derived streams before its biased meter could be identified at all, because the junction's tank and zone streams form an equivalent set within one balance. The PCA is static or lagged, with no kernel, recursive or moving-window form, no partial least squares, and contributions that point where a fault shows. Its limits and the chart limits assume independent samples, which a closed-loop record violates, and the measured false-alarm rate of 1.25 detections per normal day is a property of these defaults on this plant, not a general rate. The loop indices are screening tools that need a measured valve position and a declared delay, and the Harris index of a limit cycle depends on the autoregressive estimator. The alarm metrics treat each tag alone, with no shelving, state-based suppression or priority, and their thresholds are teaching values. Every record is generated, so the workbench has not been tested on a plant's own historian. Each record is also stylized. A real station meter would drift rather than shift in one block, while the declared step shows how a gross error spreads into a balance and how elimination removes it. The cross-check of the workbench's Jacobi solver against the storage-ring simulator's decomposition waits for that simulator.

**OPEN CHALLENGE A Monitor Trained in One Operating Mode**

A fault classifier learned from historical data assumes that new data follow the same probability distribution as its training data, which is seldom verified in practice. The Tennessee Eastman plant has six modes of operation, set by the product mass ratio and the production rate, and each mode changes the statistics of the measured series, so a model trained in one mode may not generalize to another. A benchmark built on simulations of the plant tests methods that adapt a classifier with unlabeled data from the new mode [22]. Progress would be a monitor that keeps its false-alarm rate when the plant changes mode, without a labeled fault record for every mode.

## 13.10 Conclusion

This chapter studied a browser workbench for monitoring and diagnosing plant records, with five tools acting on records generated by other simulators of the collection with every change declared. On the water-network record a 3 % meter shift raises the global test to 313.67 and every measurement test above its limit, and serial elimination names the station meter with a gross error of 2.35 L/s,

after which the balances close at 0.135. On the Tennessee Eastman records $Q$ detects the cooling water step at 48.05 h while $T^2$ never does, a contribution plot reaches the feed only 1.8 h after the feed ratio step, and the D feed temperature step is not detectable at the measured false-alarm rate of 1.25 detections per normal day over 61 seeds. The loop records are told apart by the valve plot (93.9 % against 0.4 % held) and not by their indices, and rationalization takes an alarm hour from 494 annunciations to 14. Independent checks disagree with two cells of a published run-length table and trace a gap in a published example to its simulated errors. The methods are linear and assume independent samples, and every record is generated. A fuller workbench would add bilinear and dynamic reconciliation, recursive and kernel monitoring, a stiction test without the valve position, state-based alarm suppression and a recorded plant historian. The workbench shows that a statistic above its limit is the start of a diagnosis and not its answer.

**KEY POINTS**

- A statistic above its limit is the start of a diagnosis and not its answer.
- One gross error spreads into every stream that shares a balance with it: a 3 % shift on one flowmeter puts all seven measurement tests above their limit.
- Limits assume independent samples, so on a closed-loop plant $Q$ lies above its 99 % limit on 4.3 % of samples on normal days, and one fault is not detectable at that false-alarm rate.
- The loop indices do not tell a sticking valve from an aggressive loop, while the valve plot does (93.9 % against 0.4 % held).

**Where else this applies.** The drifting analyzer that its own loop holds at the setpoint is the kappa loop of Volume 5, Chapter 4, and the sticking steam valve of loop A belongs to Volume 2, Chapter 4.

## References


[1] R. S. Mah, G. M. Stanley, and D. M. Downing, "Reconciliation and rectification of process flow and inventory data," *Industrial & Engineering Chemistry Process Design and Development*, vol. 15, no. 1, pp. 175–183, 1976.

[2] S. Narasimhan and C. Jordache, *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. Gulf Professional Publishing, 1999.

[3] J. V. Kresta, J. F. MacGregor, and T. E. Marlin, "Multivariate statistical monitoring of process operating performance," *The Canadian Journal of Chemical Engineering*, vol. 69, no. 1, pp. 35–47, 1991.

[4] T. J. Harris, "Assessment of control loop performance," *The Canadian Journal of Chemical Engineering*, vol. 67, no. 5, pp. 856–861, 1989.

[5] International Society of Automation, "Management of Alarm Systems for the Process Industries," ANSI/ISA-18.2-2016, Research Triangle Park, NC, USA, 2016.

[6] Health and Safety Executive, "Better alarm handling," Health and Safety Executive, HSE Information Sheet Chemicals Sheet No 6 (CHIS6), 2000, information current at 3/00. [Online]. Available: https://www.hse.gov.uk/pubns/chis6.pdf

[7] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.

[8] R. S. H. Mah and A. C. Tamhane, "Detection of gross errors in process data," *AIChE Journal*, vol. 28, no. 5, pp. 828–830, 1982.

[9] C. M. Crowe, Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of process flow rates by matrix projection. Part I: Linear case," *AIChE Journal*, vol. 29, no. 6, pp. 881–888, 1983.

[10] S. Narasimhan and N. Bhatt, "Deconstructing principal component analysis using a data reconciliation perspective," *Computers & Chemical Engineering*, vol. 77, pp. 74–84, 2015, preprint arXiv:1505.00314.

[11] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1-2, pp. 100–115, 1954.

[12] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.

[13] D. Brook and D. A. Evans, "An approach to the probability distribution of cusum run length," *Biometrika*, vol. 59, no. 3, pp. 539–549, 1972.

[14] J. M. Lucas and M. S. Saccucci, "Exponentially weighted moving average control schemes: Properties and enhancements," *Technometrics*, vol. 32, no. 1, pp. 1–12, 1990.

[15] W. F. Guthrie, "NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151)," National Institute of Standards and Technology, 2020, section 6.3.2.3.1, Cusum Average Run Length. [Online]. Available: https://www.itl.nist.gov/div898/handbook/

[16] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, 1979.

[17] W. Ku, R. H. Storer, and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 30, no. 1, pp. 179–196, 1995.

[18] N. F. Thornhill, B. Huang, and H. Zhang, "Detection of multiple oscillations in control loops," *Journal of Process Control*, vol. 13, no. 1, pp. 91–100, 2003.

[19] M. A. A. Shoukat Choudhury, N. F. Thornhill, and S. L. Shah, "Modelling valve stiction," *Control Engineering Practice*, vol. 13, no. 5, pp. 641–658, 2005.

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

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

[22] E. Fernandes Montesuma, M. Mulas, F. Ngolè Mboula, F. Corona, and A. Souloumiac, "Benchmarking domain adaptation for chemical processes on the Tennessee Eastman process," arXiv preprint 2308.11247v2, 2024, published in Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Communications in Computer and Information Science, Springer, 2026, pp. 307–322, doi:10.1007/978-3-032-25305-7_21. [Online]. Available: https://arxiv.org/abs/2308.11247

## 13.11 Exercises

**13.1** ○ Concept check · 2 min

On the water-network record of case 5 only the station flowmeter carries the declared 3 % shift. On the window from 6 to 30 h, all seven measurement tests still exceed their limit of 2.68. Why?

(a) The declared standard deviations of the zone meters are too small, so noise inflates their tests.

(b) The main break at 10 h adds a second gross error, which lands in the zone 3 meter.

(c) The Šidák correction pulls the limit below the noise level that every test carries.

(d) The error enters both balances, and reconciliation spreads it into every stream that shares a balance with the meter.

**13.2** ○ Concept check · 2 min

On the cooling water step of case 3, trained with the defaults, $Q$ detects at 48.05 h while $T^2$ never detects. Which statement explains the difference?

(a) The step moves samples off the plane of the 11 retained components, which $Q$ measures, while their scores stay inside the $T^2$ limit.

(b) $T^2$ is built from the 39 discarded eigenvalues, so it is blind to the temperatures that the step moves.

(c) The $T^2$ limit is set at 99.9 % confidence and the $Q$ limit at 99 %, so $T^2$ needs a larger move.

(d) $T^2$ must stay above its limit for three samples and $Q$ for one, so $T^2$ responds later than $Q$.

**13.3** ○ Concept check · 1 min

Over 61 normal days, $Q$ lies above its 99 % limit on 4.3 % of samples rather than 1 %. What does the chapter give as the main reason?

(a) The Jackson-Mudholkar formula underestimates the residual quantile of any data set.

(b) The plant runs in closed loop, so its samples are correlated and its operation wanders between days.

(c) The two-day training window is too long and mixes in drift, which inflates the limit.

(d) The run-of-three rule triples the share of samples that are counted above the limit.

**13.4** Theory and calculation · 8 min

The water-network tests of Section 13.4 use a Šidák-corrected normal limit and a $\chi^2$ global test at $\alpha = 0.05$. The window means from 6 to 30 h ($n = 97$ samples) are, in L/s: pump curves 75.54, station meter 77.85, tank inflow $-2.37$, zone meters 29.36, 24.07 and 21.53, and trunk leakage 2.86.

(a) Compute the Šidák limit $z_c$ for $m = 7$ and $m = 6$ redundant measurements, and the Bonferroni limit $\Phi^{-1}(1 - \alpha/(2m))$ for $m = 7$. *Checkpoint:* 2.68, 2.63 and 2.69.

(b) The cumulative distribution of $\chi^2(2)$ is $1 - e^{-x/2}$, and a $\chi^2(1)$ variable is the square of a standard normal one. Find the 0.95 quantile of each. *Checkpoint:* 5.99 and 3.84.

(c) With the station meter removed, the station flow can be estimated by two independent routes: the pump curves ($\sigma = 1.7$ L/s per sample) and the sum of the junction outflows (tank 1.8, trunk 0.86 and zone meters 0.1, 0.093 and 0.076 L/s). Find each route's estimate and standard error over the window, combine them by inverse-variance weights, and compute the station meter's gross error. *Checkpoint:* 75.54 L/s $\pm$ 0.173 and 75.45 $\pm$ 0.203; combined 75.50 L/s $\pm$ 0.132; gross error 2.35 L/s.

**13.5** Theory and calculation · 6 min

The product-flow chart of case 4 (XMEAS 17) is trained with mean 22.9494 and standard deviation 0.1169 m$^3$/h, with the defaults $\lambda = 0.2$, $L = 2.962$, $k = 0.5$ and $h = 5$ of Equation (13.5).

(a) Compute the EWMA limit $L\sqrt{\lambda/(2-\lambda)}$. *Checkpoint:* 0.987.

(b) At the sample before the first signal, $C^- = 3.25$ and $w = -0.80$. The next sample has $z = -2.46$. Find the product flow it reads, and the new $C^-$ and $w$. Do the charts signal? *Checkpoint:* 22.662 m$^3$/h; $C^- = 5.21$, $w = -1.132$; both signal.

(c) The in-control run length of Shewhart rule 1 is $1/p$, with $p$ the two-sided probability beyond $3\sigma$. Compute it, and convert it and the EWMA's 499.3 samples into hours at the record's 3 min sampling. *Checkpoint:* 370.4 samples, 18.5 h; 25.0 h.

**13.6** Theory and calculation · 6 min

A detection on the Monitor tab is a run of three consecutive samples above a limit. The Tennessee Eastman records are sampled every 3 min.

(a) If samples were independent and each lay above the limit with probability $p$, a run of three starts at a given sample with probability about $(1 - p)p^3$. How many runs of three start per day at the nominal $p = 0.01$, and how many years pass between them on average? *Checkpoint:* $4.75 \times 10^{-4}$ per day, about 5.8 years.

(b) Repeat with the measured share $p = 0.043$ of normal days. *Checkpoint:* 0.037 per day.

(c) Over 61 normal days the defaults raised 1.25 detection episodes per day. By what factor does that exceed (b), and what does the factor say about the record? *Checkpoint:* About 34.

**13.7** App challenge · 8 min

Use scenario 5 · The water network: find the meter that lies. The Reconcile tab opens with the declared balances, the window 6 to 30 h and $\alpha = 0.05$. The record is a fixed file and the tools are deterministic, so every value below repeats to the digits the app shows.

(a) Press Reconcile. Read the global test $\gamma$ and its limit, and the test $z$ of the station flowmeter, the pump curves and the zone 1 meter. *Checkpoint:* $\gamma = 313.67$ against $\chi^2(2)$ 5.99; $z$ 17.71, 13.20 and 11.68.

(b) Press Serial elimination. Read which stream is removed, its estimated gross error and the new $\gamma$. *Checkpoint:* Station flowmeter, 2.35; $\gamma$ 0.13 against 3.84.

(c) Set the Kind of Trunk main leakage (law) to unmeasured and press Serial elimination again. Read the tiles and the elimination step, and explain the tie. *Checkpoint:* $\gamma = 177.21$ against $\chi^2(1)$ 3.84, one reduced balance; the pump curves are removed, tied with the station flowmeter.

**13.8** App challenge · 8 min

Use scenario 3 · A cooling water step: the two statistics disagree. On the Monitor tab keep the defaults (training 0 to 48 h, cumulative variance 50 %, lags 0, confidence 99 %, three in a row). The record is a fixed file, so the values repeat exactly.

(a) Press Train. Read $k$, both limits, both first detections and the share of samples above each limit after training. *Checkpoint:* $k = 11$; 25.22 and 42.59; $Q$ 48.05 h, $T^2$ none; 12.5 % and 0.8 %.

(b) Set Lags (dynamic PCA) to 2 and train again. Read $k$ and both detections. *Checkpoint:* $k = 17$; $Q$ 48.05 h, $T^2$ 48.10 h.

(c) Return Lags to 0, set Confidence to 99.9 % and train. What happens, and what does this say about raising the limit to cut the false alarms of case 1? *Checkpoint:* Neither statistic detects.

**13.9** App challenge · 7 min

Use scenario 8 · An alarm flood: rationalize it. The Alarms tab opens with no deadband, filter or delays. The metrics are counts over a fixed record and repeat exactly.

(a) Read the hour's annunciations, the number of ten-minute periods in flood, and the chattering tags and the annunciations they carry. *Checkpoint:* 494; 6; 5 tags with 473.

(b) Set Deadband 2 %, Filter $\tau$ 5 s, On-delay 15 s and Off-delay 60 s. Read the same metrics. Does the setting meet the case's target? *Checkpoint:* 14 annunciations, no flood, no chattering tag; yes.

(c) Set Deadband 1 %, Filter $\tau$ 0 s, On-delay 10 s and Off-delay 30 s. Read the total and the busiest period. *Checkpoint:* 16, with one flood of 11 between 30 and 40 min.

**13.10** Application area · 4 min

A water utility reads its station flowmeter for production reports. The meter of case 5 reads 3 % high from 4 h on.

(a) With the gross error of 2.35 L/s that elimination estimates, how much water does the station overstate per day? *Checkpoint:* About 203 $m^3$ per day.

(b) With the trunk leakage declared unmeasured, the pump curves and the station meter tie. What must a utility measure or model so that a biased station meter can be named?

**13.11** Application area · 4 min

In the digester record, the laboratory samples every 4 h and reports 1 h later, and the analyzer samples every 15 min and reports 20 min later. The six paired differences (analyzer less laboratory) at 0, 4, 8, 12, 16 and 20 h are 0.05, 0.67, 2.60, 3.37, 4.35 and 5.93.

(a) Fit a least-squares line to the differences against time. Compare its slope with the declared drift. *Checkpoint:* 0.2944 kappa/h against 0.3 kappa/h.

(b) An engineer pairs each laboratory result with the analyzer value on the screen when the result arrives. How far apart in time were the two samples taken, and how much bias does the drift alone add to the pair? *Checkpoint:* 30 min; 0.15 kappa.

**13.12** Review the agent · 5 min

An AI assistant reads the loop table of case 7 and writes: “Loop A has a decay ratio of 0.904 and a Harris index of 0.0095, against 0.790 and 0.341 for loop B, so loop A is the more badly tuned of the two. Halve its controller gain, and leave loop B alone.” Find the flaw, and say how you would check it in the app.