

2005

Likelihood and Bayesian Methods for Accurate Identification of Measurement Biases in Pseudo Steady-State Processes

Sriram Devanathan
Iowa State University

Stephen B. Vardeman
Iowa State University, vardeman@iastate.edu

Derrick K. Rollins Sr.
Iowa State University, drollins@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/imse_pubs

 Part of the [Industrial Engineering Commons](#), [Process Control and Systems Commons](#), [Statistics and Probability Commons](#), and the [Systems Engineering Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/imse_pubs/144. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

Likelihood and Bayesian Methods for Accurate Identification of Measurement Biases in Pseudo Steady-State Processes

Sriram Devanathan¹, Stephen B. Vardeman^{2,3} and Derrick K. Rollins, Sr.^{1,2,*}

¹Department of Chemical Engineering, Iowa State University, Ames, Iowa 50011

²Departments of Statistics, Ames, Iowa 50011

³Department of Industrial and Manufacturing Systems and Engineering, Ames, Iowa 50011

*Author to whom correspondence should be sent

Abstract

Two new approaches are presented for improved identification of measurement biases in linear pseudo steady-state processes. Both are designed to detect a change in the mean of a measured variable leading to an inference regarding the presence of a biased measurement. The first method is based on a likelihood ratio test for the presence of a mean shift. The second is based on a Bayesian decision rule (relying on prior distributions for unknown parameters) for the detection of a mean shift. The performance of these two methods is compared with that of a method given by Devanathan et al.¹. For the process studied, both techniques were found to have higher identification power than the method of Devanathan et al. and appears to have excellent but slightly lower type I error performance than the Devanathan et al. method.

Keywords: data reconciliation, gross error detection, likelihood ratio test, Bayesian statistics

1. Introduction

An important factor in process quality, safety and economy in chemical plant operations is the accuracy of measured process variables. Inaccurate process variables can lead to poor process control which can cause poor product quality, undetected material losses that can severely effect costs and environmental safety, and plant explosions that can lead to lost of life, equipment, and revenues. Ideally, one would like estimates of true values of process variables to have zero systematic deviation and minimal random deviation from the true values. When process measurements are biased, (i.e., systematically deviate) due to instrument miscalibrations or malfunctions, it is necessary to detect these biases and remove them. Techniques used to detect and identify systematic errors (such as those caused by biases) are termed Gross Error Detection (GED) methods. Over the past four decades a number of methods have been developed to identify and (mathematically) remove biases under various conditions. A summary of this literature can be found in Rollins et al.², and Bagajewicz and Rollins³.

For linear steady-state processes, identification strategies have traditionally been based on testing for material balance closure around each node in a process network (see Ripps⁴; Nogita⁵; Romagnoli and Stephanopoulos⁶; and Rosenberg et al.⁷). When there is closure at a node, the inference is that none of the measured variables associated with that node are biased. (The alternative inference is that there is a material leak at that node.) Rollins and Davis⁸ used such a nodal strategy for accurate identification of measurement biases. However, for certain combinations of biases, this approach and other commonly used methods such as the Serial Compensation Strategy (SCS) (Narasimhan and Mah⁹, also see Keller et al.¹⁰) perform poorly and are unable to completely or accurately identify the biases (e.g., the best conclusion could be

that at least two of the three suspect variables are biased). For such situations Devanathan et al. presented the Imbalance Correlation Strategy (ICS) which was shown to have a high probability of correct identification (large power) and low probability of false identification (small type I error probabilities). The ICS is based on observing changes in the sample correlation between the material balance at each node (interconnecting unit) and the measured variables associated with that node. However, for small sample sizes, the ICS appears to have low power for certain combinations of biases due to the cancellation of the effects of multiple biases in a material balance.

This article presents two new methods which are capable of completely identifying multiple measurement biases for these special cases with high probabilities of correct identification. These two techniques (unlike the nodal strategies) do not involve the use of process physical constraints, such as material and energy balances. Hence, they are not affected by the presence of leaks or error (bias) cancellation, and are applicable even for the case of nonlinear process constraints. However, these methods are limited to bias detection and identification. Therefore, other methods will be needed for bias estimation (see Mah and Tamhane¹¹; Narasimhan and Mah⁹; and Rollins and Davis⁸).

This article is organized as follows. First, the mathematical models are presented in Section 2 and the statistical tests are given in Sections 3 and 4 for the likelihood ratio test (LRT) and the Bayesian method (BM), respectively. Following these sections, performances of both the LRT and BM are evaluated using simulated data and the results are compared in Section 5. In the Concluding Remarks, we make recommendations for the cases where on-line testing may be important and also for processes that have bilinear or nonlinear constraints (i.e., simultaneous

mass, energy and component mass balances) (see Crowe¹²; and Rollins and Roelfs¹³) and give a brief discussion regarding the choice of technique for a given problem.

2. The Measurement Model

This section first presents the measurement model used in the study. The pseudo steady-state model following from the work of Rollins and Davis¹⁴ can be represented by

$$\text{with } y_{ij} = \mu_i + \delta_{ij} + \lambda_{ij} + \varepsilon_{ij} \quad (1)$$

$$\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon_i}^2) \quad (2)$$

$$\lambda_{ij} \sim N(0, \sigma_{\lambda_i}^2) \quad (3)$$

where y_{ij} is the measured value of variable i at the j th time instant; μ_i is the steady state true value of variable i ; λ_{ij} is the true value of the process deviation of variable i from μ_i at the j th time instant; δ_{ij} is the measurement bias of variable i at time j ; and ε_{ij} is the random error of variable i at the j th time instant. In this article it is assumed that the ε_{ij} 's for a fixed j are normally distributed with mean 0 and known variances and covariances. Additionally, the ε vectors for different time points j are assumed to be independent. Furthermore, ε_i is assumed to be independent of ε_k for $i \neq k$, and λ_{ij} is assumed to be independent of λ_{kj} for different measured variables i and k . Finally, the ε 's are assumed to be independent of the λ 's (i.e., measurement error is independent of process variability). In practice one will need to rely on an external source of knowledge to support the pseudo steady-state assumption.

This work initially confines attention to one process variable at a time, say variable i .

With $\mu_{ij} = \mu_i + \delta_{ij}$, Eq. 1 can be rewritten as

$$y_{ij} = \mu_{ij} + \eta_{ij} \quad (4)$$

where

$$\eta_{ij} = \lambda_{ij} \varepsilon_{ij}, \quad (5)$$

$$\sigma_{\eta_i}^2 = \sigma_{\varepsilon_i}^2 \lambda_{ij}^2, \quad (6)$$

$$\eta_{ij} \sim N(0, \sigma_{\eta_i}^2), \quad (7)$$

$$y_{ij} \sim N(\mu_{ij}, \sigma_{\eta_i}^2). \quad (8)$$

Based on the above, the density of y_{ij} is given by Eq. 8 and the joint density of $y_{i1}, y_{i2}, \dots, y_{in}$ is a product of its marginal densities, i.e., $\prod_{j=1}^n f(y_{ij} | \mu_{ij}, \sigma_{\eta_i}^2)$. Under the conditions given in the above model, the next two sections outline the decision rules involved in the identification of biased variables.

3. Likelihood Ratio Test For A Mean Shift

Let n denote the total number of available observations (i.e., the sample size), and M denote the number of observations before the initiation of a bias. Then the number of observations after the initiation is $n - M$. Suppose that $\delta_{ij} = 0$ for $j \leq M$ and $\delta_{ij} = \delta_i$ for $j > M$. When a bias occurs in the measurement of a process variable, the mean of that variable undergoes a shift. By comparing appropriate estimates of the mean of y_i it is possible to determine if there has indeed been a shift and if (a change in) measurement bias has occurred in variable i .

For the measurement model of this study it can be shown that if M is known, the maximum likelihood estimate (MLE) of μ_i is the sample average of the observations with $j \leq M$.

That is, the MLE for μ_i is

$$\bar{y}_{i|M} = \frac{\sum_{j=1}^M y_{ij}}{M}, \quad (9)$$

and the MLE for $\mu_i + \delta_i$ is

$$\bar{y}_{i2M} = \frac{\sum_{j=1}^n y_{ij}}{n \& M} \quad (10)$$

The objective is to determine if δ_i is significantly different from zero. To achieve this, a statistical hypothesis testing procedure is considered. The test statistic is based on the differences of the sample averages $(\bar{y}_{i2m} \& \bar{y}_{i1m}) = d_{im}$ for $m = 1, \dots, n - 1$, where M is set to a value m . Since the y_{ij} 's are normally distributed, it follows from the previous two equations that

$$\bar{y}_{i1M} \sim N\left(\mu_i, \frac{\sigma_{\eta i}^2}{M}\right) \quad (11)$$

and

$$\bar{y}_{i2M} \sim N\left(\mu_i + \delta_i, \frac{\sigma_{\eta i}^2}{n \& M}\right) \quad (12)$$

Furthermore, with $m = M$, $d_{im} \sim N\left(\delta_i, \sigma_{d_{im}}^2\right)$, where under the assumption that observations are independent in time,

$$\sigma_{d_{im}}^2 = \frac{\sigma_{\eta i}^2}{m} + \frac{\sigma_{\eta i}^2}{n \& m} \quad (13)$$

The hypotheses to be tested are $H_0: M = n$ (no bias) versus $H_a: M < n$ (bias).

Now, let $T_{im} = d_{im} / \sigma_{d_{im}}$. Note that T_{im} has a standard normal distribution under H_0 , i.e., $T_{im} \sim N(0, 1)$ for all m . In the implementation of the proposed likelihood ratio testing method, m is varied from 1 through $(n - 1)$, each time computing the value of T_{im} . The decision rule is based on

$$T_{i,max} = \text{maximum}^* T_{im}^*, \quad 1 \# m \# n \& 1 \quad (14)$$

Since $T_{i,max}$ is the maximum absolute value of a number of (correlated) standard normal variables

it does not itself follow a standard normal distribution. However, for a single variable i , the critical value of $T_{i,max}$ can be easily determined from simulation for a desired Type I Error probability and for a given sample size (n).

To summarize, the likelihood ratio test for a mean shift (i.e., for identifying a biased measured variable) is based on the following sequence of steps for a variable i :

1. Choose a value for the time of occurrence (TOC) of the bias, m , from the range $\{1, \dots, (n - 1)\}$ starting with $m = 1$.
2. Split the set of observations into two groups ($\{1, \dots, m\}$ and $\{m + 1, \dots, n\}$).
3. Compute the MLE for the mean of y_i in each group (i.e., the group sample average).
4. Compute the difference (d_{im}) of the sample averages for the two groups and standardize to obtain T_{im} .
5. Vary m from 1 to $m - 1$ each time going from step 1 to step 4.
6. Determine the maximum of T_{im} over all $m = 1, \dots, (n - 1)$ and denote this maximum by $T_{i,max}$.
7. Compare the observed value of $T_{i,max}$ to an appropriate percentage point of the null distribution of $T_{i,max}$. If the observed value exceeds this small upper percentage point the conclusion is that there is (a change in) bias in the variable under consideration.

In the analysis of a process network where there are several variables that are potentially biased, one then needs to repeat steps 1 through 7 for each suspect variable.

4. A Bayesian Approach To Testing For A Mean Shift

In this approach, one assigns prior distributions to the unknown parameters, μ_i , $\mu_i + \delta_i$, and M defined in the previous section. In a chemical plant, process history and the expertise of process engineers could be used in choosing the priors. (It may be noted here that as more and more data become available, i.e., as the sample size increases, the particular choice made for a prior becomes less important.). Once the priors are selected, and data are “in hand,” one needs to get the conditional distribution of a parameter of interest *given* the data. This conditional distribution then becomes the basis of inference for the parameter of interest.

Based on Eq. 8, let the marginal density of y_{ij} be denoted by $f(y_{ij} | \mu_i, \sigma_{\eta i}^2)$. Then, for a sample size n , with observations being independent in time, the joint distribution of $\{y_{i1}, \dots, y_{in}\}$ is given by

$$f(y_{i1}, \dots, y_{in} | \mu_i, \mu_i + \delta_i, m, \sigma_{\eta i}^2) = \prod_{j=1}^m f(y_{ij} | \mu_i, m, \sigma_{\eta i}^2) @ \prod_{j=m+1}^n f(y_{ij} | \mu_i + \delta_i, m, \sigma_{\eta i}^2) \quad (15)$$

where the parameters δ_i and m account for the change in the mean of y_i due to the initiation of a bias (δ_i) at some point of time $M + 1$ for $M \in \{1, \dots, n\}$.

For this study the following (independent) prior distributions were chosen for μ_i and $\mu_i + \delta_i$:

$$\mu_i \sim N(0, \tau^2) \quad (16)$$

$$\mu_i + \delta_i \sim N(0, \tau^2) \quad (17)$$

where τ^2 is a known (input) parameter. (However, the actual value of τ^2 is not critical to this analysis. For simplicity, the analysis and results of this work used $\tau^2 = 1$.) Thus, the distributions of μ_i and $\mu_i + \delta_i$ have the densities:

$$f(\mu_i^* 0, \tau^2) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{\mu_i^2}{2\tau^2}\right] \quad (18)$$

$$f(\mu_i^*) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{\mu_i^2}{2\tau^2}\right]. \quad (19)$$

The following probability mass function was used for M (We use M to denote the random variable and m to denote a value for the variable):

$$g(m) = \begin{cases} \frac{1}{n} & m = 1, 2, \dots, n-1 \\ k & m = n. \end{cases} \quad (20)$$

(Ultimately the value of the weight, k , is varied to get a desired test level, i.e., a desired probability of false identification when there are no biases.) With these (independent) priors the joint distribution of $\{y_{i1}, \dots, y_{in}, \mu_i, \mu_i^* | m\}$ is specified by

$$f(y_{i1}, \dots, y_{in}, \mu_i, \mu_i^* | m) = \prod_{j=1}^{m-1} f(y_{ij} | \mu_i, \sigma_{\eta i}^2) \prod_{j=M}^n f(y_i^* | \mu_i, \sigma_{\eta}^2) \quad (21)$$

$$\otimes f_1(\mu_i^* 0, \tau^2) \otimes f_2(\mu_i^*, \tau^2) \otimes g(m).$$

Integrating this expression successively with respect to μ_i and μ_i^* with limits $-\infty$ to $+\infty$ results in a function of m proportional to the conditional probability mass function for M given the data $\{y_{i1}, \dots, y_{in}\}$. The details of this derivation are given in Appendix A. Denoting the conditional distribution of M as $h(m | y_{i1}, \dots, y_{in})$, the decision rule for detecting a bias in variable i is:

- C Compute the value of $h(m | data)$ varying M from 1 to n .
- C Infer $M < n$, i.e., a bias has occurred, if a value of m maximizing $h(m | data)$ (i.e., the mode of the conditional distribution) $< n$.

In order to compare the performance of this Bayesian method with that the other methods the following hypotheses, based on the parameter M , were used in making an inference on the presence of the bias:

$H_0: M = n$, or equivalently, H_0 : (no change) in bias

$H_a: 1 \neq M < n$. or equivalently, H_a : a (change in) bias has occurred

In summary, the proposed Bayesian method for the identification of measurement biases in process variable follows these general lines:

- C Determine the distribution (joint density of observations) of the process variable given the parameters (known and unknown).
- C Select appropriate priors for all the unknown parameters based on process knowledge or historical data.
- C Determine the conditional density of the parameter of interest, M , given the process data and the other parameters.
- C Follow a decision rule for identifying a bias based on an analysis of the conditional distribution.

5. Simulation Study and Results

In this section the performance of the two methods is evaluated based on simulated process data and compared to the performance of the Imbalance Correlation Strategy (ICS) presented by Devanathan et al.¹ Data for the study were generated using a FORTRAN-NAG (Numerical Algorithm Group) subroutine for random number generation. For ease of comparison across the techniques, the same performance measures were used as those employed in Devanathan et al. The definitions of these measures are reproduced below, for convenience.

The first part of this section presents a comparison of the three techniques (the Likelihood Ratio Test (LRT) , the Bayesian Method (BM), and ICS) for the case where a single variable is considered. In this part the factors that are varied are M (time of initiation of the bias), δ_i (the magnitude of the bias), and n (the sample size). The second part of this section consists of an analysis done for the situation of two biases in the network used as an example in Devanathan et al.

As in the studies conducted for the ICS (Devanathan et al.), it is assumed that observations are available on each process variable. In addition, the following conditions are true for the simulation study:

- C The process is in pseudo steady-state.
- C When the expected value of a measured variable undergoes a shift, the shift is due to the occurrence of a bias.
- C $\sigma_{\eta_i}^2 = 1, i = 1, \dots, p$ where p is the number of process variables.
- C Each result is obtained from 10,000 simulated cases.
- C Once a bias occurs it stays constant.

Following Devanathan et al., two performance measures are used in evaluating the proposed methods. The first one, power (P_i), represents the method's ability to correctly identify a particular biased variable i , and is given by

$$P_i = \frac{\text{number of nonzero } \delta_i\text{'s correctly identified}}{\text{number of nonzero } \delta_i\text{'s simulated}} \quad (22)$$

The second measure gives the probability of falsely concluding that a variable is biased and is called the Average Type I error. The *AVTI* is defined as

$$AVTI = \frac{\text{number of zero } \delta_i\text{'s wrongly identified}}{\text{total number of simulations}} \quad (23)$$

As mentioned earlier, the study involved observing the effect of several parameters on each method's ability to correctly identify biases. In addition to the bias magnitude, M , and n , the overall test level (denoted by α) is also varied (0.05 and 0.01). To ensure fairness in the comparisons, results for the three methods are always compared at the same level of α for both single variable and network comparisons. For the ICS and LRT, critical values were chosen for the maximum absolute sample correlation and $T_{i,max}$, respectively, so that only 500 or 100 of those data sets (i.e., simulation cases) produced values above the critical values. For the BM k was chosen so that only 500 or 100 of the data sets had posterior modes for M less than n . The network used for this study is the same as that in Devanathan et al.¹ (which was originally given in Narasimhan and Mah⁹) and is shown in Figure 1.

Results for a single variable, namely, the mass flow rate for stream number 1 in Figure 1, are presented in Table 1. In this situation, exactly one correlation is computed for the ICS, exactly one LRT performed, and one BM decision made. The critical value in each case is based on performing just one test and **not multiple tests for the whole network**. The critical value for each test is given in Table 1. This table shows the effect of varying δ_i (3.0, 4.0, and 5.0) on the power. A value of $\delta_i = 3.0$ means that the magnitude of the bias is three (3) times the standard deviation of the variable. It can be seen that, for $M = 20$, the LRT and the BM, power is very high (0.9992) even for $\delta_i = 3.0$. Table 1 shows that, for $M = 20$, the ICS, the power is low for $\delta_i = 3.0$, but as δ_i increases, power is comparable to that of the LRT and BM. Figure 2 gives the

LRT power function as δ varies from 0 to 3 for this case. This curve was developed from simulated cases from eleven (11) values of δ from 0 to 3. The BM should give a similar power curve for these conditions. As expected, as δ increases, the power increases, and this curve has the shape that one would expect. For specific conditions, one can develop this power curve from simulations under the model given in Section 2.

To observe the effect of M , three values of M (3, 15 and 25) for a fixed n (equal to 30) were used. Table 1 shows that power decreases when either M or $n - M$ decreases. The reason for the decrease in power when M is small is that \bar{y}_1 will not be as precise as an estimate of μ_i as when M has a higher value (say 15). Similarly when $n - M$ is low, \bar{y}_2 will not be as precise as an estimate of $\mu_i + \delta_i$ as when $n - M$ has a higher value (say 15). Once again, note that the ICS has lower power. This is not surprising since this table shows that power is low for $\delta_i = 3$. In summary, for the case testing for a single nonzero δ_i , the performances of the LRT method and the Bayesian method are comparable, while the ICS consistently has lower power for small values of δ_i .

Finally, the two new methods presented here are compared with the ICS on the complete seven (7) stream process for various combinations of two biases. The conditions used in this study are identical to those presented in the study of Devanathan et al. Table 2 shows that both the LRT and the BM have very high power (approximately 1.0) for all the cases. Table 3 shows that the ICS consistently has lower AVTI than either of the other two methods. However, the AVTI for the LRT and the BM is still very low (around 0.0167).

6. Concluding Remarks

This article presented two new techniques that can accurately detect mean shifts in process variables and are thereby capable of accurate identification of measurement biases for pseudo steady-state processes. These techniques are applicable for both linear and nonlinear process constraints. A further advantage is that the presence of physical leaks does not confound the identification issue, while it might do so in the case of nodal strategies. These advantages are possible because each process variable is tested individually for biases, and the nodal constraints are not used at all. For large process networks, the two methods should ideally be used after a nodal strategy, such as the Linear Combinations Technique (Rollins et al.²), has narrowed down the list of suspect variables, and further analysis is unable to complete the identification. Such a combination of a nodal strategy and tests on individual variables would minimize the analytical and computational requirements.

An important and widely prevalent problem in the chemical process industry is the identification of biases when constraints are bilinear or nonlinear. Process constraints are said to be bilinear when the components of conservation equations are products of two measured variables, such as mass flow rate and temperature in an enthalpy balance. Similarly, when the components of constraint equations are measured variables raised to powers other than one (as might be seen in chemical reactions with high order kinetics), the underlying process is said to have nonlinear constraints. Since the methods proposed in this work do not use physical constraints, they are not subject to the complexity of treating non-linear constraints.

Notation

AVTI = average Type I error given by Eq. 23

BM = Bayesian method

d_{im} = difference in sample averages, $\bar{y}_{i2m} - \bar{y}_{i1m}$

$f(x)$ = the marginal probability distribution function for x

k = adjustable parameter in Eq. 20 to give the desired test level

LRT = likelihood ratio test

m = an actual value of M

$M+1$ = the true time of occurrence of bias

n = total number of observations

p = number of process variables

P_i = the probability of a correct identification of bias for variable i given by Eq. 22

$T_{i,max}$ = test statistic given by Eq. 14

y_{ij} = measured value of variable i at the j th time instant

\bar{y}_{i1m} = sample average for variable i given by Eq. 11

\bar{y}_{i2m} = sample average for variable i given by Eq. 12

δ_i = measurement bias of variable i

ε_{ij} = random error of variable i at the j th time instant

η_{ij} = total random deviation for variable i at time instant j

λ_{ij} = true value of process deviation of variable i at the j th time instant

σ^2 = variance

μ_i = steady-state true value of variable i before the occurrence of bias

$\mu_i^)$ = steady-state true value of variable i after the occurrence of bias

τ^2 = variance assumed for μ_i and $\mu_i + \delta_i$ in priors

Literature Cited

1. [Devanathan, S., D. K. Rollins and S.B. Vardeman, \(2000\) "A New Approach for Improved Identification of Systematic Measurement Errors," *Computers and Chemical Engineering*, **24**\(12\), 2755.](#)
2. [Rollins, D. K., Y. Cheng and S. Devanathan, \(1996\) "Intelligent Selection of Hypothesis Tests to Enhance Gross Error Identification," *Comp. & Chem. Eng.*, **20**\(5\), 517.](#)
3. Bagajewicz, M. and D. K. Rollins, (2002) "Data Reconciliation," Section 5.2 in the *Instrument Engineers' Handbook, 3rd Edition, Volume 3: "Digital Systems,"* 691.
4. [Ripps, D. L., \(1965\) "Adjustment of Experimental Data," *Chem. Eng. Prog. Symp. Ser.*, **61**, 8.](#)
5. [Nogita, S. \(1972\) "Statistical Test and Adjustment of Process Data," *Ind. Eng. Chem. Process Des.*, **2** 197.](#)
6. [Romagnoli J. and G. Stephanopoulos, \(1980\) "On the Rectification of Measurement Errors for Complex Chemical Plants," *Chem. Eng. Sci.*, **35**\(5\), 1067.](#)
7. [Rosenberg, J., R. S. H. Mah and C. Iordache, \(1987\) "Evaluation of Schemes for Detecting and Identifying Gross Errors in Process Data," *Ind. Eng. Chem. Res.*, **26**, 55.](#)
8. [Rollins, D. K., and J. F. Davis, \(1992\) "Unbiased Estimation of Gross Errors in Process Measurements." *AIChE J.*, **38**\(4\), 563.](#)
9. [Narasimhan, S., and R. S. H. Mah, \(1987\) "Generalized Likelihood Ratios for Gross Error](#)

- Identification.” *AIChE J.*, **33**(9), 1514.
10. Keller, J.Y., M. Darouach, and G. Krzakala, (1994), “Multiple Biases or Process Leaks in Linear Steady-State Systems, *Computers & Chemical Engineering Journal*, **18**(10), 1001.
11. Mah, R. S. H. and A. C. Tamhane, (1988) “Detection of Gross Errors in Process Data,” *AIChE J.*, **34**, 541.
12. Crowe, C. M., (1986) “Reconciliation of Process Flow Rates by Matrix Projection. Part II: The Nonlinear Case,” *AIChE J.*, **32**(4), 616.
13. Rollins, D. K., and S. D. Roelfs, (1992) "Gross Error Detection When Constraints Are Bilinear," *AIChE J.*, **38**, 1295.
14. Rollins, D. K., and J. F. Davis, (1993) “Gross Error Detection When Variance-Covariance Matrices are Unknown” *AIChE J.*, **39**(8), 1335.

Appendix

This section gives a derivation of the conditional distribution of the change point parameter M (the number of unbiased measurements) for use in the Bayesian decision rule. For n observation on variable y_i , the joint distribution of $y_{i1}, y_{i2}, \dots, y_{in}$ is given by

$$f(y_{i1}, \dots, y_{in} | \mu_i, \mu_i^*, \sigma_\eta^2) \cdot \prod_{j=1}^m f(y_{ij} | \mu_i, \sigma_\eta^2) @ \prod_{j=m+1}^n f(y_{ij} | \mu_i^*, \sigma_\eta^2) \quad (\text{A.1})$$

With the independent priors specified earlier, the joint distribution of $\{y_{i1}, \dots, y_{in}, \mu_i, \mu_i^* | M\}$ is given by

$$f(y_{i1}, \dots, y_{in}, \mu_i, \mu_i^*, m) \cdot \prod_{j=1}^m f(y_{ij}^* | \mu_i, \sigma_\eta^2) @ \prod_{j=m+1}^n f(y_{ij}^* | \mu_i^*, \sigma_\eta^2) \\ @ f_1(\mu_i^* | 0, \tau^2) @ f_2(\mu_i^* | 0, \tau^2) @ g(m). \quad (\text{A.2})$$

Integrating this expression successively with respect to μ_i and μ_i with limits -4 to $+4$ results in a function of m proportional to the conditional probability mass function for M given the data $\{y_{i1}, \dots, y_{in}\}$. Using Eq. 10 and substituting for the marginal probability density functions (pdf's), Eq. (A.2) reduces to

$$f(y_{i1}, \dots, y_{in}, \mu_i, \mu_i, m) = c_1 \exp \left[\sum_{j=1}^m \frac{(y_{ij} - \mu_i)^2}{2\sigma^2} - \frac{\mu_i^2}{2\tau^2} - \sum_{j=m+1}^n \frac{(y_{ij} - \mu_i)^2}{2\sigma^2} - \frac{\mu_i^2}{2\tau^2} \right] \quad (\text{A.3})$$

where c_1 is a constant that can be excluded for the purposes of integration.

The integration can be done in two parts, first with respect to μ_i and then with respect to μ_i . The first term considered is

$$\begin{aligned} & \exp \left[\sum_{j=1}^m \frac{(y_{ij} - \mu_i)^2}{2\sigma^2} - \frac{\mu_i^2}{2\tau^2} \right], \\ & \exp \left[\frac{1}{2} \left(\sum_{j=1}^m \frac{(y_{ij}^2 - \mu_i^2 - 2\mu_i y_{ij})}{\sigma^2} - \frac{\mu_i^2 \sigma^2 / \tau^2}{\sigma^2} \right) \right], \\ & \exp \left[\frac{1}{2} \frac{\sum_{j=1}^m [y_{ij}^2 - \mu_i^2 - 2\mu_i y_{ij} - \mu_i^2 (\sigma^2 / m\tau^2)]}{\sigma^2} \right], \\ & \exp \left[\frac{1}{2\sigma^2} \sum_{j=1}^m \left(\mu_i^2 \left(1 - \frac{\sigma^2}{m\tau^2} \right) - 2\mu_i y_{ij} - y_{ij}^2 \right) \right], \\ & \exp \left[\frac{1}{2\sigma^2} \left[\mu_i^2 m b - 2\mu_i \sum_{j=1}^m y_{ij} - \sum_{j=1}^m y_{ij}^2 \right] \right] \end{aligned} \quad (\text{A.4})$$

where

$$b = 1 - \frac{\sigma^2}{m\tau^2}$$

With a some algebra, the last line in Eq. (A.4) can be rewritten as follows:

$$F_1 = \exp\left[\frac{1}{2\sigma^2} \left(\mu_i \sqrt{mb} + \sum_{j=1}^m y_{ij} (\sqrt{mb})^{\&1} \right)^2\right] \otimes \exp\left[\frac{1}{2\sigma^2} \left(\sum_{j=1}^m y_{ij}^2 + \left(\sum_{j=1}^m y_{ij} \right)^2 [mb]^{\&1} \right)\right]. \quad (\text{A.5})$$

Now, the integration of the first line of Eq. (A.4) is achieved more easily by integrating F_1 in Eq. (A.5). Denoting this integration term by I_1 , we have

$$I_1 = \int_{-\infty}^{\infty} F_1 d\mu_i = \int_{-\infty}^{\infty} \exp\left[\frac{1}{2\sigma^2} \left(\sum_{j=1}^m y_{ij}^2 + \left(\sum_{j=1}^m y_{ij} \right)^2 [mb]^{\&1} \right)\right] \otimes \exp\left[\frac{1}{2\sigma^2 [mb]^{\&1}} \left(\mu_i + [mb]^{\&1} \sum_{j=1}^m y_{ij} \right)^2\right] d\mu_i \quad (\text{A.6})$$

It can be seen that the integral term in Eq. (A.6) represents the area under a normal distribution with mean $[mb]^{\&1} \sum_{j=1}^m y_{ij}$, and variance $\sigma^2 [mb]^{\&1}$, when multiplied by the constant term

$$\frac{1}{\sqrt{2\pi [mb]^{\&1}} \sigma}$$

Thus, Eq. (A.6) reduces to

$$I_1 = \exp\left[\frac{1}{2\sigma^2} \left(\sum_{j=1}^m y_{ij}^2 + \left(\sum_{j=1}^m y_{ij} \right)^2 [mb]^{\&1} \right)\right] \otimes \sqrt{2\pi [mb]^{\&1}} \sigma \quad (\text{A.7})$$

where the constant b is defined earlier.

Now consider the second part in Eq. (A.3) which involves $\mu_i \setminus \setminus$ but not μ_i :

$$\begin{aligned}
& \exp \left[\sum_{j=m+1}^n \frac{(y_{ij} - \mu_i)^2}{2\sigma^2} + \frac{\mu_i^2}{2\tau^2} \right], \\
& \exp \left[\frac{1}{2} \left(\sum_{j=m+1}^n \frac{(y_{ij}^2 - \mu_i^2 + 2\mu_i y_{ij})}{\sigma^2} - \frac{\mu_i^2 \sigma^2 / \tau^2}{\sigma^2} \right) \right], \\
& \exp \left[\frac{1}{2} \frac{\sum_{j=m+1}^n [y_{ij}^2 - \mu_i^2 + 2\mu_i y_{ij} - \mu_i^2 \sigma^2 / ((n-m)\tau^2)]}{\sigma^2} \right], \tag{A.8} \\
& \exp \left[\frac{1}{2\sigma^2} \left(\mu_i^2 \left((n-m) - \frac{\sigma^2}{\tau^2} \right) - \sum_{j=m+1}^n y_{ij}^2 + 2\mu_i \sum_{j=m+1}^n y_{ij} \right) \right], \\
& \exp \left[\frac{1}{2\sigma^2} \left(\mu_i^2 c - \sum_{j=m+1}^n y_{ij}^2 + 2\mu_i \sum_{j=m+1}^n y_{ij} \right) \right]
\end{aligned}$$

where

$$c = (n-m) - \frac{\sigma^2}{\tau^2}$$

Using a few algebraic manipulations to express terms in the exponential in Eq. (A.8),

we get

$$\begin{aligned}
& \exp \left[\frac{1}{2\sigma^2} \left[(\mu_i \sqrt{c})^2 + 2 \left(\prod_{j=m+1}^n y_{ij} \right) c^{1/2} \mu_i + \prod_{j=m+1}^n y_{ij}^2 c \right] \right] @ \\
& \exp \left[\frac{1}{2\sigma^2} \left[\prod_{j=m+1}^n y_{ij}^2 + \left(\prod_{j=m+1}^n y_{ij} \right)^2 c \right] \right] \\
& \exp \left[\frac{1}{2\sigma^2} \left[\mu_i \sqrt{c} + \left(\prod_{j=m+1}^n y_{ij} \right) c^{1/2} \right]^2 \right] @ \\
& \exp \left[\frac{1}{2\sigma^2} \left[\prod_{j=m+1}^n y_{ij}^2 + \left(\prod_{j=m+1}^n y_{ij} \right)^2 c \right] \right]
\end{aligned} \tag{A.9}$$

Once again, the integration of Eq. (A.7) is achieved by integrating F_2 with respect to μ_i between limits of -4 and 4 once the integral is expressed in the form of a cumulative distribution function for a normal distribution.

$$\begin{aligned}
I_2 &= \int_{-4}^4 F_2 d\mu_i \exp \left[\frac{1}{2\sigma^2} \left(\prod_{j=m+1}^n y_{ij}^2 + \left(\prod_{j=m+1}^n y_{ij} \right)^2 c \right) \right] @ \\
&= \int_{-4}^4 \exp \left[\frac{1}{2\sigma^2 c} \left(\mu_i + \left(\prod_{j=m+1}^n y_{ij} \right) c^{1/2} \right)^2 \right] d\mu_i
\end{aligned} \tag{A.10}$$

Note that the integral term in Eq. (A.10) represents the area under a normal distribution with mean $c^{1/2} \prod_{j=m+1}^n y_{ij}$, and variance $c^{-1} \sigma^2$, when multiplied by the constant term

$$\frac{1}{\sqrt{2\pi c^{-1} \sigma^2}}.$$

Thus, Eq. (A.10) reduces to

$$I_2 = \sqrt{2\pi c^{&l}} \sigma \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{j=1}^n y_{ij}^2 + \left(\sum_{j=1}^n y_{ij}^2 \right) c^{&l} \right) \right] \quad (\text{A.11})$$

Using Eq. (A.7) and Eq. (A.11), the conditional distribution of M can now be determined and the Bayesian decision rule implemented.

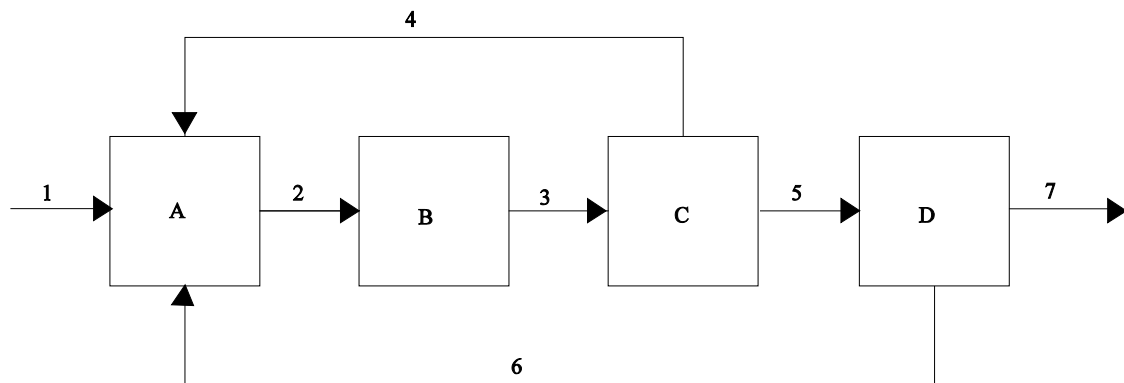


Figure 1. Process network used in this study (see Devanathan et al.¹ and Narasimhan and Mah⁹).

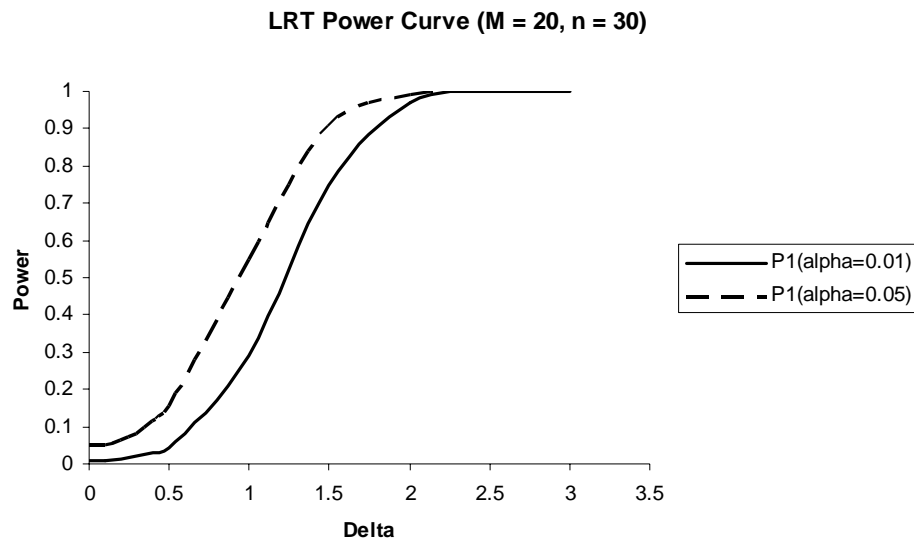


Figure 2. LRT power curve for $M = 20$, $n = 30$ for $\alpha = 0.01$ and $\alpha = 0.05$ as a function of δ .

Table 1. LRT, BM and ICS results ($n = 30$).

Method	M = 20			$\delta_1 = 3$			
	α	δ_1	P_1	M	α	P_1	
LRT $T_{critical,0.05}$ = 2.69 $T_{critical,0.01}$ = 3.23	0.05	3	0.9992	3	0.05	0.9936	
	0.05	4	1.0000	15	0.05	1.0000	
	0.05	5	1.0000	25	0.05	0.9998	
	0.01	3	0.9978	3	0.01	0.9764	
	0.01	4	1.0000	15	0.01	1.0000	
	0.01	5	1.0000	25	0.01	0.9996	
	BM $T_{critical, 0.05}$ = 0.20 $T_{critical,0.01}$ = 0.41	0.05	3	1.0000	3	0.05	0.9912
		0.05	4	1.0000	15	0.05	1.0000
		0.05	5	1.0000	25	0.05	1.0000
0.01		3	1.0000	3	0.01	0.9822	
0.01		4	1.0000	15	0.01	1.0000	
0.01		5	1.0000	25	0.01	0.9998	

Table 2. Comparison of results. $n = 30$, $\alpha = 0.05$, $\delta_i = 7$, $\delta_j = 4$.

i	j	P_i	P_j for ICS	P_j for LRT	P_j for BM
1	6	1.00	1.00	1.00	1.00
1	7	1.00	0.81	1.00	1.00
6	7	1.00	1.00	1.00	1.00
2	3	1.00	0.75	1.00	1.00
2	4	1.00	0.75	1.00	1.00
3	4	1.00	0.63	1.00	1.00
4	5	1.00	1.00	1.00	1.00
4	6	1.00	1.00	1.00	1.00
5	6	1.00	0.62	1.00	1.00

Table 3. Comparison of results. $n = 30$, $\alpha = 0.05$, $\delta_i = 7$, $\delta_j = 4$.

i	j	AVTI for ICS	AVTI for LRT	AVTI for BM
1	6	0.0005	0.0173	0.0156
1	7	0.0002	0.0175	0.0167
6	7	0.0002	0.0167	0.0159
2	3	0.0003	0.0163	0.0172
2	4	0.0001	0.0165	0.0162
3	4	0.0006	0.0179	0.0165
4	5	0.0001	0.0175	0.0158
4	6	0.0002	0.0167	0.0166
5	6	0.0004	0.0179	0.0177