E. B. Dahlin

On-Line Identification of Process Dynamics

Abstract: Two methods similar to regression analysis are applied to the estimation of parameters in a dynamic process equation. One method uses a residual based directly on the differential equation for the process model. The second method forms the residual from the integro-differential equations derived by integration of the original differential equation with respect to time. The two methods are shown to be complementary in their sensitivity to process and measurement disturbances as well as to errors in the estimate of the process variable reference values. Certain parameters that are very sensitive with one method are shown to be much less sensitive with the other method. A combined method is developed which utilizes each one of the constituent techniques to estimate the parameter for which it has the highest accuracy. This not only permits identification with higher overall parameter accuracy, but also under many practical circumstances gives a convergent solution when one of the constituent methods would give a diverging solution having no practical value for updating control coefficients in an adaptive controller. It is shown that the estimate of the magnitude of a pole in a transfer function can be significantly improved by prefiltering the process input and output data with the same lowpass filter.

The paper presents a theoretical and experimental evaluation of the identification methods. Formulas are derived for the variances of the parameters permitting an estimation of the parameter accuracy in a particular test. The test data used has been collected from different control loops on paper machines. The disturbance level on some of the variables is very strong and a test signal-to-noise ratio as low as 0.5 can be encountered. The method is currently used in routine operation in adaptive paper machine control. It is in use on four control loops and has been tested on several other loops as well.

Introduction

This paper is concerned with methods for on-line identification of industrial processes; that is, it discusses methods of establishing and updating a dynamic model of a process during normal operation of that process. The identification method proposed in this paper was developed specifically for on-line updating of the control-law parameters in a digital controller used for moisture and basis-weight regulation in paper-making machines. The paper-making process is typical of those processes for which on-line identification is appropriate in that:

1. The process has time-varying dynamics. The variation in dynamics is caused by such factors as changes in wire-drainage characteristics, degradation of dryer felt, ambient air properties, and change of weight and composition from one paper grade to another.
2. The accuracy with which control parameters are selected for moisture and basis weight has a strong influence on operating economy.
3. Manual selection of control parameters by trial and error methods is not practical because of the normal strong interaction between moisture and basis-weight loops and because of the high disturbance level.

The length of computation time is a prime consideration in the selection of an on-line identification method. For any method this time depends upon the number and type of dynamic parameters that need estimation. With some methods certain parameters can be estimated by the relatively rapid solution of linear equations, while other parameters require the solution of non-linear equations. For example, with the least-squares criteria, estimation of a pole and process gain can be accomplished by the solution of two linear equations, whereas estimation of transport delay requires a time-consuming search procedure. As a more drastic example, if one is using a stochastic process model (which simultaneously relates both the manipulated and stochastic input signals to the output), none of the parameters can be estimated by the solution of linear equations in the least-squares criteria unless the disturbance signals are stochastically independent and represent white random processes. Some criteria in the maximum probability methods have the same effects on computation time as those just mentioned.[1]

Clearly, then, one must design the model structure for a particular application with an eye on computation time. Yet the model structure will also depend to a considerable extent on the control law to be used. A control law selected for its ability to handle a variety of different spectral properties of noise will not need the identification of all the parameters in a stochastic process model. The opposite is true if the control law is designed for optimal control with accurately known disturbance properties. Without further elaboration, one can conclude that the selection of an on-line identification method and a control law are highly interrelated and require an understanding of control loop sensitivity and the range and frequency of variation of the different process parameters. A general discussion of the needs for parameter identification and the needs for adaptive control is found in Ref. 2.

The techniques presented in Refs. 1, 3, 4, 5 all have been applied to process control problems. They share the common characteristic that parameters describing the process disturbances or their effects on the output are identified simultaneously with the parameters relating manipulated inputs to the process output. This is a larger task than that attempted by the method in this paper where only the parameters of the latter category are identified. The relative merits of one approach versus another can again only be judged by considering the identification method in conjunction with a control law. Reference 6 discusses general concepts of the design of a digital controller/process identifier pair. Åström, in this issue, presents a case history of the design of a controller and identifier for the paper-making process.[1]

Experimental work on process identification by frequency-domain methods has been presented, for example, in Refs. 7 and 8. It appears that these methods require much longer periods of data collection than was felt desirable for the present application. Hence, several well-known time-domain identification methods were examined for their applicability to on-line use in a paper machine. These methods utilize maximum probability criteria[3,4] as well as least-squares error criteria, in various formulations, as used in conventional regression analysis. When applied to a process model with only deterministic process variables, the latter methods are particularly attractive from a computational point of view. However, when applied to typical paper machine data that contained strong process noise, the most common of the least-squares criteria (described in Refs. 9–13) was found to produce large estimation errors. In fact, for certain process data, such as those collected from a particular paper machine dryer, the least-squares criteria consistently failed to provide even the correct polarity of the dominant pole in the transfer function. For other processes, such as a particular fourdrinier system (basis weight/stock flow dynamics), the method worked quite well, but only if one selected proper prefiltering of the data. Since filter characteristics have a critical effect on the estimated value of the pole, one is faced with the additional problem of filter selection.

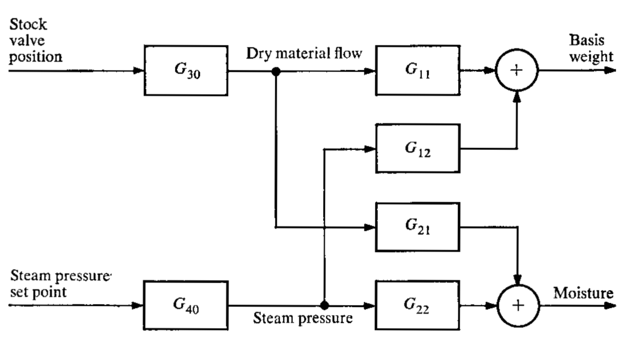
Evaluation of the conventional least-squares method as a potential technique for the paper-making process identification led to an investigation of a different least-squares method in which observation errors or residuals are formed from a time-integrated version of the original differential equation of the process. This method is equivalent to a special form of filtering (namely, a pure integrator) but has the unique property that initial conditions never cease to influence the filtered data. In estimating the pole of a typical transfer function used to characterize a portion of the paper machine, the method was found to be excellent, no matter what the filter characteristics and even if strong process disturbances were present. The method thus seems to offer the means for determining a parameter that cannot always be determined accurately by the conventional method. Since the two methods of forming residuals have other valuable complementary properties, a computation algorithm was developed which uses both methods in combination with computer-selected filters. This algorithm is referred to as the hybrid method.

Later in the paper some typical results obtained by the conventional least-squares method, by the integration method and by the hybrid method are presented. Since literature on identification focuses on the asymptotic properties of the least-squares or maximum probability estimators, some formulas will be presented for the parameter variance with finite observation time. Such formulas are of considerable practical interest because in many applications with moderate noise level one can successfully identify dynamic parameters using a short observation time and a single step function as process excitation. The derivations of these formulas are presented in a development report.[14]

The process model

In typical applications of on-line identification, processes having multiple inputs and multiple outputs are encountered. In paper machine control, for example, one is particularly interested in the transfer functions relating the stock-valve signal to basis weight and to the moisture of the sheet leaving the dryer. Equally important are the transfer functions relating the set-point signal of a steam pressure controller to basis weight and moisture. While identification can be accomplished from a knowledge of these four transfer functions, it can be done with much higher accuracy if one determines the transfer functions relating the set-point signal to steam pressure and relating dry material flow to the stock-valve signal. In Fig. 1 these two transfer functions are labelled G_{40} and G_{30}, respectively. The output of G_{30} is related to basis weight through

Figure 1 Process model partitioned into single-input, single-output elements.



the transfer function G_{11} and to moisture through the function G_{21}; the output of G_{40} is related to basis weight through G_{12} and to moisture through G_{22}. Even though this method converts the orginal problem into a new one in which more transfer functions have to be analyzed, the new functions are of lower order and, hence, can be handled with more accuracy. An additional advantage is that the method efficiently reduces the effects of disturbances.

The individual transfer function to be used in the following analysis is of first order and has three unknown parameters: gain, pole, and transport delay. Many industrial processes can be divided into elements having transfer functions of this form, and the model is therefore of significant practical importance. The six transfer functions mentioned above for the paper machine are of this category.

The principles to be described are applicable also to other types of transfer functions. During identification experiments using a double-pole transfer function for a process without a transport delay, it was found preferable to use the transport delay term of the single-pole formulation as an approximate description of the largest pole (the one having the smallest time constant). This led to longer computation time, but it had the advantage over solutions based on a double-pole model of permitting identification with data from processes having high disturbance levels.

The single-pole transfer function of interest in the present analysis is given in sampled-data form by the equation

$$\frac{1}{T}(y_i - y_{i-1}) + A(y_{i-1} - y_r) + B(x_{i-\Gamma-1} - x_r) = 0, \tag{1}$$

where A is the pole, B is the gain, Γ is the number of sampling intervals equal to the process transport delay time τ, T is the length of the sampling interval, y_i and x_i are the i^{th} samples of the model output and input, respectively, and y_r and x_r are reference values for the output and input, respectively. These reference values are the steady-state values, which are, in general, not known, but are estimated by a procedure separate from that used to estimate the parameters A, B, and τ.

The same transfer function can also be modeled in a form obtained by summing the terms of Eq. (1) over the intervals from 1 to n:

$$y_n - y_0 + AT\sum_{i=1}^{n}(y_{i-1} - y_r) + BT\sum_{i=1}^{n}(x_{i-\Gamma-1} - x_r) = 0, \tag{2}$$

where y_n and y_0 are the final and initial values of the model output, respectively.

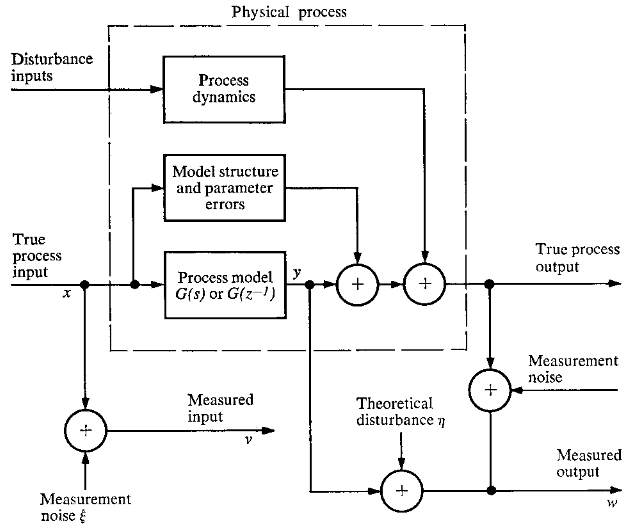


Figure 2 Process and model variables.

The formulations given by (1) and (2) have different properties and both are used in complementary fashion in the proposed identification scheme. It is worth noting here that both (1) and (2) are written in a form, which, at a high sampling rate, enables the parameters A, B, and τ to retain the same physical significance as they would in an equivalent continuous model of the system. This is convenient since the corresponding parameters in an actual continuous process have special physical significance.

Now, in order to use the models of (1) and (2) for on-line identification, it is necessary to relate the input and output values of the model to those values that can actually be observed in the process. Note that the model output y would have the same value as the measured output w if there were no errors in the model structure and parameter estimates and if no disturbance inputs and output measurement noise existed. Also, the model input x would have the same value as the measured input v if the input measurement noise were zero. Hence, by definition,

$$\begin{aligned} v &\equiv x + \xi \\ w &\equiv y + \eta, \end{aligned} \tag{3}$$

where ξ is the input measurement noise and η is a hypothetical disturbance signal assumed to account for all the factors that produce a difference between model output and observed output. The relation between the process model and the observed variables is illustrated in the structure shown in Fig. 2.

Since a theoretical justification of the model structure is not strictly necessary for process control applications, no details of the theoretical basis for the model are presented in this paper. However, the author has performed a detailed analysis of the dryer dynamics based on the physical principles involved, and the work in Ref. 15 gives a theoretical foundation for the basis weight dynamics. In each case there seems to be a good theoretical justification for the simple model structure used here. Nevertheless, the decisive engineering proof of the validity of an identification result is obviously to analyze the control system behavior when the controller is designed using that result. In other words, if the control system behaves as expected, the identification result is, by definition, satisfactory. In the author's opinion, the second best criterion for judging the validity of an identification result is to compare the actual process output to the model output when both process and model have the same input. This comparison can be done numerically and also graphically by superimposing plots of the two sets of output data. This paper relies primarily upon the latter form of presentation, allowing the reader to judge for himself the validity of an identification result.

The comparison of model and process output avoids a difficulty often encountered when using a residual test for validity analysis. The well-known necessary requirement for a valid identification result by residual methods is that the residual have essentially a zero mean and be essentially a white random process. The author's experience has been that in many cases these conditions may be met but the identification result is nevertheless useless for controller tuning purposes. (Figure 8g is an example of one such case encountered in the paper machine study.) A comparison of model and process output in each case eliminated an acceptance of such erroneous estimates.

Performance criteria used for identification

It can be concluded from Fig. 2 that it is not possible to distinguish between the individual effects of measurement noise, model parameter errors, model structure errors, and disturbance inputs by observation of v and w without making special assumptions or other measurements. Normally the input measurement noise ξ is the least significant of the various error sources in the parameter estimation. Therefore, if one assumes ξ negligible, the model output, y, can be computed from the observed input time series v. It is then meaningful (and extremely useful) in deciding between alternative model structures and parameter values to determine the model output error, ϵ, by comparing the model output with the observed output. This procedure is referred to as model output comparison, and is an essential feature of the identification method proposed in this paper.

The model output error is given at each sample time by the expression

$$\epsilon_i = w_i - y_i^*. \tag{4}$$

The term y_i^* is the model output, which is computed under the assumption of negligible input measurement noise by the expression

$$y_i^* = T \sum_{j=0}^{i} g_{i-j}(v_{j-\Gamma} - v_r) + (w_0 - w_r)e^{-iAT} + w_r, \tag{5}$$

where g_i is the model impulse response in sampled data form, w_r and v_r are estimated reference values for their respective series, and w_0 is the estimated initial value of the process output. Actually, the model output is more simply computed by solution of the difference equation corresponding to the convolution equation (5) than by solution of (5) itself. The difference equation is, of course, similar to (1).

In certain circumstances, to be discussed, it is advantageous to estimate parameters by using the criterion of minimum variance of model output error,

$$Y = \frac{1}{N-1} \sum_{i=1}^{N} \epsilon_i^2, \tag{6}$$

where N is the number of samples for each variable.

Criteria of a second type can be established by employing "residuals" derived from the process model equations (1) and (2). Residuals derived from Eq. (1) are used to formulate an estimation method to be called the "derivative method," (the name chosen because the residual has the same dimension as the time derivative of the process output), and residuals from Eq. (2) are the basis for an estimation method called the "integral method." The form of the residuals for both methods is given by

$$R_n = A\alpha_n + B\beta_n + \gamma_n, \tag{7}$$

where, for the derivative method

$$\alpha_n = w_{n-1} - w_r$$

$$\beta_n = v_{n-\Gamma-1} - v_r$$

$$\gamma_n = \frac{1}{T}(w_n - w_{n-1}), \tag{8}$$

and for the integral method

$$\alpha_n = T \sum_{i=1}^{n} (w_{i-1} - w_r)$$

$$\beta_n = T \sum_{i=1}^{n} (v_{i-\Gamma-1} - v_r)$$

$$\gamma_n = w_n - w_0. \tag{9}$$

Minimizing the variance, J, of the residual provides a criterion for parameter estimation; the variance is given by

$$J = \frac{1}{N-1} \sum_{n=1}^{N} R_n^2. \tag{10}$$

For any given τ estimates of A and B can be obtained by minimizing J with respect to A and B. Setting the partial derivatives $\partial J/\partial A$ and $\partial J/\partial B$ equal to zero leads to the familiar normal equation of least-squares analysis. Thus, using either the derivative or the integral method, A and B are computed from

$$\begin{bmatrix} A \\ B \end{bmatrix} = M^{\ddagger} \begin{bmatrix} -\sum_{n=1}^{N} \alpha_n \gamma_n T \\ -\sum_{n=1}^{N} \beta_n \gamma_n T \end{bmatrix}, \tag{11}$$

where $M\ddagger$ is the pseudo-inverse of the matrix[16]

$$M = \begin{bmatrix} \sum_{n=1}^{N} \alpha_n^2 T & \sum_{n=1}^{N} \alpha_n \beta_n T \\ \sum_{n=1}^{N} \alpha_n \beta_n T & \sum_{n=1}^{N} \beta_n^2 T \end{bmatrix}. \tag{12}$$

The matrix M is a so-called gram-matrix, which is invertible whenever α and β are linearly independent.[17] Since α refers to the process output and β to its input, this linear independence is assured in the absence of noise for any open-loop testing with any input perturbation if the system has a finite pole (which is true for almost any physical process). Experience has shown that a non-invertible normal matrix M is seldom encountered even in the presence of noise when open-loop data are used. Therefore the methods in the following are developed on the assumption that M^{-1} exists. Discussions concerning the inversion of M and its implication upon identification methods can be found in Ref. 18.

Since the derivative of J with respect to τ cannot be determined, τ is estimated by a search procedure rather than analytically. In general, the search procedure involves selecting a value for τ, calculating A and B from Eq. (11), and determining whether the three values are located at the global minimum of either J or Y. No problem with local minima has ever appeared in this search and it appears that the starting value for τ can be almost arbitrarily selected.

It will be shown in the experimental analysis that minimizing J with respect to τ yields a good estimate for τ only if A and B are calculated by the derivative method. However if the derivative method diverges or gives a $\tau \leqq 0$, the integral method must be used. In this case the accuracy of the estimate for τ is better if the model output variance, Y, rather than the residual variance, J, is minimized with respect to τ.

Accuracy of parameter estimation

Certain aspects of accuracy in parameter estimation with the residual performance criteria will now be analyzed. In particular, estimation accuracies obtained with the derivative and integral methods will be compared. The following analysis is limited to the case where the disturbance amplitudes and errors in the estimation of the reference levels for x and y are small compared with the process perturbations. The following inequalities are assumed valid for almost any i:

$$\begin{aligned} &\delta x_r \equiv x_r - v_r \ll \delta x_i \equiv x_{i-\Gamma-1} - x_r, \\ &\delta y_r \equiv y_r - w_r \ll \delta y_i \equiv y_{i-1} - y_r, \\ &\delta y_0 \equiv y_0 - w_0 \ll \delta y_i, \\ &\eta \ll \delta y_i, \quad \xi \ll \delta x_i, \quad e^{-AT(N-\Gamma)} \ll 1. \end{aligned} \tag{13}$$

Considering that the true values of A and B are known, one can derive expressions for the deviations from these values which will be caused by disturbances and reference errors. Defining parameter variation as δA and δB for A and B respectively, it can be shown[14] that

$$\begin{bmatrix} \delta A \\ \delta B \end{bmatrix} = M^{-1} \bar{a}, \tag{14}$$

where, for the derivative method,

$$\bar{a} = \begin{bmatrix} -T \sum_{i=1}^{N} \delta y_i [A(\eta_{i-1} + \delta y_r) \\ \quad + B(\xi_{i-\Gamma-1} + \delta x_r) + (\eta_i - \eta_{i-1})/T] \\ -T \sum_{i=1}^{N} \delta x_i [A(\eta_{i-1} + \delta y_r) \\ \quad + B(\xi_{i-\Gamma-1} + \delta x_r) + (\eta_i - \eta_{i-1})/T] \end{bmatrix} \tag{15}$$

and for the integral method,

$$\bar{a} = \begin{bmatrix} -T \sum_{i=1}^{N} \gamma_i \sum_{j=1}^{i} T\, \delta y_j \\ -T \sum_{i=1}^{N} \gamma_i \sum_{j=1}^{i} T\, \delta x_j \end{bmatrix}, \tag{16}$$

where

$$\gamma_i = A \sum_{j=1}^{i} T(\eta_j + \delta y_r) + B \sum_{j=1}^{i} T(\xi_{j-\Gamma} + \delta x_r) + \eta_i + \delta y_0. \tag{17}$$

From Eqs. (14)–(17) it can be seen that derivative and integral estimates of A and B are unbiased if the disturbances and reference errors have zero mean. Consider now the error sources δx_r and δy_r alone and assume that the noise sources are zero. Clearly, $\bar{a} = 0$ for the derivative method if

$$\sum_{i=1}^{N} \delta y_i = \sum_{i=1}^{N} \delta x_i = 0, \tag{18}$$

and for the integral method if

$$\sum_{i=1}^{N} i \sum_{j=1}^{i} \delta y_j = \sum_{i=1}^{N} i \sum_{j=1}^{i} \delta x_j = 0. \tag{19}$$

These results show that it is, in principle, simple to select a test signal wave form so that the parameter estimation is independent of errors in the reference values. If such a wave form is not used, the integral method is the more vulnerable to reference value errors. For a test signal consisting of a single step with the amplitude $\hat{x}$, one can compute M and $\bar{a}$ and obtain the following result for the derivative method:

$$\frac{\delta A}{A} = -\delta y_r \frac{2}{K\hat{x}} \frac{1}{1 + \dfrac{2}{A(V-\tau)}}, \tag{20}$$

and for the integral method:

$$\frac{\delta A}{A} = -\delta y_r (V-\tau) \frac{2}{K\hat{x}} \times \frac{1 - \dfrac{2\tau}{A(V-\tau)^2}}{1 + \dfrac{18}{A(V-\tau)} - \dfrac{12\tau^2}{A^2(V-\tau)^4}}, \tag{21}$$

where $V = NT$ is the length of the test interval, and $K = B/A$ is the steady-state gain.

Equation (21) shows that the relative error in the estimation of A increases almost linearly with the length of the test interval for the integral method, while Eq. (20) shows that the error is almost independent of test interval length for the derivative method.

It is also of great interest to determine the standard deviation for the estimation of A when no reference errors exist and, thus, the only error source is the disturbance η. This standard deviation has been computed by the author in Ref. 14 for a test signal consisting of a single step function. The results based on a single step have been found to reflect the interesting properties quite well, even when the identifier uses multiple steps with fairly large separation (as is typically done with the on-line method). When samples of the random variable η are uncorrelated, one can derive the following result for the derivative method:

$$\frac{\sigma_A}{A} = \frac{\sigma_\eta}{K\hat{x}} \sqrt{10AT} \left[\frac{1 + \dfrac{4}{3A(V-\tau)}}{1 + \dfrac{2}{A(V-\tau)}} \right]^{1/2} \tag{22}$$

and for the integral method:

$$\frac{\sigma_A}{A} = \frac{\sigma_\eta}{K\hat{x}} AT\sqrt{F_1 F_2} \times \frac{\sqrt{4.8(V-\tau)/T}}{1 + \dfrac{18}{A(V-\tau)} - \dfrac{12\tau^2}{A^2(V-\tau)^4}}, \tag{23}$$

where F_1 and F_2, which are near unity for large $(V-\tau)$, are defined by

$$F_1 = 1 + \frac{4\tau^2}{A^2(V-\tau)^4} - \frac{4\tau}{A(V-\tau)}, \tag{24}$$

$$F_2 = 1 + \frac{1.88}{A(V-\tau)} + \frac{2.5}{A^2(V-\tau)^4}, \tag{25}$$

and where σ_A and σ_η are the standard deviations for A and η, respectively.

It can be concluded from this theoretical analysis that if the sampling rate is chosen so that $AT \ll 1$, and if $V - \tau$ is not too large, the integral method will have higher accuracy than the derivative method in defining the pole A and the gain B. For typical paper machine data the difference in accuracy is quite marked; if, for example, $T = 1$, $V = 100$, $\tau = 0$, $A = 0.01$, $\sigma_\eta = 1$, and $K\hat{x} = 1$, the theoretical variance σ_A/A will be 0.303 for the derivative method and 0.028 for the integral method.

Data prefiltering

It was shown in the previous section that random disturbances can seriously affect the accuracy in estimating the pole, especially with the derivative method. Is it possible to prefilter the data in order to reduce this estimation error?

Suppose the time series, v and w, are both filtered with identical linear filters. Also assume the filtering to be such that all the effects of measurement noise and process disturbances are eliminated. One can easily see that the time series obtained by filtering w will be the true process output which would occur if the true process input were the filtered version of v. Thus the fact that the time series are distorted does not matter if a complete noise cancellation occurs.

Of course, we cannot meet this requirement exactly, but even by letting the filters make a partial cancellation of the noise, one can achieve a very significant improvement in the estimation, as will be shown with experimental data.

What form of filter is most suitable for this purpose? A useful filter can be derived heuristically without difficulty. The purpose of the filter is to improve the estimate of the parameter A. In order to improve the estimate of A as much as possible, the filter should improve the signal-to-noise ratio on the time interval where the data is sensitive to the magnitude of A. Since the typical excitation of the process is a series of step functions, consider the response of the model for a single step input. The data having maximum sensitivity to a variation of A can be shown to be collected at the time t_1 defined by:

$$t_1 = \tau + \frac{1}{A}. \tag{26}$$

The problem is now a familiar one. We want to select a filter function so that the maximum signal-to-noise ratio for a step input occurs at time t_1.

The amount of distortion on the signal waveform is immaterial. The solution to this problem is given by Davenport and Root,[19] on page 244. Considering the special case of white noise, the optimal filter impulse response is the so-called matched filter. The impulse response for this filter is the mirror image of the model output with the mirror in the plane, $t = t_1$, and is defined by the following function:

$$h(t) = K\hat{x}(1 - e^{-1}e^{At})[1 - u(1/A)], \tag{27}$$

where

$\hat{x}$ = amplitude of the process input step function (occurring at time, $t = 0$),

$u(t)$ = unit step function.

This filter is somewhat clumsy in practical use. Since it is of lowpass character, an exponential filter is a reasonable substitute and is simpler to apply. The exponential filter, which is equivalent to the matched filter, is chosen by matching the first and second moment of the impulse response functions. On this basis the best exponential filter is defined by the impulse response $h_1(t)$ given by

$$h_1(t) = \mathcal{H}_1 e^{-\mathcal{H}_2 t}, \tag{28}$$

where

$$\begin{aligned} \mathcal{H}_1 &= 0.915 K\hat{x} \\ \mathcal{H}_2 &= 2.49 A. \end{aligned} \tag{29}$$

Since the best filter requires prior knowledge of A an iteration method is applied to find the best value of $\mathcal{H}_2$ in the method described in the next Section. The $\mathcal{H}_2$ values found experimentally agree closely with Eq. (29) and the simple reasoning used here to justify the filter selection is well supported by practical results.

The hybrid identification method

Theoretical and experimental analysis of the derivative and integral methods conducted with numerous sets of data demonstrated that the methods have complementary properties. It was found furthermore that neither method alone had sufficient reliability and accuracy to handle the identification of all the transfer functions of main interest in paper machine control. It was, however, possible to combine the methods and use data filtering in such a manner that these transfer functions could be safely handled in routine on-line identification. This combination, to be called the hybrid method, will be briefly described in this Section. The properties of the derivative and integral methods that were the basis for the chosen structure of the hybrid method are summarized in Table 1.

Table 1 Comparison of several features in derivative and integral methods.

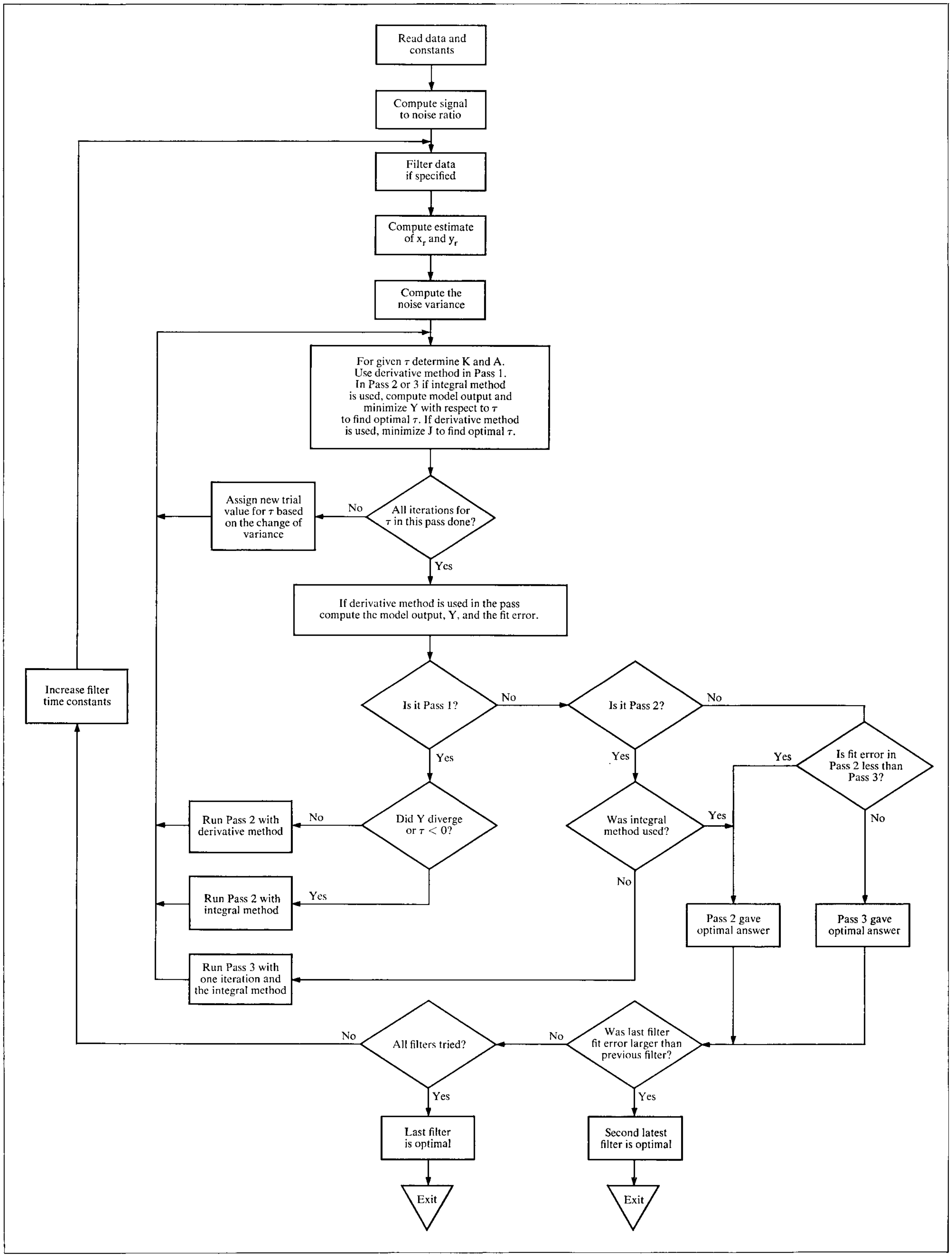
1. Divergence of solution	
DERIV:	Possible (Occurs for Series III experimental data).
INTEG:	Possible (Occurs for Series I experimental data).
2. Errors in estimation of A *and* B	
DERIV:	Quite high in the presence of disturbance. Independent of length of test pulses.
INTEG:	Small if the test transients are not too far apart and the pole is small. Errors increase with length of test pulses.
3. Errors in estimation of τ	
DERIV:	Very small, particularly if test pulses have fast rise time and the process time constant is small.
INTEG:	Large in the presence of disturbances.
4. Effect of errors in estimate of x_r *and* y_r	
DERIV:	Parameter errors are small and independent of length of experiment.
INTEG:	Parameter errors increase rapidly with length of experiment.
5. Effect of data filtering	
DERIV:	Prefiltering has excellant ability to improve estimates of A and B.
INTEG:	Prefiltering is of value.

The first property in Table 1 indicates that both methods may, under the influence of noise, provide an answer which is said to diverge. Divergence of a set of parameter values implies that the model output variance, Y, approaches infinity for large observation intervals. (Even when the values diverge, the residual variance, J, typically remains finite and of reasonable magnitude; see Figs. 8f and 8g, for example.) A divergent answer is clearly unusable for controller tuning purposes. It should be noted, though, that while divergence can occur with both methods, the experiments showed that the data which diverged with the derivative method were from a different process than the data which diverged with the integral method.

The second feature in Table 1 indicates that the estimation of A and B is normally done best with the integral method. However, the integral method was found experimentally to be inaccurate in the estimation of τ if this variable was chosen for minimum J. A strong improvement in the estimation of τ occurs if it is chosen for minimum Y when the integral method is used. The integral method requires more computation time, and it is therefore always preferable to try to acquire the value of τ by the derivative method. Table 1 also summarizes the sensitivity to errors in reference values and the benefits of filtering.

Figure 3 illustrates the basic logic structure in a computer program for the hybrid method. The first block in the

Figure 3 Hybrid identification method, basic flow chart.



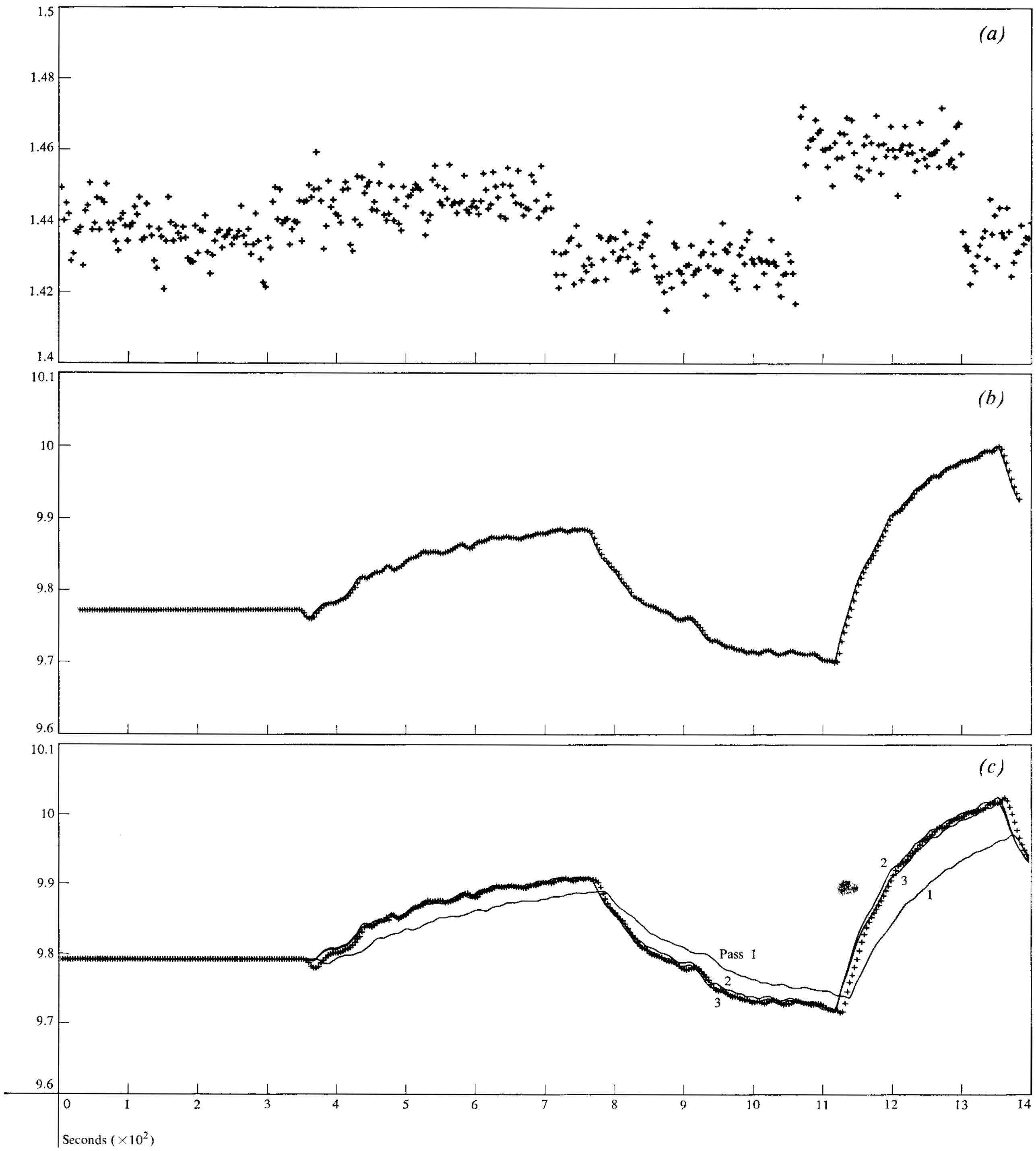


Figure 4 Test problem; (a) input signal, (b) solution by integral method, (c) solution by hybrid method. (Pass 2 is equivalent to solution by derivative method). In this, as well as subsequent figures, the crosses, +, refer to observed process output, while the smooth-line curves refer to computed output from the process model.

program refers to the reading of input data from the process. The data to be used in the program are collected during two different time periods, the relaxation period and the transient period. During the relaxation period the process is maintained in open-loop operation and the manipulated input variable, x, stays constant as near to the steady-state value as possible. (The controller that is typically used in conjunction with the identifier maintains and updates a good estimate of this value.) During the transient, which immediately follows the relaxation period, the value of x is varied.

After the input data have been read into the computer and converted to suitable units, the signal-to-noise ratio of the data collected during the transient period is computed according to the definition:

$$(S/N) = \sqrt{(\sigma^2_{wt} - \sigma^2_{wr})/\sigma^2_{wr}}, \tag{30}$$

where

σ^2_{wt} = the variance of the observed output, w, during the transient period, and

σ^2_{wr} = the variance of w during the relaxation period.

The signal-to-noise ratio is used to judge whether or not the test transient is of acceptable magnitude.

In the next block the program has entered a loop for iteration with respect to the filter constant $\mathcal{K}_2$. This parameter is varied in a monotonic sequence starting with a very large value or zero time constant (meaning there is no filtering) and changing to successively decreasing $\mathcal{K}_2$ (larger time constant). The filter chosen as best is the one that provides the smallest fit-error, E_f, defined by

$$E_f = (Y - \sigma^2_{wr})/\sigma^2_{wr}, \tag{31}$$

where σ^2_{wr} has the same meaning as above, but this time refers to the data from the relaxation period *after* it has been filtered.

Next the computer estimates reference values for x_r and y_r, which estimates have been defined as the observable variables, v_r and w_r. These estimates are based on a formula expressing the values at the end of the relaxation period of the best straight lines fitted to the input and output data, respectively. ("Best" is defined on a least-square error basis.) This method gives some compensation for drift occurring during the relaxation period.

After the noise variance is computed, the program enters an intermediate loop where the derivative and integral methods are compared as prospective means for solution, and an inner loop where the optimal value of τ is found by an iterative search procedure. Each iteration in the intermediate loop is called a "pass." In Pass 1, which has relatively few iterations with respect to τ in the inner loop, the derivative method is always used in order to determine whether it is a good prospect for solution. If Y does not diverge and if τ is not equal to or less than zero, the derivative method is used again in Pass 2. In Pass 2, there are more iterations than in Pass 1 with respect to τ, and an accurate solution for τ is obtained. In Pass 3 the integral method is used with the value of τ required during the previous pass. Finally, the fit errors computed during Pass 2 and Pass 3 are compared and the values of A and K (where $K = B/A$) that give the smaller fit error are selected as the best parameter set for the particular $\mathcal{K}_2$ value used in that outer loop iteration. However, if during Pass 1, Y diverges or τ is equal to or less than zero, the derivative method is immediately judged a poor prospect, and Pass 2 then uses the integral method with one or more iterations in the inner loop. In these iterations the value of Y is computed for each value of τ and the optimum τ is chosen as the one that minimizes Y.

Table 2 Identification results for test problem.

	Parameters K	A	τ
Runge-Fox parameters	4.05	0.0114	65.0
Derivative method	4.19	0.0094	56.3
Integral method	4.08	0.0113	60.0
Integral method with filter (40 sec time constant)	4.22	0.0083	80.0

Test problem using data not influenced by process disturbances or measurement noise

A special test problem having a known answer was created for evaluation of the identification methods. The observed process input shown in Fig. 4a was used for the computation of the model output, y^*_i, for a particular set fo model parameters. In subsequent experiments the input data used here were considered to be the observed process input and the computed model output was considered to be the observed process output. This test problem, then, has a typical process input waveform and an ideal, observed output time series, not influenced by process disturbances or measurement noise. The computation of y^*_i was done by a Runge-Fox[20] integration routine with the parameter values shown in the first row of Table 2.

When the derivative and integral methods were applied to this test problem, the parameter values shown in the second and third rows of Table 2 were determined. Slight parameter errors do exist, but they are probably due to the size of the sampling interval (3.49 seconds). Even so, Figs. 4b and 4c show that the identified parameters provide a model output that closely resembles the observed output. Note in Fig 4c that the output obtained using the results of Pass 2 of the hybrid method is the same as that obtained using the derivative method alone Although Pass 3 gives the optimal solution for this problem, the result is only slightly better than that obtained from Pass 2.

Equations (15)–(21) have shown that errors in the estimation of x_r and y_r will, for a large observation interval, cause much larger errors in the values of parameters estimated by the integral method than by the derivative method. This theoretical conclusion was checked experimentally on the test problem. Errors in the reference values, v_r and w_r, were purposely introduced and the

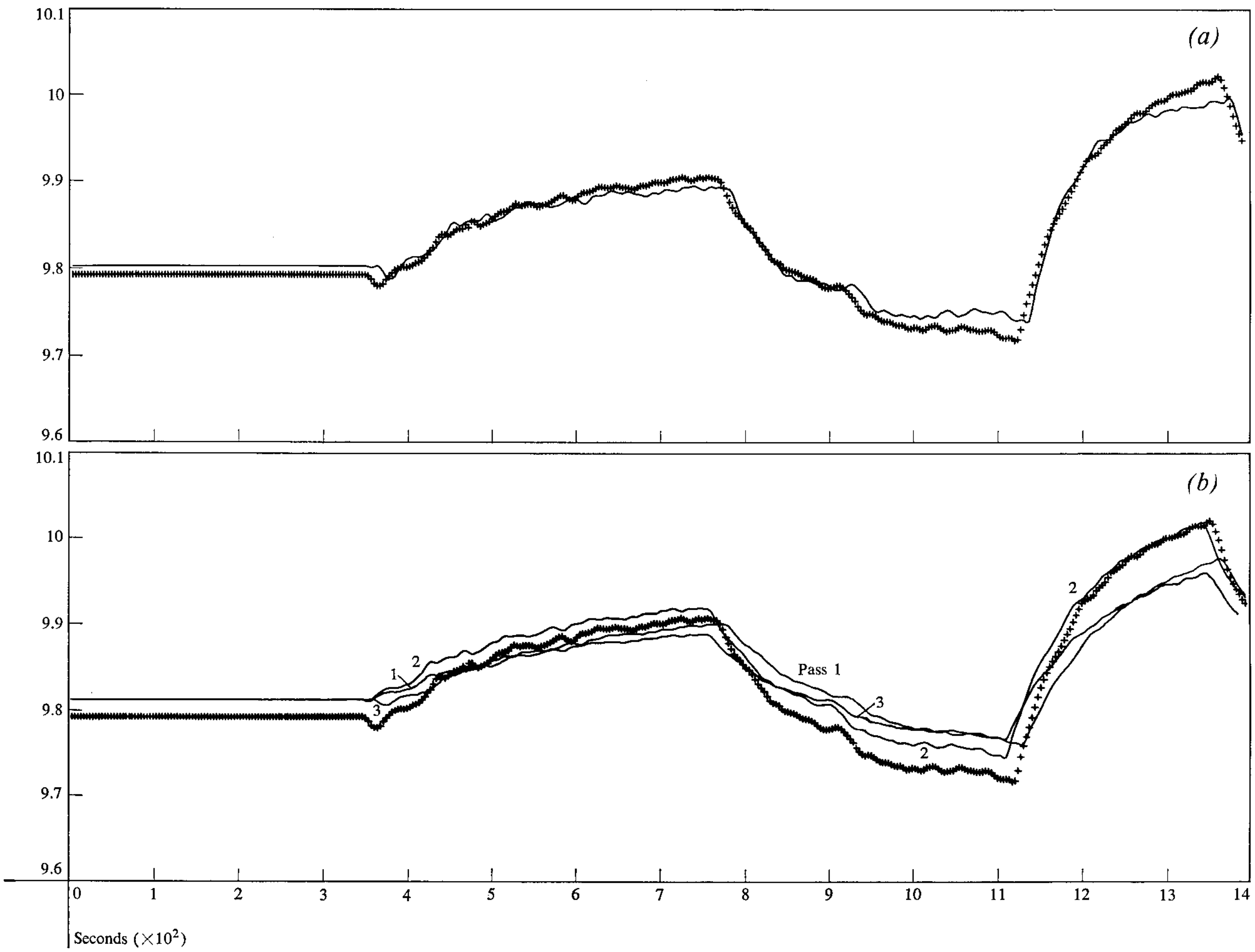


Figure 5 Effect of reference value errors on model output; (a) solution by integral method with reference error of 0.01 units, (b) solution by hybrid method with reference error of 0.02 units. (Pass 2 is equivalent to solution by derivative method.)

model output as well as the parameter values were determined as shown in Figs. 5 and 6. In Fig. 5a the integral method was used to compute the model output for a reference error, $w_r - y_r$, of 0.01 units. In Fig. 5b the reference error was 0.02 units and the model output was computed by the hybrid and derivative methods (the derivative method solution is identical with that of Pass 2 of the hybrid method). The Pass 2 solution is optimal and fits quite well. The variation of parameter values for a range of reference errors is shown in Fig. 6. Note that the derivative method provides a constant estimate for τ regardless of the magnitude of the reference error. The other parameters are also much less influenced by the reference error when estimated by the derivative method rather than the integral method. This test, as well as others, has confirmed that the derivative (and hybrid) method is far less sensitive than the integral method to reference errors. Using Eqs. (20) and (21) with numeric values corresponding to the experimental data of Fig. 6 (average $K\hat{x} = 0.65$, $A = 0.0114$, $V - \tau = 1000$, $\tau = 65$), one obtains the theoretically expected results shown as dashed lines in Fig. 6. The theoretical and experimental results for the relative error in A agree closely for the derivative method of estimation.

The test problem was also used to check the effect of data prefiltering. The fourth row in Table 2 shows that a relatively heavy filter did not affect the result very severely. However, because of sampling effects, the identification was not completely independent of the filter. A filter with a time constant (by definition, the inverse of the parameter $\mathcal{K}_2$) of 40 seconds was used for this and subsequent experiments.

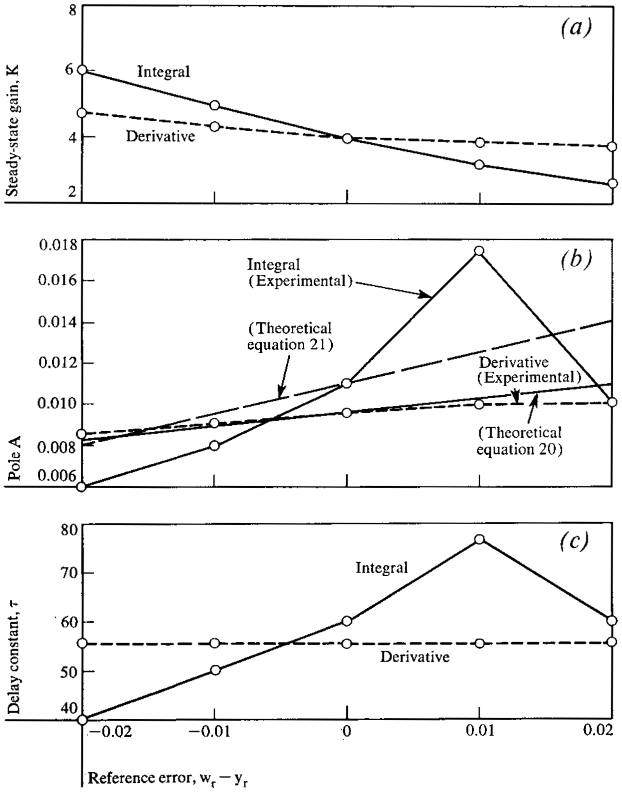


Figure 6 Effect of reference errors on the estimates of parameter values.

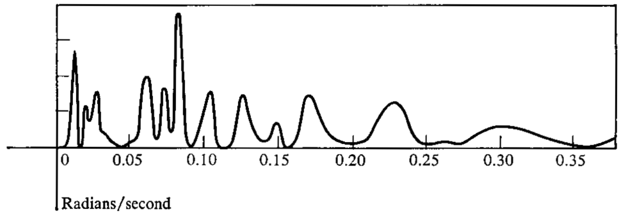


Figure 7 Power spectrum for Series I process output with constant manipulated input.

Experimental evaluation with process data

Typical results from the experimental evaluation of the identification methods with actual process data are presented in this Section. The methods have so far been applied to eight different transfer functions involving such variables as steam temperature, steam pressure, moisture, basis weight, stock flow, consistency, and stock valve position. The total number of on-line and off-line identifications made to date exceeds 100.

Some results of experiments with three series of data from different control loops will be discussed. Each series contains the data from several identification experiments, or runs. The series have purposely been chosen from highly disturbed control loops where the identification is difficult. For these data the results clearly show the differences between the derivative, integral, and hybrid techniques. The disturbances for all three series have a broad bandwidth, and the formulas derived under the assumption of white noise should be well applicable. A typical power spectrum is shown in Fig. 7.

Series I of the experiments has basis weight as the output variable and dry material flow as the input. Series II uses data taken from a different paper machine. It, also, has basis weight as the output variable, but uses the stock valve position signal as the input. Series III data are taken from a paper machine that was producing a grade of paper typically having very strong moisture disturbances; reel moisture is the output and dryer steam pressure is the input. The sampling time used in Series I is 3.49 seconds, and in Series II and III, 3.6 seconds.

Series I, basis-weight/dry-material-flow transfer function

Table 3 summarizes some of the results from the analysis of the Series I data. The signal-to-noise ratio varies significantly between the several runs. In Run 1 the test signal selected was very large. This led to a loss of head in the stock line, which gave the transient a strong droop, as can be seen in the output graphs of Fig. 8a. Run 2 has a normal excitation; the input signal is shown in Fig. 8b and the output in Figs. 8c and 8d. Run 3 uses a signal having only one pulse of relatively low amplitude; in the output graphs for this run notice that Pass 3 of the hybrid method diverged (Fig. 8e), as did the integral method. In Run 4, which has a very low signal-to-noise ratio (0.55), the integral method diverged again. (Fig. 8f). This is interesting because if acceptance of the identification were based on the residual being normal with zero mean, as it is in this case, (Fig. 8g), the model would give a response having almost no similarity to the observed process output. Divergence of the solution causes the identification by the integral method to be rejected. However, the derivative method did not diverge in Run 4 (Fig. 8h) although the accuracy was low when no filter was used (Table 3).

The values for K and τ obtained by the derivative method are in excellent agreement among Runs 1–3. In spite of the spread of values for A among the Runs, the model output curves for the derivative method (these curves are the same as the Pass 2 curves for the hybrid method) fit quite well to the observed output.

The model output shown in Fig. 8d was obtained with the integral method, where τ was estimated by minimizing the residual variance, J. Although the estimate for K was good, the values for A and τ were notably small. This result, coupled with the fact that the integral method

Figure 8 Series I experiments; (a) Run 1, hybrid solution, (b) Run 2, process input, (c) Run 2, integral solution, (d) Run 2, hybrid solution.

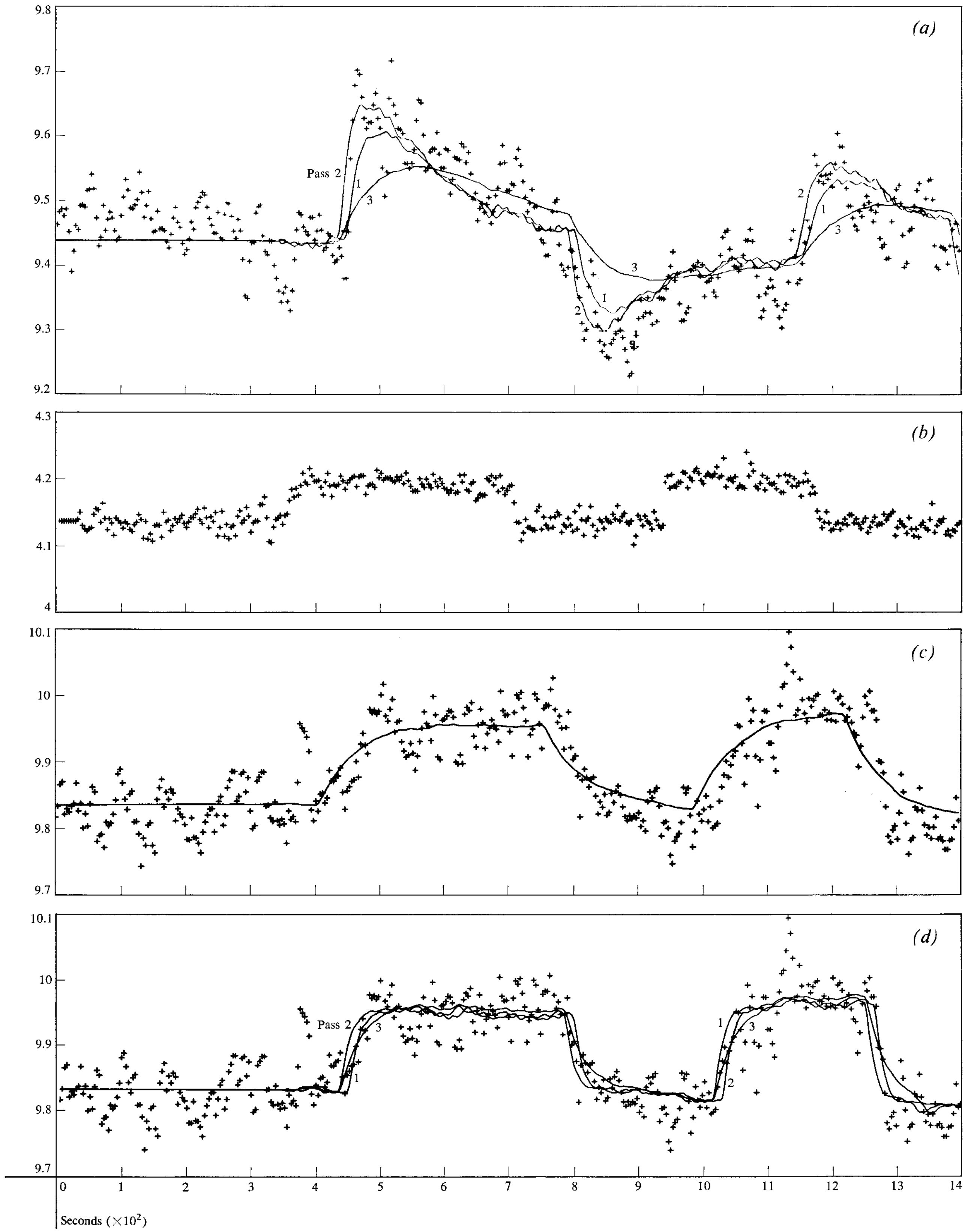


Figure 8 (e) Run 3, hybrid solution, (f) Run 4, integral solution, (g) Run 4, process residual with integral solution, (h) Run 4, derivative solution.

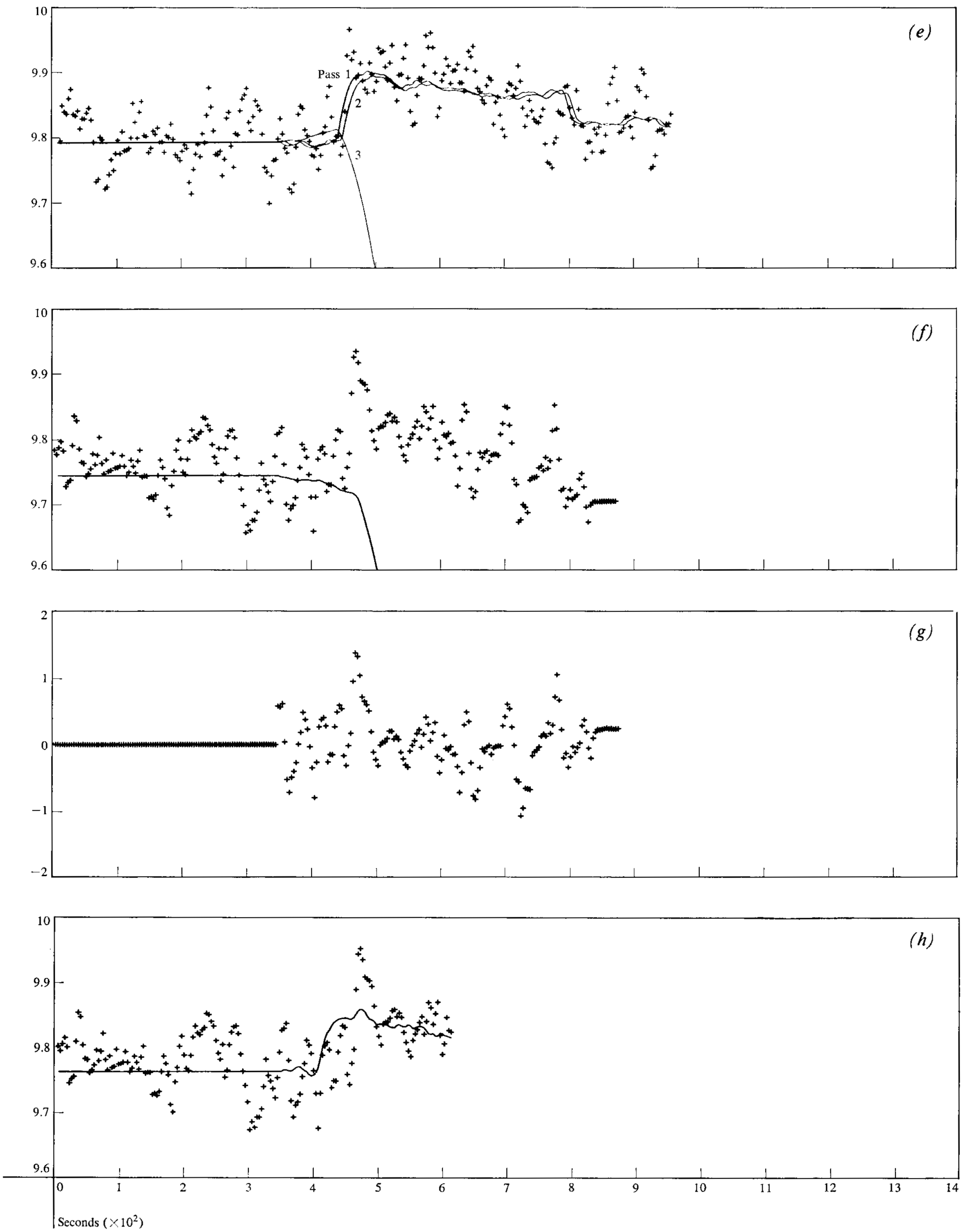


Table 3 Identification results for Series I experimental data.

Run number	*Experimental conditions* Noise variance	S/N ratio	Number of data points
1	0.00139	2.26	400
2	0.00115	1.91	400
3	0.00155	0.89	275
4	0.00150	0.55	175

Run number	*Parameters* K	A	τ	E_f
Integral method, no filter (τ from residual variance)				
1	2.35	0.00964	50.0	
2	2.40	0.0168	50.0	
3*	2.29	−0.0174	135.0	
4*	2.42	−0.463	110.0	
Integral method, no filter (τ from model output variance)				
1	2.28	0.00692	100.0	
2	2.48	0.0169	49.2	
3	2.94	0.00751	78.1	
Integral method, 20 sec filter (τ from residual variance)				
1*	−0.0003	−0.00004	−135.0	
2	2.43	0.0171	50.0	
3	2.16	0.00915	10.0	
4	2.04	0.0233	60.0	
Derivative method, no filter (Pass 2 in hybrid method)				
1	2.34	0.0520	90.6	0.65
2	2.33	0.0786	87.5	0.35
3	2.34	0.0762	93.8	0.066
4	1.83	0.0556	50.0	—
Derivative method, 20 sec filter				
1	2.61	0.0193	85.0	—
2	2.47	0.0301	80.0	—
3	2.19	0.0179	85.0	—
4	1.97	0.0498	87.5	—
Hybrid method Pass 3, no filter (using integral method with assigned τ†)				
1‡	2.24	0.00928	90.6	2.05
2§	2.51	0.0288	87.5	0.30
3§	2.20	−0.0222	93.8	$>10^{38}$
Hybrid method, optimal filter, 10 sec				
1‡	2.54	0.0288	87.5	—
2§	2.43	0.0329	84.4	—
3‡	2.20	0.0274	56.25	—

* Solution diverged.
† Values determined in Pass 2
‡ Pass 2 gave optimal solution.
§ Pass 3 gave optimal solution.

diverged in both Runs 3 and 4, indicates the unreliability of this method for a process having high noise, large τ, and large A. Even when used as part of the hybrid method, which assures a good estimate for τ, the integral method improved the estimate of A only in Run 2, as can be seen in Table 3. Still, the fit error for hybrid Pass 3 (integral method) was only slightly better than for Pass 2 (derivative method), i.e., 0.30 vs 0.35. Thus, the integral method did not contribute much to the value of the hybrid technique in the Series I experiments. It will be shown that the Series II and III experiemnts are highly different in this regard, however.

Table 3 shows that the integral method gave a much better estimate for τ (Runs 1 and 3) when that parameter was estimated by minimizing the model output variance rather than by minimizing residual variance.

Eqs. (22) and (23) are now applied to calculate the magnitude of the theoretical standard deviation for the estimated value of A. The following values are used: $\sigma_\eta = 0.00115$, $\tau = 90$, $A = 0.03$, $K_\pm = 0.12$, and $(V - \tau) = 400$. The last two values apply for the first step of Run 2. The calculations give $\sigma_A/A = 0.260$ and 0.248 for the derivative and integral methods, respectively. The observed standard deviation for all runs in Series I gave similar σ_A/A values, namely, 0.21 and 0.21 for the two methods, respectively.

For the large pulsewidth which was used in the identification experiments, the integral method thus has only slightly better theoretical accuracy for the pole estimate than the derivative method. We can conclude that it probably would have been better to use much narrower pulses in which case the hybrid method should be able to provide higher pole accuracy than the derivative method.

Prefiltering of the data has no particular value for Series I data when the integral method is used (Table 3). For the derivative method the opposite is true. Note that the inaccurate estimation of τ for the low signal in Run 4 is considerably improved by prefiltering.

The effects of different filter time constants with the derivative method are shown in Figs. 9 and 10. Note that the minimum fit error in Fig. 9a occurs approximately at 10 seconds time constant (which then defines the experimentally acquired optimal filter). The corresponding value of A for Run 1 is 0.0288 at 10 seconds. The optimal filter by Eq. (29) should have $1/\mathcal{K}_2 = 13.8$ seconds. This is quite good agreement between the theoretical and experimental filter values since the difference is not much more than one sample interval (i.e., 3.49 seconds).

The variation in K and τ with the filter constant is shown in Figs. 9c and 9d. Run 3, which has the smallest signal-to-noise ratio among the runs illustrated, breaks the pattern of the τ curves by producing an exceptionally low value at 10 seconds. Otherwise, the character of the filter influence on estimated parameter values is fairly uniform among the runs. Only A is strongly affected by the filter.

Fig. 9e shows that the function Z, defined by

$$Z \equiv K/(\tau + 1/A), \tag{32}$$

is approximately constant even for the largest filter of 80 seconds, for which τ has changed appreciably. Invariance of Z is a valuable property, since Z has more influence upon control loop characteristics than any one of the parameters K, A and τ alone. (The invariance in Z also was found to apply to the different solutions while searching for the optimal τ.)

The model output curves in Fig. 10 show how well the hybrid method will fit the data for the different filters. Pass 3 is optimal when no filter is used and with a filter time constant of 20 seconds; otherwise Pass 2 has slightly lower fit error. When a filter with 80 seconds time constant is used, the waveform is flattened out to the extent that even the derivative method loses its ability to recognize the length of the transport delay accurately. The most significant conclusions for the Series I data analysis are:

1. The hybrid method (with filter iterations) is usable at a signal-to-noise ratio as low as 0.5.
2. The data prefiltering technique is very beneficial for the estimation of A.
3. The filter must have an accurately chosen time constant, which is best found by the iterative scheme proposed, unless A is approximately known. In that case, Eq. (29) will provide the needed filter time constant.
4. The formulas for the standard deviation of the estimated value of A give approximately the same result as the observed standard deviation for both the derivative and integral method.
5. Bias errors exist which make the integral method give estimates for τ that are too low (with residual performance criterion) and the derivative method (without filtering) gives estimates of A that are too large.
6. The hybrid method eliminates these bias errors. For this series this is primarily done by selection of the best filter and obtaining the answer by the derivative method.

- *Series II, basis-weight/stock-valve-position transfer function*

Series II has the most difficult data for identification purposes. One reason is that the signal-to-noise ratio is very low. Two runs having the highest signal-to-noise ratios (1.05 and 0.74) among 12 identification runs made during two consecutive weeks will be illustrated. A second difficulty was caused by a nonpartitioned process model. Rather than introducing dry material flow measurement as shown in Fig. 1, only the stock valve position signal as input and basis weight as output were measured. This led to high sensitivity of the results to valve sticking and consistency disturbances. The accuracy in the estimation undoubtedly was decreased by these factors. Nevertheless, Series II clearly illustrates many characteristic features of the proposed identification method.

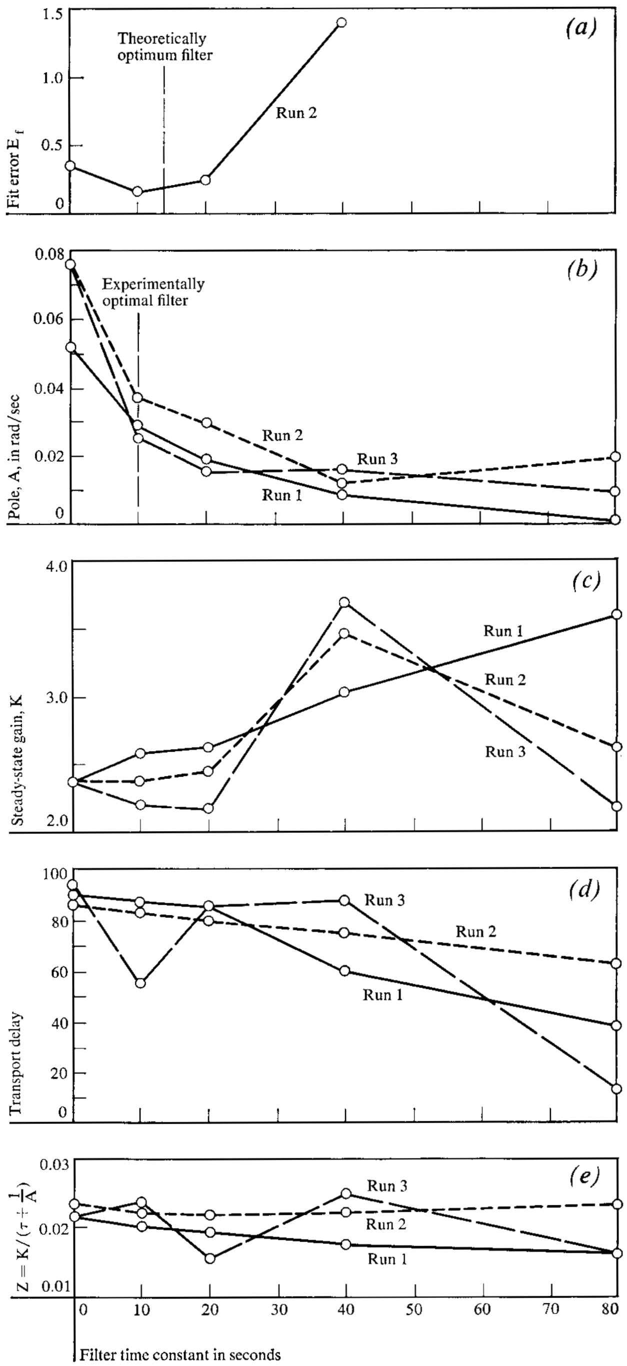
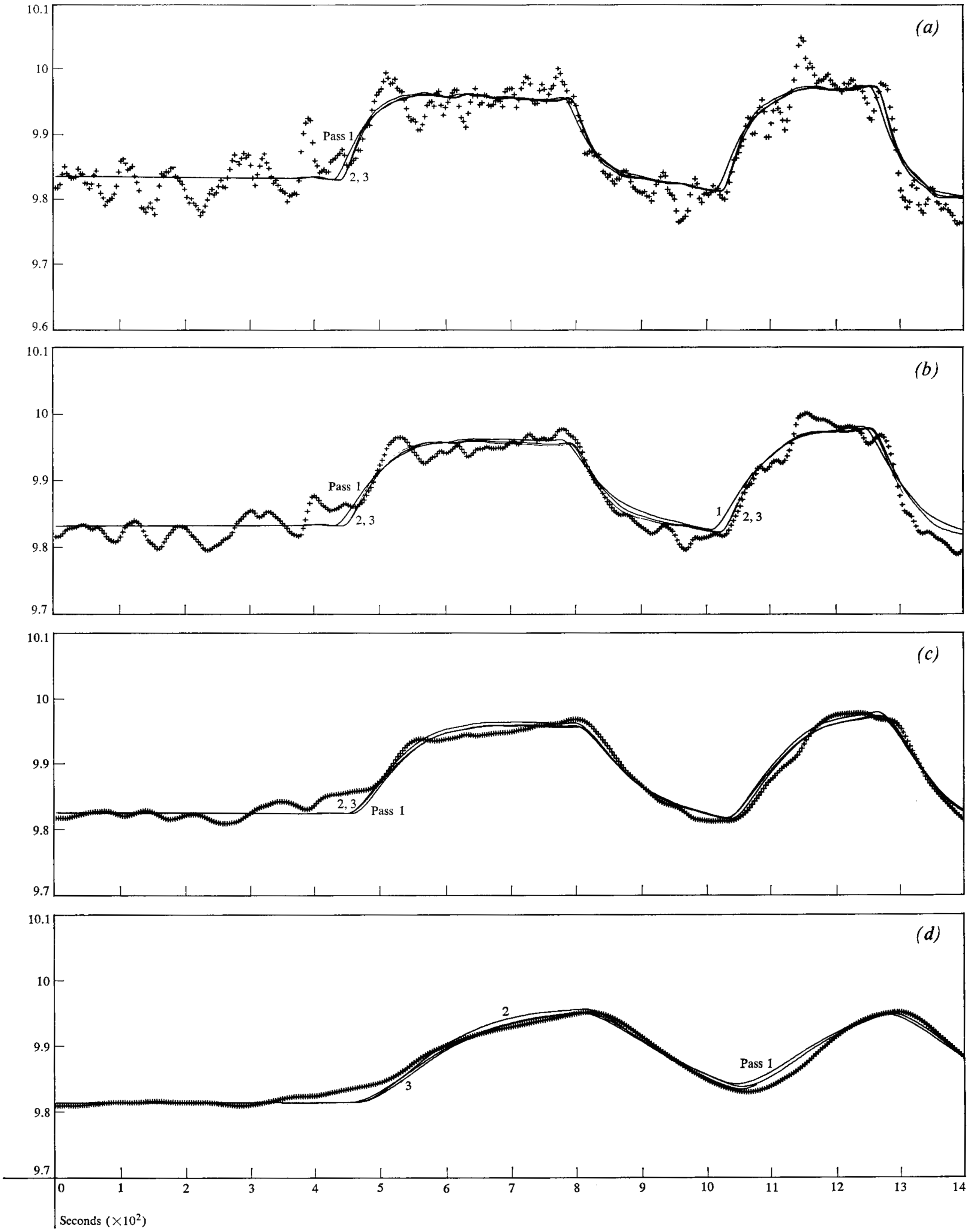


Figure 9 Effect of filter time constant on estimation of parameter values for Series I data.

Figs. 11 and 12 show the results of identification by the derivative, integral and hybrid methods for two runs from Series II. The curves shown for the hybrid method in Fig. 11 are the result of Pass 3 by this method (Pass 2 is identical

Figure 10 Effect of filter time constant with hybrid method used for Series I data; (a) filter time constant, 10 seconds (optimal value), (b) 20 seconds, (c) 40 seconds, (d) 80 seconds.



with the derivative method). Pass 3 is optimal (better than Pass 2) only for Run 1 at zero and 40 seconds time constant.

The theoretically best filter was determined by using $A = 0.011$ in Eq. (29). Fig. 11a shows that in Series II, as was the case in Series I, the filter value affects the estimate of A very strongly when the derivative method is used. Without filtering, that method gives an A-value which is far too large (Figs. 12b and c illustrate how Pass 2 exaggerates the steepness of the process response). Even without filtering, Pass 3 by the hybrid method improved the A estimate considerably, however (Fig. 11a). Figs. 11a and 11b show how the answers for A and K by the integral, derivative and hybrid methods all converge to nearly the same value for each run when filter time constant approaches the theoretically best value. The pole estimate actually converges approximately to the same value for both Run 1 and 2, but K approaches different values for the two runs. The latter effect is probably due to the nonlinearities previously mentioned. The estimates of τ also appear to converge, although the effect is not so clear as for A and K. The same trends were seen in the results on many more runs in Series II besides those shown.

The model response by the integral method in Run 1 is shown in Fig. 12a. The response by the hybrid method without filters for Runs 1 and 2 are shown in 12b and 12c, respectively; and with nearly optimal filter value (40 seconds) for Runs 1 and 2 in Figs. 12d and 12e, respectively.

The analysis of the experiments lead to the following conclusions:

1. The derivative technique must be backed up with the integral technique or else a filter must be used to give an acceptable pole estimate.
2. Not only the derivative technique but also the integral technique is improved by the use of the theoretically optimal filter.
3. The hybrid method worked successfully for all runs in this series, one of which had a signal-to-noise ratio as low as 0.35.

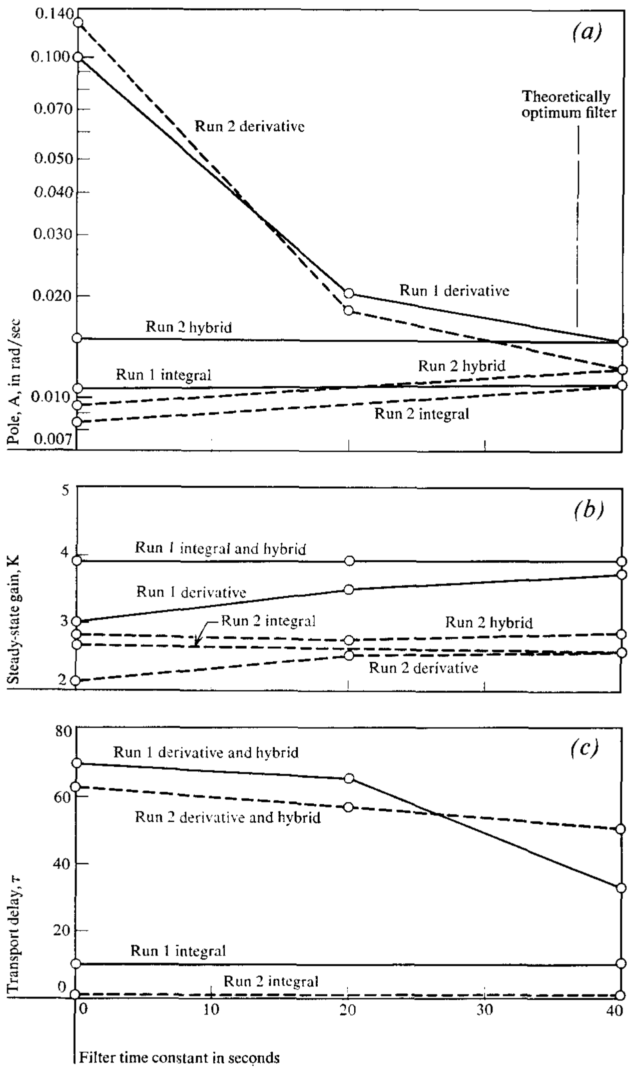


Figure 11 Effect of filter time constant on estimation of parameter values for Series II data.

• *Series III, reel-moisture/dryer-steam-pressure transfer function*

Series III uses data from a dryer having moisture as the output signal and steam pressure as input. This process has a long time constant and a relatively short transport delay. The integral method was tried on 17 identification experiments as well as on normal on-line runs, and always gave a satisfactory answer. A typical result is shown in Fig. 13. The derivative method and Pass 1 of the hybrid method (which uses the derivative technique) always diverged in the test on four sets of data having the highest signal-to-noise ratio among the 17 sets. The hybrid method then automatically selected the integral method in Pass 2 for all runs. Convergence and reasonable parameter values were acquired for all runs.

The conclusion from Series III is that the integral method not only backs up the pole estimate for the derivative method in the event a proper filter is not selected, but in certain applications it is the only workable technique.

The runs in Series I, II and III typically utilized four iterations for τ in Pass 1, 12 iterations in Pass 2, and of course, one iteration in Pass 3 by the hybrid method. The solution time per τ-iteration with 400 data points for each variable was about two seconds on the IBM 7094 computer, and three minutes on the IBM 1710 computer.

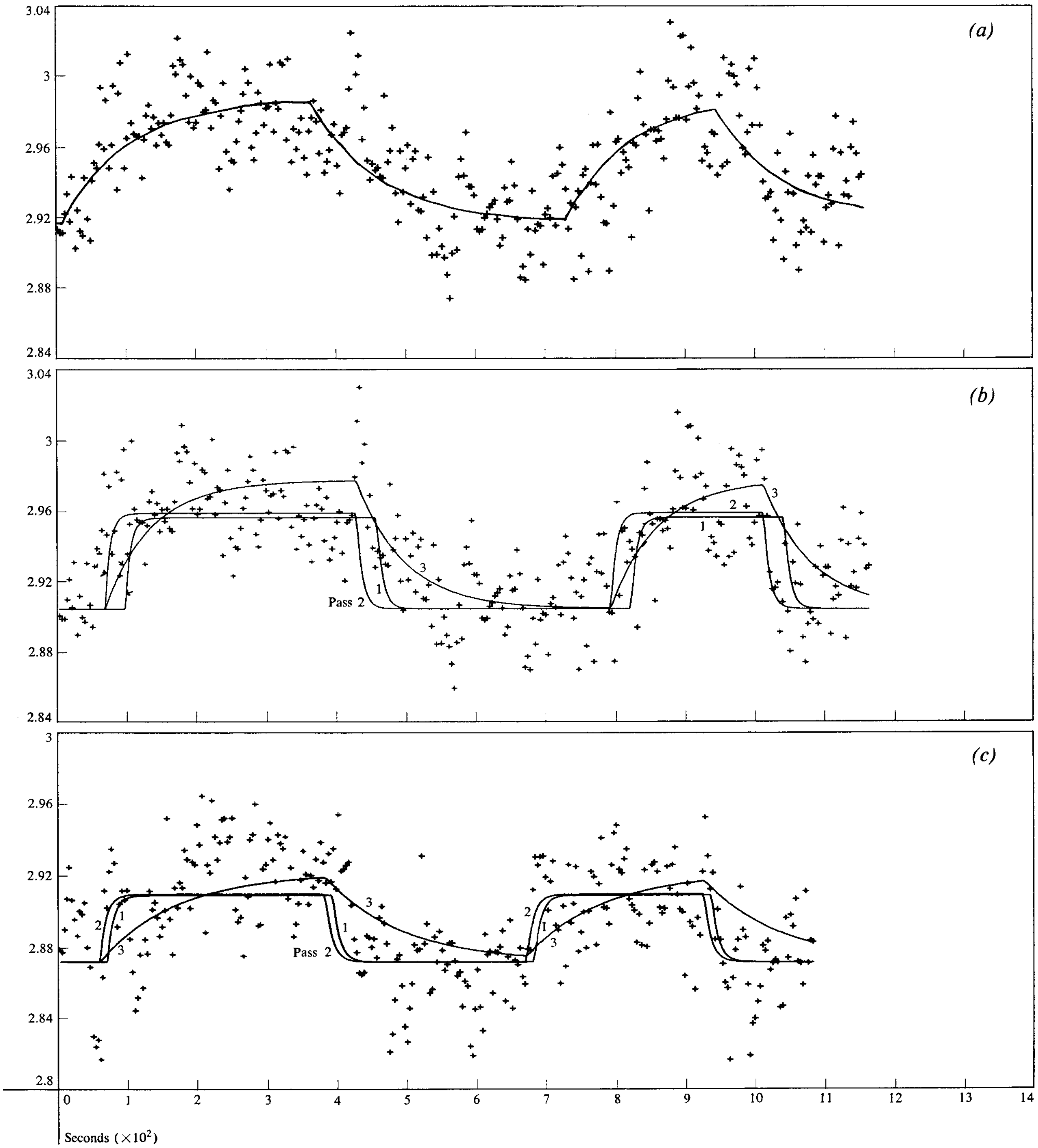


Figure 12 Series II experiments; (a) Run 1, integral solution, (b) Run 1, hybrid solution with no filtering, (c) Run 2, hybrid solution with no filtering.

Conclusions

The hybrid identification technique has been proven to work reliably on several different processes with strong disturbances. In the testing it has so far converged to a reasonable answer every time it was used even though many of the test cases have had a signal-to-noise ratio as low as 0.5 and occasionally much lower.

The hybrid method proposed here is far superior to its consitituent techniques, one of which (the derivative

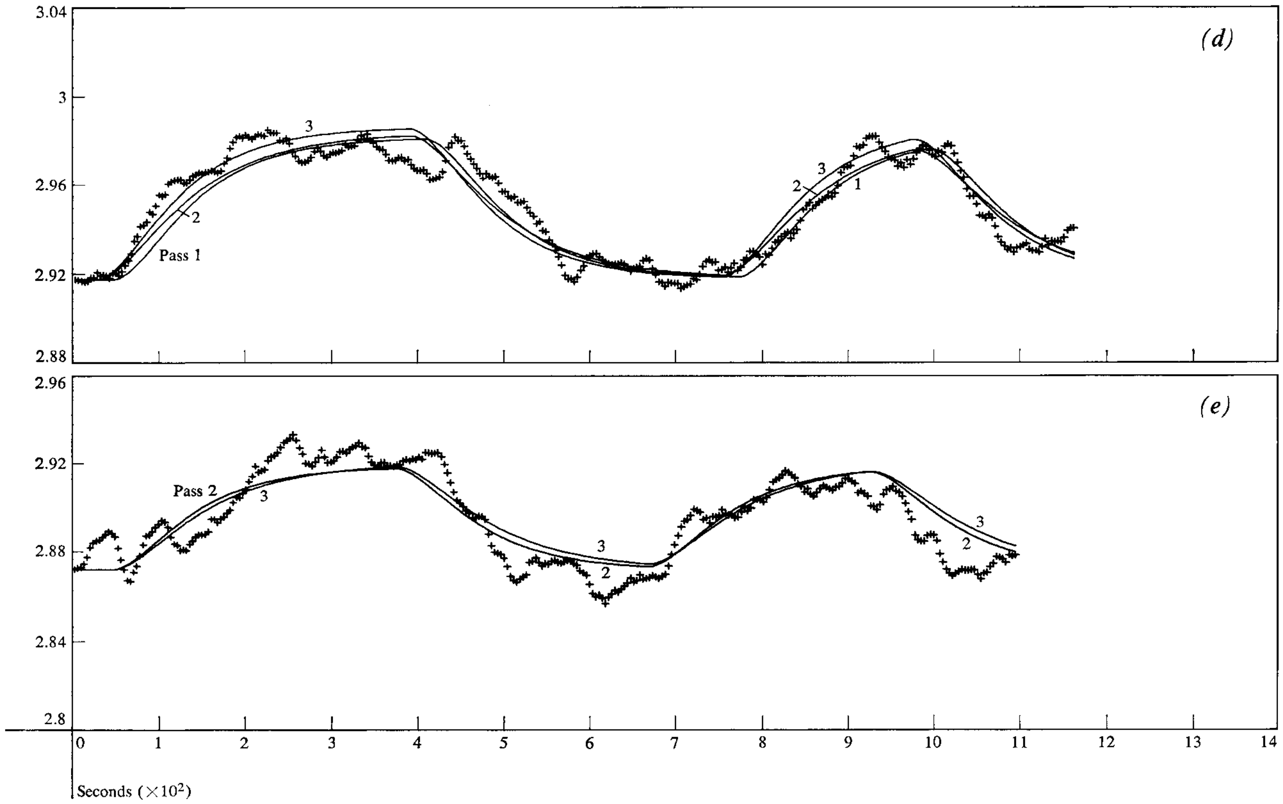


Figure 12 (d) Run 1, hybrid solution with optimal, 40-second filter time constant, (e) Run 2, hybrid solution with optimal, 40-second filter time constant.

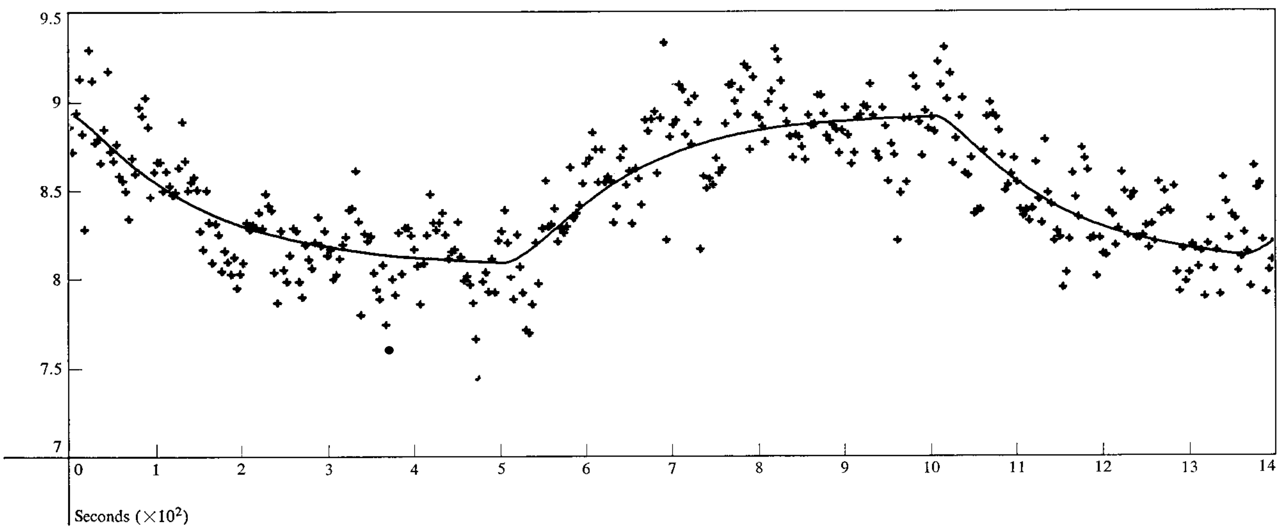


Figure 13 Series III experiment, solution by integral method.

method) has been much described in literature. The advantage of the hybrid method is most obvious from the fact that it alone could handle *all* the dynamic data presented in this paper.

The critical weakness of the derivative method is its inaccuracy in estimation of poles in the presence of strong process noise. This fact was demonstrated experimentally and theoretically, as was the fact that the integral technique is much better in this regard if the time between the transients in the test signal is not chosen too long.

The formula derived for quasi-optimal data filtering in the presence of wide bandwidth noise agrees very well with the observed results. The formula can, therefore, be used for preselection of the filter value if the process pole is approximately known. In such cases one can then eliminate all or some of the iterations for the filter time constant beyond the first.

The character of the parameter errors due to erroneous estimates of the reference values for the input and output variables was demonstrated theoretically and experimentally. Both approaches had good agreement. It was shown that a test signal wave form can be chosen so that the effect of such errors is small. If a single-direction process perturbation is used, these conditions cannot be satisfied. The error in parameter estimation will increase proportionally to the length of the experiment for the integral method, but is independent of the same factor for the derivative method.

The value of the process partitioning technique was illustrated by non-linear phenomena which occurred during collection of the data shown here. Severe non-linear effects were cancelled out in the data of Series I, which used a partitioned model, but relatively large parameter discrepancy, for gain in particular, resulted in the non-partitioned case (Series II).

The hybrid method can be expanded simply to handle more than three parameters per transfer function element by the same principles as shown here. Minor portions of the program are affected by such changes.

Since the hybrid method requires only a fairly small amount of data, a simple experiment on the process, which causes no loss of production, has very modest computer requirements, and can handle processes with high disturbance levels; it is ideally suited for process control.

References

1. K. J. Astrom, "Computer Control of a Paper Machine—An Application of Linear Stochastic Control Theory," *IBM Journal* 11, 389 (1967), this issue.
2. J. G. Truxal, "The Concept of Adaptive Control" Chapter 1 of *Adaptive Control Systems*, Edited by E. Mishkin and L. Braun, Jr., McGraw-Hill, New York, 1961.
3. C. A. Galtieri, "Problems of Estimation in Discrete Time Processes," Research Paper No. RJ-315, IBM Research Laboratory, San Jose (1964).
4. C. A. Galtieri, "The Problem of Regulation in the Computer Control of Processes," Research Paper No. RJ-382, IBM Research Laboratory, San Jose (1966).
5. G. E. Box and G. M. Jenkins, "Some Statistical Aspects of Optimization and Control" *Journ. Royal Stat. Soc.* **24**, 297–343 (1962).
6. R. W. Koepcke, "A Discrete Design Method for Digital Control," *Control Engineering* **13**, No. 6, 83–89 (June, 1966).
7. P. E. A. Cowley, "An Active Filter for the Measurement of Process Dynamics," *Proc. ISA Eng. Conf.* (1964).
8. N. R. Goodman, S. Katz, B. H. Kramer and M. T. Kuo, "Frequency Response from Stationary Noise," *Technometrics* **3**, No. 2, 245–268 (1961).
9. M. S. Bartlett, *An Introduction to Stochastic Processes*, Cambridge University Press, London, 1962.
10. V. Grenander, and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, Wiley, New York, 1957.
11. R. E. Kalman, "Design of a Self-Optimizing System," *ASME Transactions* (February, 1958).
12. R. C. K. Lee, *Optimal Estimation Identification and Control*, M. I. T. Press, Cambridge, Massachusetts, 1964.
13. E. Parzen, "Analysis and Synthesis of Linear Models for Time Series," Technical Report No. 4, April 18, 1966; Department of Statistics, Stanford University.
14. E. B. Dahlin, "An Analysis of Methods for Process Identification," Development Report, IBM Control Systems Development Center, San Jose, California.
15. A. E. Beecher, "Dynamic Modeling Techniques in the Paper Industry," *TAPPI* **46**, No. 2, 117–120 (February, 1963).
16. R. Penrose, "A Generalized Inverse for Matrices," *Proc. Camb. Phil. Soc.* **51**, 406–413 (1955).
17. E. Sansone, *Orthogonal Functions*, Interscience Publishers, New York, 1959.
18. E. E. Fisher, "The Indentification of Linear Systems" *Proc. JACC*, 473–475 (1965).
19. W. B. Davenport, Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.
20. M. G. Salvadori, and M. L. Baron, *Numerical Methods in Engineering*, Prentice-Hall, Englewood Cliffs, N. J., 1959.

Received October 10, 1966