Computer Control of a Paper Machine : An Application of Linear Stochastic Control Theory

Åström, Karl Johan

Published in:
IBM Journal of Research and Development

DOI:
10.1147/rd.114.0389

1967

K. J. Åström*

Computer Control of a Paper Machine—an Application of Linear Stochastic Control Theory

Abstract: This paper describes an attempt to apply linear optimal control theory to computer control of an industrial process. The applicability of the theory is discussed. Particular attention is given to the problem of obtaining a mathematical model of process dynamics and disturbances. Results of actual measurements as well as results from on-line control experience are presented.

1. Introduction

The work presented in this paper was carried out at the IBM Nordic Laboratory in connection with the installation of an IBM 1710 system[1,2] in the Billerud Kraft Paper Mill at Gruvön, Sweden. Its objectives were twofold: to solve some specific control problems associated with the Billerud installation and to derive systematic methods for design of control laws for on-line computers. The second objective was justified by the fact that a considerable amount of systems work is required to install a control computer on an industrial process. This work can be substantially reduced if systematic methods for the design of control laws are available. Such a synthesis procedure has been developed. The synthesis procedure consists of a few theorems and a FORTRAN program which enables us to obtain mathematical models and control strategies directly from measured plant data. The existing identification programs are limited to multiple-input, single-output systems. For such systems we can always find a canonical form for the mathematical models and the identification problem can then be solved by one computer program. Although the mathematical solution of the identification problem is directly applicable to multivariable systems, the actual identification programs will differ with the structure of the systems. For multivariable systems we will thus need one program for each structure.

We have found the synthesis procedure to be a practical and convenient tool to obtain on-line control strategies. The procedure has been applied to design control algorithms for several loops on the paper machine such as basis weight, moisture content and refining.[3,4] In all cases we have considered the systems simply as input-output systems, and we have not exploited the particular characteristics of the paper machine. So far we have extensive practical experience only with single-input, single-output systems.

There are many possible alternatives to the method proposed in this paper. One possibility is a straightforward trial and error method: program a three-term control algorithm in the control computer and adjust its parameters until acceptable performance is obtained. This method has the decided advantage of being simple and has also been applied to most of the control loops of the Billerud system. However, in order to adjust a three-term controller the performance of the control loop has to be evaluated for a number of parameter combinations. This evaluation may take considerable time if the process dynamics are slow and if many parameter combinations have to be tried. In our particular case evaluation of the performance of the basis weight loop may take a few hours. For the refiner loops the corresponding time is ten times longer. Also, when using the trial and error approach one always wonders whether still better results could be achieved with a more complicated control law. When using the systematic approach of this paper the results of the process identifications show what can possibly be achieved and what the limitations are. When the identification programs are available the design of a control loop can be made with little effort. Our present experience indicates that in typical cases a control loop can be designed with two identification experiments and two evaluation tests.

In this paper we will describe the synthesis procedure and its application to one typical practical problem, namely basis weight control. Basis weight fluctuations

* Present address: Lund Institute of Technology, Lund, Sweden.

were investigated during a feasibility study before the system was under computer control. The fluctuations had a standard deviation of 1.3 g/m^2. In the feasibility study the target value for standard deviation of basis weight under computer control was set to 0.7 g/m^2.

The control computer was installed late in December 1964. Two experiments for determination of process dynamics were performed in March 1965 and the first operation of on-line basis weight control was done on April 28, 1965 covering a period of 10 hours. Since that date a large number of tests have been performed and the basis weight loop has been in continuous operation since the beginning of 1966. In actual operation we can now consistently achieve standard deviations of 0.5 g/m^2 wet basis weight and 0.3 g/m^2 dry basis weight.

Important phases of the project are briefly discussed in Section 2. In Section 3 we give a mathematical formulation of the basis weight problem. The solution of the mathematical problem and our synthesis procedure are given in Section 4. Section 5 describes the practical experiments that were performed in order to determine the process dynamics and in Section 6 we give practical results from on-line control operations.

2. Review of important phases of the project

The work described in this paper was initiated at an early stage during the feasibility study undertaken from May to September 1963. A great many of the tasks which might be assigned a control computer were considered in order to discover the common characteristics of the problems and to select problems suitable for closer investigation. During this phase we found that the problems were essentially of two types: Those associated with controlling the process during *normal operation* and during *grade changes*. The feasibility study indicated that the solution of both these problems and the production planning problem could economically motivate the computer installation.

Grade-change control involves making the transition from one grade to another in the shortest possible time, subject to various constraints such as a given risk of paper break and limitations of physical variables. The solution of grade change problems using optimal control theory was considered. To solve these problems we need accurate dynamic models as well as accurate descriptions of the constraints.

The problem of controlling the mill during normal operation is essentially a regulation problem. The process variables must be kept as close as possible to given reference values. Product specifications are usually set in terms of upper and/or lower limits on the variables. Because of disturbances, the set points have to be chosen well within the specified limits in order to ensure that a given amount of the production falls within the specifications. Closer control makes it possible to decrease the magnitude of the fluctuations. The set points can then be chosen closer to the specified limits with a given risk for production outside the specifications.

Moving set points closer to the specifications either reduces raw material and power consumption or increases production. Even a moderate decrease in the variations of process variables can thus realize considerable economic gains. Control of basis weight and moisture content are typical examples of normal-operation problems.

A large amount of powerful theoretical results are available for the solution of both types of problems: optimal control theory for the grade change problem and linear stochastic control theory for the regulation problem. In both cases application of these theoretical tools requires the solution of identification problems.

We decided to concentrate on the normal-operation problems rather than grade-change problems for the following reasons:

- We believed that the results would be more general.
- The identification of the non-linear models required for the application of optimal control theory is considerably more difficult than the identification of the linear model required for the steady state control.
- A solution of the regulation problem is a prerequisite for solving more complex process control problems including that of optimization.

It was thus decided to investigate the possibilities of obtaining a systematic procedure to control the process during normal operation based on linear stochastic control theory. Control of basis weight and moisture content and refining were critical problems in the particular application and it was therefore decided to use these as test cases.

Having decided to attack the regulation problem we investigated the applicability of linear stochastic control theory. The essential assumptions for this theory are:

- The process dynamics can be characterized by linear equations.
- Disturbances can be described as sample functions of second-order random processes.
- The criterion for the operation is to minimize mathematical expectation of a quadratic form in state variables and control variables.

Admittedly, in practice there sometimes are "upsets," which give rise to large deviations. These upsets may originate in many different ways; for example, from equipment or instrument malfunctions, which all require particular corrective actions. It is questionable whether these types of disturbances can be described by probabilistic models. In the following we disregard these upsets.

During the feasibility study we verified that during

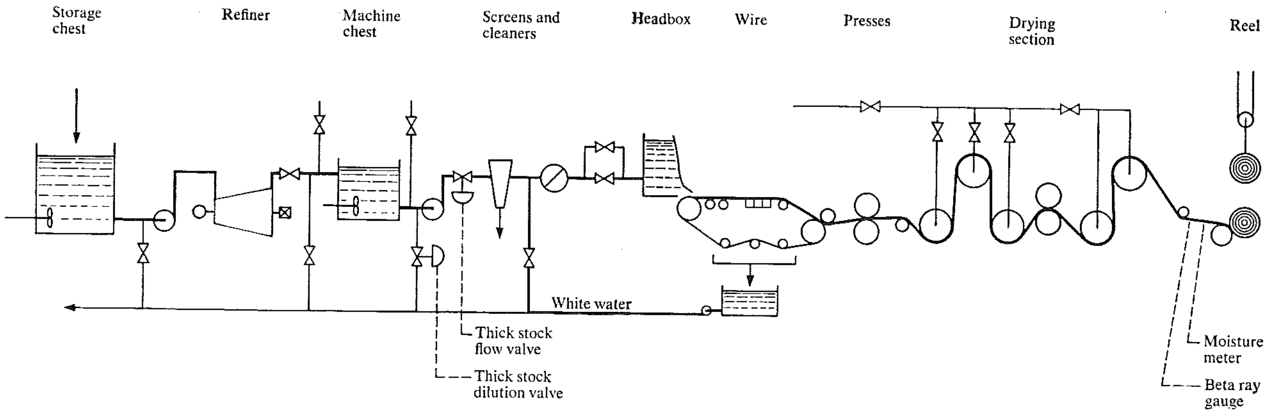


Figure 1 Simplified diagram of a kraft paper machine.

normal operation the probability distributions of the fluctuations of the process variables were close to the normal distribution. The problem of operating as close as possible to the specifications with a given risk of paper-break is then equivalent to controlling the process in such a way that the variance of the output is as small as possible.

During the feasibility study and the initial phase of the project we also performed experiments which indicated that the disturbances occurring during normal operation were so small that the system could be described by linear equations. The indications obtained in these early experiments were subsequently verified by experiments on the computer-controlled system.

Linear stochastic control theory was found applicable. The chief difficulty in the application of this theory to a practical industrial problem is to obtain adequate models of the process and its disturbances. Once the models are obtained the optimal control algorithms are given by known formulas.[5–11] Computer programs to generate the algorithms from the coefficients of the mathematical model are also available.[5] The mathematical models are obtained from experiments on the process. The input variables are systematically perturbed and the corresponding variations in the output variables are observed. From this data we obtain information about the process dynamics and the disturbances.

In order to obtain mathematical models for the process dynamics and the disturbances we have developed an identification procedure,[12,13] which has been programmed in FORTRAN. The experimental data are entered into the program and the computation results in a mathematical model of the process and the disturbances.

We were able to test part of the solution, modelling of disturbances and design of optimal filters, on a quality-control problem[1] before the control computer was installed. This successful test weighed heavily in the decision to continue work on real-time control. As work progressed it was discovered[14] that for the special case of steady-state control the optimal control algorithms could be derived very simply from the results of the identification.

Our investigations have thus resulted in a procedure for synthesizing control strategies for on-line control. This procedure is based on linear stochastic control theory and is particularly well suited for a situation where:

- The problem is to keep process variables close to fixed operating points.
- The process is linear with time delay, but time invariant.
- Disturbances are unavoidable and can be characterized as samples of stationary random processes with rational power spectra.
- The process cannot be taken out of production but process analyses have to be performed under normal operating conditions.

3. Mathematical formulation of the basis weight regulation problem

In Fig. 1 we show a simplified diagram of the parts of the paper machine that are of interest for basis weight control. Thick stock, i.e., a water fibre mixture with a fibre concentration of about 3% comes from the machine chest. The thick stock is diluted with white water so that the headbox concentration is reduced to 0.2 to 0.5%. On the wire, the fibres are separated from the water and a web is formed. Water is pressed out of the paper web in the presses, and the paper is then dried on steam-heated cylinders in the dryer section.

In this particular case it is possible to influence the basis weight by varying the thick stock flow and/or the thick stock consistency (fibre concentration in thick stock). Both these variables will directly influence the amount of fibres flowing out of the headbox and thus also the basis weight. The control variables are manipulated via set points of analog regulators which control the thick stock flow valve and the thick stock dilution valve shown in Fig. 1.

Basis weight is measured by a beta-ray gauge set at a fixed position at the dry end. The output of this instrument will be proportional to the mass of fibres and water per unit area, i.e., the *wet basis weight*. This is because the coefficient of absorption of beta rays in fibres and water is approximatively the same. In order to obtain *dry basis weight*, i.e., the mass of fibres per unit area, the beta-ray gauge reading has to be compensated for the moisture in the paper sheet. Moisture is measured by a capacitance gauge. In our particular case this gauge can traverse the paper web although it is normally set at a fixed position.

There is also a beta-ray gauge before the drying section. Basis weight is also measured by the machine tender and in the test-laboratories. When a reel of paper is produced its weight and size are determined giving a very accurate value of the average basis weight of the reel. This information is used to calibrate the other gauges. An analysis of the information sources for basis weight has shown that:

- The reel weight and dimension information can be used to compensate the drift in the beta-ray gauge.
- The high frequency fluctuations in the moisture gauge and the beta-ray gauge signal have similar characteristics and a good estimate of dry basis weight is

$$y = \mathrm{WSP}\,(1 - \mathrm{MSP}), \tag{1}$$

  where WSP is the calibrated beta-ray gauge signal and MSP is the signal from the moisture gauge. The difference between dry basis weight and the estimate y of Eq. (1) is essentially a stationary random process which contains many high frequencies.
- The estimate of dry basis weight given by (1) is improved very slightly when laboratory measurements are taken into account.

Fluctuations in basis weight during normal operation have been investigated. There are variations of weight in both the machine direction and the cross direction of flow. In our case it was found that the cross-direction profile is stable if certain precautions are taken. The fluctuations observed can be described as normal random processes. There is a considerable amount of low-frequency variation. Data in the records have been divided into samples covering about five hours for analysis. Before carrying out a time-series analysis the trend is removed. In Fig. 2 we show the covariance function of basis weight variations in a typical case. In all cases studied we found that the variations in basis weight had a standard deviation equal to or greater than 1.3 g/m^2.

We investigated the possibilities of controlling the basis weight by careful regulation of machine speed, thick stock flow and consistency. Experiments have also been performed to establish the correlation of fluctuations in basis weight with fluctuations in thick stick flow and consistency.

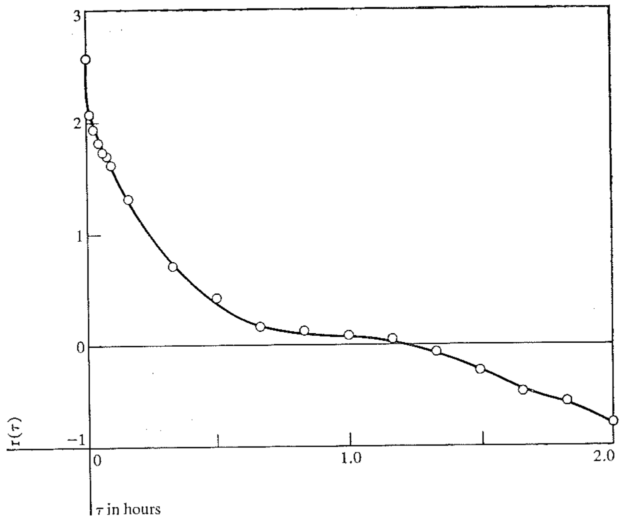


Figure 2 Covariance function of fluctuations in wet basis weight obtained during the feasibility study, before control computer was installed.

The results of these investigations have shown that in this particular application it is not possible to keep the basis weight constant by careful regulation of machine speed and fibre flow. It was therefore decided to control basis weight by feedback from the measurements at the dry end of the machine to thick stock flow or thick stock consistency.

The dynamics of the paper machine are such that there is a coupling between dry basis weight and moisture content. An increase of thick stock flow or thick stock consistency results in an increase in moisture content as well as dry basis weight. A change of steam pressure in the drying section will, however, influence the moisture content of the paper but not the dry basis weight. This coupling was not known to us initially but the first identification experiment showed the effect clearly. The special character of the coupling implies that the loop for controlling dry basis weight can be considered as a single-input, single-output system. The control actions of the basis weight loop (thick stock flow) will, however, introduce disturbances in moisture content. These disturbances can be eliminated by using thick stock flow as an input to the moisture control loop. Control of *dry* basis weight can thus be considered as a system with one input, thick stock flow, and one output, the variable y as given by Eq. (1). Notice, however, that the moisture control system must be considered as a system with two inputs and one output. As long as there is only one output there are no serious structural problems and the basic identification algorithm can be easily modified. See Ref. 12, p. 94, and Ref. 13.

It is explained in Section 2 why the variance of dry basis weight is a good measure of the performance of the basis weight control loop. Now if the estimate y, of

Eq. (1), differs from dry basis weight by a quantity which is essentially high frequency noise, minimizing the variance of dry basis weight is equivalent to minimizing the variance of y. Notice that this reasoning fails if the deviation conains low frequencies!

The criterion is thus to control the system in such a way that the variance of the output signal y is minimal. To complete the formulation of the control problem we now need a description of the process dynamics and the characteristics of the disturbances.

The corrections that are required to control the process during normal operation are so small that the system can be described by linear differential equations with a delayed input. The time-delay T_d depends on the time it takes to transport the fibres along the paper machine. The equations determining the dynamics can be partly determined by continuity equations for the flow. The degree of mixing in the tanks is uncertain, however. There is also a rather complicated mechanism that determines the amount of fibres that pass through the wire. A direct derivation will thus give a very uncertain model for the process dynamics.[3,15]

Since a digital computer is to be used to implement the control law we will consider a discrete time model directly. If it is assumed that the sampling interval T_s is chosen so that T_d is an integral multiple of T_s and if we also assume that the control signal is constant over the sampling interval the process dynamics can be expressed by the general linear model

$$\begin{aligned} y(t) &+ a_1'y(t - T_s) + \cdots + a_l'y(t - lT_s) \\ &= b_0'u(t - T_d - T_s) + \cdots \\ &\quad + b_{l-1}'u(t - T_d - lT_s). \end{aligned} \tag{2}$$

Now let the sampling interval T_s be the time unit. Introduce the shift operator

$$zy(t) = y(t + 1) \tag{3}$$

and the polynomials

$$A'(z) = 1 + a_1'z + \cdots + a_l'z^l$$

$$B'(z) = b_0' + b_1'z + \cdots + b_{l-1}'z^{l-1}.$$

Equation (2) can now be written

$$y(t) = \frac{B'(z^{-1})}{A'(z^{-1})}\, u(t - k), \tag{4}$$

where

$$T_d = (k - 1)T_s. \tag{5}$$

Equation (4) would apply if there were no disturbances. Because of linearity the disturbances can always be represented as an equivalent disturbance $d(t)$ in the output:

$$y(t) = \frac{B'(z^{-1})}{A'(z^{-1})}\, u(t - k) + d(t). \tag{6}$$

If the disturbance $d(t)$ is a stationary* random process with a rational power spectral density it can always be represented as

$$d(t) = \lambda\, \frac{C'(z^{-1})}{D'(z^{-1})}\, e(t), \tag{7}$$

where $\{e(t), t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of independent, equally distributed random variables. The polynomials $C'(z^{-1})$ and $D'(z^{-1})$ can always be chosen so that the functions $z^m\, C'(z^{-1})$ and $z^m\, D'(z^{-1})$ have no zeros outside the unit circle.

Introducing (7) into (6) and writing the two terms on common denominators we thus find that the input-output relation can be described by the model

$$A(z^{-1})y(t) = B(z^{-1})u(t - k) + \lambda C(z^{-1})e(t), \tag{8}$$

where $\{e(t)\}$ is a sequence of normal equally distributed random variables, $e(t)$ is independent of $e(s)$ for $s \neq t$ and $e(t)$ is also independent of $y(t)$ and $u(t - k)$. Hence,

$$A(z) = 1 + a_1z + \cdots + a_nz^n$$

$$B(z) = b_0 + b_1z + \cdots + b_{n-1}z^{n-1}$$

$$C(z) = 1 + c_1z + \cdots + c_nz^n.$$

The polynomials A, B and C of Eq. (8) are formally of the n^{th} order. This is no loss in generality because we can always put trailing coefficients equal to zero.

Equation (8) is thus the general model for a linear sampled n^{th} order system with a time-delay that is an integral multiple of the sampling interval and which is subject to disturbances that are stationary random processes with rational power spectra.

4. Solution of the mathematical problem

We now consider the problem of controlling the system, Eq. (8) in such a way that the variance of the output $y(t)$ is as small as possible. If the coefficients of the polynomials $A(z)$, $B(z)$ and $C(z)$ are known this problem is a straightforward application of linear stochastic control theory. In this particular case we can, however, derive the results directly. The direct derivation is simple; it gives insight into the problem and provides also a suitable algorithm for computing the control strategy.

In a practical problem we also have the additional problem of determining the coefficients of the model. The problem can thus be conveniently divided into two subproblems:

- To determine a mathematical model of the process and the disturbances (identification problem).

* If the disturbance is drifting we can often, in practice, take time differences of $d(t)$ until a stationary process is obtained.

- To determine the minimum variance control law for a system governed by the mathematical model (control problem).

These two problems will be discussed below.

- *Process identification*

As stated in Section 3, it is very difficult to derive the mathematical model from first principles. Instead we have determined the model (8) directly from measurements on the process. When making the measurements, the control variable of the process is perturbed and the resulting variations in the output are observed. On the basis of recorded input-output pairs $\{u(t), y(t), t = 1, 2, \cdots, N\}$ we then determine a model (8) of the process and the disturbances. Since the identification technique is discussed elsewhere,[1,13] we will not go into any details of the identification problem in this paper.

Let it suffice to say that the problem is solved by determining the maximum likelihood estimate of the parameters $\theta = (a_1, a_2, \cdots, a_n, b_0, b_1, \cdots, b_{n-1}, c_1, c_2, \cdots, c_n)$ of the model based on a sequence of input-output pairs $\{u(t), y(t), t = 1, 2, \cdots, N\}$. It has been shown[12] that maximizing the likelihood function is equivalent to minimizing the loss function

$$V(\theta) = \frac{1}{2} \sum_{t=1}^{N} \epsilon^2(t), \tag{9}$$

where the numbers $\epsilon(t)$ are related to the input-output signal by the equation

$$C(z^{-1})\epsilon(t) = A(z^{-1})y(t) - B(z^{-1})u(t-k). \tag{10}$$

The numbers $\epsilon(t)$ can be interpreted as being one-step-ahead prediction errors.

When we have found a $\theta = \hat{\theta}$ such that $V(\hat{\theta})$ is minimal we get the maximum likeihood estimate of λ from

$$\hat{\lambda}^2 = \frac{2V(\hat{\theta})}{N}. \tag{11}$$

The identification problem is thus reduced to a problem of finding the minimum of a function of several variables.

The function V is minimized recursively by a gradient routine which involves computation of the gradient V_θ of V with respect to the parameters as well as the matrix of second partial derivatives $V_{\theta\theta}$. Due to the particular choice of model structure the computation of the derivatives of the loss function can be done very economically. In fact for large N the computations increase only linearly with the order of the model.

To obtain a starting value for the maximizing algorithm we set $c_i = 0$. The function V is then quadratic in a_i and b_i and the algorithm converges in one step giving the least squares or the Kalman estimate.[16] This is then taken as the starting point for the gradient routine. To investigate whether $V(\theta)$ has local minima we also choose several other starting points. FORTRAN programs for the identification procedure are available.[12]

It is shown in Ref. 12 that the maximum likelihood estimate is consistent, asymptotically normal, and efficient under mild conditions. The conditions are closely related to the information matrix. An estimate of this matrix is provided by

$$\hat{I} = \lambda^{-2} V_{\theta\theta}. \tag{12}$$

The matrix $V_{\theta\theta}$ which was computed in order to get a fast convergence for the gradient routine will thus have a physical interpretation.

- *Minimum variance control strategy*

We will now consider the control problem; i.e., we will find a control algorithm, expressing $u(t)$ as a function of the observed outputs up to time t, $y(t)$, $y(t-1)$, $\cdots$ and the previous control signals $u(t-1)$, $u(t-2)$, $\cdots$, such that the variance of the output $y(t)$ of the system (8) is as small as possible.

Theorem

The minimum variance control algorithm for the system (8) is given by

$$u(t) = -\frac{F(z^{-1})}{E(z^{-1})B(z^{-1})} y(t), \tag{13}$$

where $E(z)$ and $F(z)$ are polynomials

$$E(z) = 1 + e_1 z + e_2 z^2 + \cdots + e_{k-1} z^{k-1}$$

$$F(z) = f_0 + f_1 z + f_2 z^2 + \cdots + f_{n-1} z^{n-1},$$

which satisfy the identity

$$C(z) = A(z)E(z) + z^k F(z). \tag{14}$$

The control error with the optimal strategy is a moving average

$$\begin{aligned} y(t) &= \lambda E(z^{-1})e(t) \\ &= \lambda\{e(t) + e_1 e(t-1) + \cdots \\ &\qquad + e_{k-1}e(t-k+1)\} \end{aligned} \tag{15}$$

and the minimum variance of y is

$$\text{Min var } y(t) = \lambda^2\{1 + e_1^2 + e_2^2 + \cdots + e_{k-1}^2\}. \tag{16}$$

Proof

The system equation (8) gives

$$y(t) = A^{-1}(z^{-1})B(z^{-1})u(t-k) + \lambda A^{-1}(z^{-1})C(z^{-1})e(t).$$

Using the identity (14) we find

$$\begin{aligned} y(t) &= \lambda E(z^{-1})e(t) \\ &\quad + C^{-1}(z^{-1})z^{-k}\{F(z^{-1})y(t) \\ &\quad + B(z^{-1})E(z^{-1})u(t)\}. \end{aligned} \tag{17}$$

The first term of the right member of this equation is

$$\lambda E(z^{-1})e(t) = \lambda\{e(t) + e_1 e(t-1) + \cdots + e_{k-1}e(t-k+1)\},$$

and the last term of the equation is a function of

$$y(t-k),\quad y(t-k-1),\ \cdots,$$
$$u(t-k),\quad u(t-k-1),\ \cdots.$$

Now it was postulated that the control law must be such that $u(t)$ is a function of $y(t)$, $y(t-1)$, $\cdots$, $u(t-1)$, $u(t-2)$, $\cdots$. Because of the assumption of independence of $e(t)$ and $e(s)$ for $t \neq s$, the two terms of the right member of (17) are now independent and we get

$$\operatorname{var} y(t) \geq \operatorname{var}\{\lambda E(z^{-1})e(t)\}^2 = \lambda^2\{1 + e_1^2 + \cdots + e_{k-1}^2\}.$$

Equality is obtained for the control law (13) which proves the first part of the theorem. Introducing the control law (13) into (8) we find

$$y(t) = \lambda E(z^{-1})e(t),$$

and the second part of the theorem is also proven. *Q.E.D.*

Remark 1

Notice that the theorem still holds true if it is assumed only that $e(t)$ and $e(s)$ are uncorrelated for $t \neq s$ but a linear control law is postulated.

Remark 2

Notice that the last term of the right member of (17), i.e., the quantity

$$\hat{y}(t \mid t-k) = C^{-1}(z^{-1})\{F(z^{-1})y(t-k) + B(z^{-1})E(z^{-1})u(t-k)\},$$

can be interpreted as the k-step-ahead prediction of $y(t)$ based on $y(t-k)$, $y(t-k-1)$, $\cdots$ $u(t-k)$, $u(t-k-1)$, $\cdots$ and that

$$\tilde{y}(t \mid t-k) = \lambda E(z^{-1})e(t)$$

is the k-step-ahead prediction error. The equation (17) is thus a Wold decomposition. It generalizes the well-known formula for autoregressions to processes with rational spectral densities. The theorem thus states that the minimum variance equals the variance of the k step-ahead prediction error, and the optimal control law is obtained simply by requiring that the prediction of the output k steps ahead should equal the desired output.

Notice the crucial importance of the number k. Back-tracking we find that k physically corresponds to the sum of the transportation delay T_d and one sampling interval T_s. We thus find that what could possibly be achieved is limited by the transportation delay of the process T_d, the sampling interval T_s and the characteristics of the disturbance. Hence the minimal variance of the output equals the variance of the error when predicting the output over an interval which equals the sum of the transportation delay T_d and one sampling interval T_s. This result is important from a practical point of view because it gives the influence of the sampling interval on the results and can thus serve as a guideline for the choice of the sampling interval.

Remark 3

Notice that the control error (15) is a moving average of order k. The correlation function for the control error will thus vanish for lags greater than $k-1$. If k is known this fact can be exploited to test whether the system is optimally controlled simply by computing the correlation function for the control variable. The observation can be also used for on-line tuning of the control loops.

- *Sensitivity*

It is well-known that optimal solutions under special circumstances may be very sensitive to parameter variations. We shall therefore investigate this matter in our particular case. To do so we shall assume that the system is actually governed by the equation

$$A^{\circ}(z^{-1})y(t) = z^{-k}B^{\circ}(z^{-1})u(t) + \lambda^{\circ}C^{\circ}(z^{-1})e(t), \tag{18}$$

but that the control law is calculated under the assumption that the system model is

$$A(z^{-1})y(t) = z^{-k}B(z^{-1})u(t) + \lambda C(z^{-1})e(t), \tag{19}$$

where the coefficients of A, B, and C differ slightly from those of A°, B°, and C°.

Notice that the orders n of the models (18) and (19) are the same. The minimal variance control strategy for the model (19) is

$$u(t) = -\frac{F(z^{-1})}{B(z^{-1})E(z^{-1})}\,y(t), \tag{20}$$

where $E(z)$ and $F(z)$ are given by the identity (14).

We shall now investigate what happens if the system (18) is controlled with the control law (20). Introducing (20) into (18) we get

$$(A^{\circ}BE + z^{-k}FB^{\circ})y = \lambda^{\circ}C^{\circ}BEe, \tag{21}$$

where to facilitate writing, we have dropped the argument z^{-1} of A°, B, B°, C°, E, F and the argument t of y and e. Using the identity (14), Eq. (21) can now be written

$$\{B^{\circ}C + (A^{\circ}B - B^{\circ}A)E\}y = \lambda^{\circ}BC^{\circ}Ee. \tag{21$'$}$$

The dynamical system represented by this equation thus has a number of modes[17] equal to the degree of the polynomial

$$\{B^{\circ}C + (A^{\circ}B - B^{\circ}A)E\}(z). \tag{22}$$

Now if $A = A^\circ$, $B = B^\circ$ and $C = C^\circ$ this polynomial reduces to $B^\circ C^\circ$. For small perturbations in the parameters the modes of the system (21′) are thus close to the modes associated with $B^\circ C^\circ$. Furthermore when the design parameters equal the true parameters the factor $B^\circ C^\circ$ cancels in (21′). This fact implies that the modes associated with $B^\circ C^\circ$ are uncoupled to the input e if $A = A^\circ$, $B = B^\circ$ and $C = C^\circ$ or that the corresponding state-variables under the same conditions are not controllable from the input e (See Refs. 17, 18). Now if the control law is calculated from a model which deviates from the true model, the input might excite all the modes associated with the polynomial (22). This is not a serious matter if the modes are stable. However, if some modes are unstable it is possible to get infinitely large errors if the model used for designing the control law deviates from the actual model by an arbitrarily small amount. This situation will occur if the function

$$z^{2n} B^\circ(z^{-1}) C^\circ(z^{-1})$$

has zeros outside or on the unit circle. It follows from the representation theorems for stationary random processes[19] that $z^n C^\circ(z^{-1})$ can always be chosen to have zeros inside or on the unit circle. As far as C° is concerned the only critical case would be if $C^\circ(z^{-1})$ had a zero on the unit circle. The polynomial $z^n B^\circ(z^{-1})$ will have zeros outside the unit circle if the system dynamics is nonminimum phase. Hence, if either the dynamical system to be controlled is nonminimum phase or if the numerator of the spectral density of the disturbances has a zero on the unit circle, the minimum variance control law will be extremely sensitive to variations in the model parameters.

In these situations it is of great practical interest to derive control laws which are insensitive to parameter variations whose variances are close to the minimal variances. This can be done as follows.

To fix the ideas we will assume that $B(z)$ can be factored as

$$B(z) = B_1(z) B_2(z), \tag{23}$$

where $z^{n1} B_1(z^{-1})$ has all zeros inside the unit circle and $z^{n2} B_2(z^{-1})$ has all zeros outside the unit circle.

When resolving the identity (14) we impose the additional requirement that $F(z)$ contain $B_2(z)$ as a factor; i.e., we use the identity

$$C(z) = A(z) E'(z) + z^k B_2(z) F'(z). \tag{24}$$

Going through the arguments used when deriving the theorem we find the control law

$$u(t) = -\frac{F'(z^{-1})}{B_1(z^{-1}) E'(z^{-1})} y(t), \tag{25}$$

which gives the control error

$$\begin{aligned} y(t) = \lambda\{&e(t) + e_1 e(t-1) + \cdots \\ &+ e_{k-1} e(t-k+1) + e'_k e(t-k) + \cdots \\ &+ e'_{k+n_2-1} e(t-k-n_2+1)\}. \end{aligned} \tag{26}$$

The control law (25) thus gives an error with the variance

$$\operatorname{Var} y = \operatorname{Min} (\operatorname{Var} y) + \lambda^2 \{ e_k'^2 + \cdots + e'^2_{n+n_2-1} \}.$$

The control law (25) is not extremely sensitive to variations in system parameters. To realize this we assume again that the system is governed by the model $(A^\circ, B^\circ, C^\circ, \lambda^\circ)$ but that the control law is calculated from the model (A, B, C, λ) with slightly different parameters. The equation describing the controlled system then becomes

$$(A^\circ B_1 E' + z^{-k} B^\circ F') y = \lambda C^\circ B_1 E' e.$$

When the parameters equal the true parameters the characteristic equation of the system becomes

$$B_1^\circ(z^{-1}) C^\circ(z^{-1}) = 0,$$

and it now follows from the definition of B_1° and the assumption made on C° that all modes are stable when the design parameters equal the actual parameters. The stability for small perturbations of the parameters now follows by continuity.

- *Calculation of minimal variance control strategies*

The theorem stated earlier may be used conveniently for the actual calculation of minimum variance control strategies. Equating coefficients of different powers of z in Eq. (14) we get the following set of equations for determining the coefficients e_i and f_i:

$$\begin{aligned}
c_1 &= a_1 + e_1 \\
c_2 &= a_2 + a_1 e_1 + e_2 \\
&\vdots \\
c_{k-1} &= a_{k-1} + a_{k-2} e_1 + a_{k-3} e_2 + \cdots + a_1 e_{k-2} + e_{k-1} \\
c_k &= a_k + a_{k-1} e_1 + a_{k-2} e_2 + \cdots + a_1 e_{k-1} + f_0 \\
c_{k+1} &= a_{k+1} + a_k e_1 + a_{k-1} e_2 + \cdots + a_2 e_{k-1} + f_1 \\
&\vdots \\
c_n &= a_n + a_{n-1} e_1 + a_{n-2} e_2 + \cdots + a_{n-k+1} e_{k-1} + f_{n-k} \\
0 &= a_n e_1 + a_{n-1} e_2 + \cdots + a_{n-k+2} e_{k-1} + f_{n-k+1} \\
&\vdots \\
0 &= a_n e_{k-1} + f_n.
\end{aligned} \tag{27}$$

From these equations we can compute the coefficients of the E and F polynomials, $e_1, e_2, \cdots, e_{k-1}, f_0, f_1, \cdots, f_{n-1}$ recursively. Notice that these coefficients are independent of $B(z)$. When the polynomials $E(z)$ and $F(z)$ are known, the minimum variance control strategy is then given directly by (13).

We will now give some explicit examples of the calculation of control algorithms.

Example 1

Consider the following system

$$y(t) = \frac{24.9z^{-4}}{(1 - 0.55z^{-1} + 0.20z^{-2})(1 - z^{-1})} \nabla u(t) + \lambda \frac{1 - 0.77z^{-1} + 0.352z^{-2}}{(1 - 0.55z^{-1} + 0.20z^{-2})(1 - z^{-1})} e(t). \tag{28}$$

Hence,

$$A(z) = 1 - 1.55z + 0.75z^2 - 0.20z^3$$
$$B(z) = 24.9$$
$$C(z) = 1 - 0.77z + 0.352z^2.$$

Solving Eqs. (27) recursively we find

$$e_1 = 0.78 \qquad f_0 = 0.90$$
$$e_2 = 0.81 \qquad f_1 = -0.49$$
$$e_3 = 0.87 \qquad f_2 = 0.17.$$

The minimum variance control law thus becomes

$$\nabla u(t) = -\frac{0.91 - 0.49z^{-1} + 0.17z^{-2}}{24.9\{1 + 0.78z^{-1} + 0.81z^{-2} + 0.87z^{-3}\}} y(t) = -\frac{0.65}{24.9}(1.4)\frac{1 - 0.55z^{-1} + 0.19z^{-2}}{1 + 0.78z^{-1} + 0.81z^{-2} + 0.87z^{-3}} y(t)$$

where we have normalized the gain by the inverse of the low-frequency gain of the system (28).
The minimum variance is

$$\lambda^2\{1 + e_1^2 + e_2^2 + e_3^2\} = 3.02\lambda^2.$$

Example 2

Consider the system

$$y(t) = \frac{6.39 + 20.2z^{-1}}{(1 - 0.64z^{-1} + 0.22z^{-2})(1 - z^{-1})} \nabla u(t-3) + \lambda \frac{1 - 0.82z^{-1} + 0.21z^{-2}}{(1 - 0.64z^{-1} + 0.22z^{-2})(1 - z^{-1})} e(t). \tag{29}$$

Hence,

$$A(z) = 1 - 1.64z + 0.86z^2 - 0.22z^3$$
$$B(z) = 6.39 + 20.2z$$
$$C(z) = 1 - 0.82z + 0.21z^2.$$

In this case the polynomial $z\,B(z^{-1})$ has a zero outside the unit circle and the system is thus nonminimum phase. Proceeding as in Example 1, we find

$$e_1 = 0.82 \qquad f_0 = 0.64$$
$$e_2 = 0.69 \qquad f_1 = -0.41$$
$$f_2 = 0.15.$$

The minimum variance control law is thus

$$\nabla u(t) = -\frac{0.64 - 0.41z^{-1} + 0.15z^{-2}}{(6.39 + 20.2z^{-1})(1 + 0.82z^{-1} + 0.69z^{-2})} y(t) = -\frac{0.58}{26.6}(4.6) \times \frac{1 - 0.64z^{-1} + 0.24z^{-2}}{1 + 3.97z^{-1} + 3.26z^{-2} + 2.16z^{-3}} y(t), \tag{30}$$

where again we have normalized the gain by the inverse of the static gain of the system. Notice in particular the very high relative gain, 4.6, of the system.
The minimum variance is

$$\lambda^2\{1 + e_1^2 + e_2^2\} = 2.14\lambda^2.$$

Because the function $z\,B(z^{-1})$ has a zero, $z = -3.15$ outside the unit circle. It follows from the discussion on sensitivity in the previous section that the minimum variance control law is useless in practice because if the design parameters only vary slightly we will get an exponentially increasing error.

We will now show how to obtain a non-optimal control law which is less sensitive to parameter variations. Proceeding as in the section on sensitivity we now use the identity (24), i.e.,

$$1 - 0.82z + 0.21z^2 = (1 - 1.64z + 0.86z^2 - 0.22z^3) \times (1 + e_1z + e_2z + e_3'z^3) + z^3(f_0' + f_1'z + f_2'z^2)(1 + 3.15z).$$

Equating coefficients of equal powers of z of both members we get

$$e_1 = 0.82 \qquad f_0' = 0.16$$
$$e_2 = 0.69 \qquad f_1' = -0.096$$
$$e_3' = 0.49 \qquad f_2' = 0.035.$$

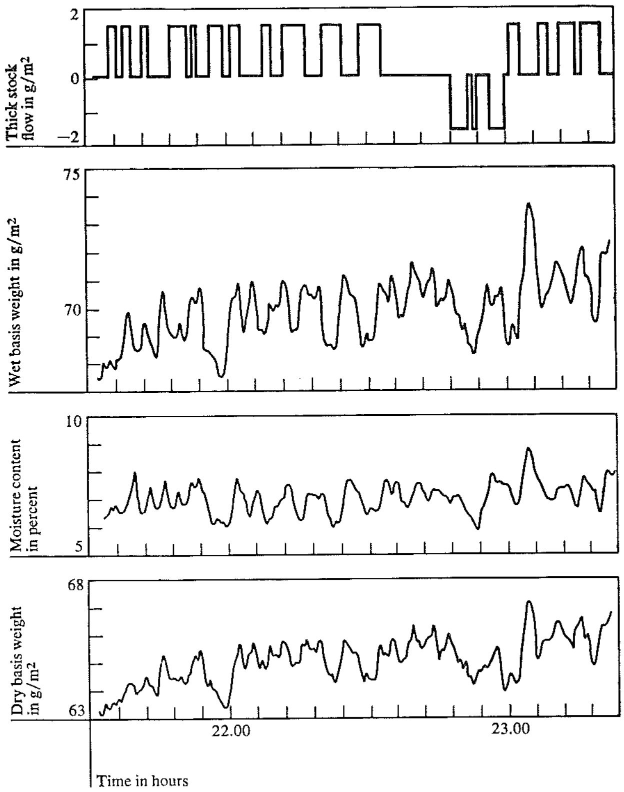


Figure 3 Results of an experiment for determination of process dynamics showing response of wet basis weight, moisture content and dry basis weight to perturbation in thick stock flow.

Notice that e_1 and e_2 are the same as before. We then obtain the control law (25), i.e.,

$$\nabla u(t)$$

$$= -\frac{0.16 - 0.096z^{-1} + 0.035z^{-2}}{6.39\{1 + 0.82z^{-1} + 0.69z^{-2} + 0.49z^{-3}\}}\, y(t)$$

$$= -\frac{0.58}{26.6}\,(0.64)$$

$$\times \frac{1 - 0.62z^{-1} + 0.22z^{-2}}{1 + 0.82z^{-1} + 0.69z^{-2} + 0.49z^{-3}}\, y(t). \quad (31)$$

Notice in particular the drastic decrease of the normalized gain as compared with (30). The variance associated with the control strategy (31) is

$$\lambda^2\{1 + e_1^2 + e_2^2 + e_3'^2\} = 2.38\lambda^2.$$

This should be compared with the minimum variance $2.14\lambda^2$ and the variance of the four-step-ahead predictor $2.55\lambda^2$. Hence, in this particular case we can find a control law which is insensitive to parameter variations at the cost of a 10% increase in the variance.

5. Practical experiments to determine process dynamics

Our identification procedure was described briefly in the previous section. The procedure has been applied extensively in connection with the Billerud project for quality control, basis weight control, moisture content control and refiner control. In this section we will present some of the practical results obtained. The examples are taken from basis weight regulation.

The control computer is used to perform the experiments. The input signal used in the experiment is represented as a sequence of numbers stored in the control computer. The numbers of the sequence are read periodically and converted to analog signals by the D/A converter and the regular D/A conversion subroutines. The output signals from the process are converted to digital numbers using the control computer's A/D converter. In this way we represent both the input and output signals by numbers that appear in the control computer in precisely the way they occur when the computer is controlling the process. The dynamics of signal transducers, transmission lines, and A/D and D/A converters are thus included in the model. Disturbances in transducers and signal converters, as well as round-off errors, are thus also included in the disturbances of the model. The whole experiment is executed by a program. The result of a typical identification experiment is illustrated in Fig. 3.

- *Choosing the input signal*

The choice of the input signal involves certain considerations. That is, it is desirable to have large signal amplitudes in order to get good estimates. Large input signals may, however, drive the system outside the linear region and may also cause unacceptably large variations in the process variables. In our particular case we had to make all experiments during normal production. This was one major reason for using a fairly sophisticated identification procedure. Notice that in order to obtain a specified accuracy it is possible to compromise between signal amplitude and length of the sample. In the identification of models required for the design of basis weight control laws we usually used samples 1 to 5 hours in length. The amplitudes of the signals shown in Fig. 3 corresponding to 1.7 g/m² peak-to-peak are typical. This number was a suitable compromise. Notice that the standard deviation during normal operation with no control is typically 1.3 g/m².

The input signal must also be chosen so that it is persistently exciting.[12,13] This is always the case if the input signal has constant spectral density. Pseudo-random binary signals have been used successfully. We have found, however, that if some knowledge of the process is available, it is desirable to tailor the test signals to the specific purpose.

Table 1 Successive parameter iterates for a first-order model relating dry basis weight to thick stock flow; $k = 4$, $N = 101$.

Step	a_1	b_0	c_1	V	$\frac{\partial V}{\partial a} \times 10^5$	$\frac{\partial V}{\partial b} \times 10^5$	$\frac{\partial V}{\partial c} \times 10^5$
0	0	0	0	6.7350	91683	39509	−91683
1	−0.0122	13.0054	0	4.1603	0	0	193777
2	−0.3924	13.9356	−0.6320	3.3764	−78727	1190	51707
3	−0.3492	14.6689	−0.6542	3.3360	1339	−69	2575
4	−0.3502	14.6468	−0.6572	3.3360	106	−3	−165
5	−0.3500	14.6468	−0.6569				

• *Examples of numerical identification*

We will now present some examples which illustrate the numerical identification procedure. These examples are based on the data shown in Fig. 3. Mathematical models which relate changes in dry basis weight (WSPO) and wet basis weight (WSP) to changes in (the set point of the) thick stock flow (regulator) will be discussed. Figure 3 shows that the output is drifting. The drift is even more pronounced in test experiments of longer duration. To take care of this drift we have used models which relate changes in the output to changes in the input; i.e.,

$$\nabla y(t) = \frac{b_0 + b_1 z^{-1} + \cdots + b_{n-1} z^{-n+1}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} \nabla u(t-k) + \lambda \frac{1 + c_1 z^{-1} + \cdots + c_n z^{-n}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} e(t), \tag{32}$$

where ∇ is the backward shift operator:

$$\nabla\, y(t) = y(t) - y(t-1).$$

Rewriting the equation we find

$$y(t) = \frac{b_0 + b_1 z^{-1} + \cdots + b_{n-1} z^{-n+1}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} u(t-k) + \lambda \frac{1 + c_1 z^{-1} + \cdots + c_n z^{-n}}{(1 - z^{-1})(1 + a_1 z^{-1} + \cdots + a_n z^{-n})} e(t). \tag{33}$$

The time interval in all cases has been 0.01 hour. All examples are based on data of Fig. 3 in the time interval 21.53 to 22.58 hours.

As stated previously, the identification procedure is carried out recursively starting with a first-order system, continuing with a second order system, etc. To obtain the value of k for a fixed order, the identification is also repeated with the input signal shifted.

Example 1—Model relating dry basis weight to thick stock flow

The first numerical example will be a model relating dry basis weight to thick stock flow. First we shall identify a first order model having the structure (32). Applying our numerical identification algorithm we get the results shown in Table 1.

Starting with the initial parameter estimate $\theta = 0$, the first step of the identification algorithm gives the Kalman estimate[12,16] of the parameters and this estimate is then successively improved until the loss function $V(\theta)$ of Eq. (32) is minimized and the maximum likelihood estimate obtained. Notice in particular the significant difference between the Kalman estimate (step 1) and the maximum likelihood estimate.

The value of the matrix of second partial derivatives at the last step of the iteration is

$$V_{\theta\theta} = \begin{bmatrix} 19.28 & -0.29 & -8.86 \\ -0.29 & 0.04 & 0.06 \\ -8.86 & 0.06 & 12.05 \end{bmatrix}.$$

Repeating the identification for different values of the time-delay k we obtain the results given in Table 2.

Table 2 Results of identification of first-order models relating dry basis weight to thick stock flow for different time-delays.

k	a_1	b_0	c_1	λ	V
3	−0.807	9.846	−0.994	0.297	4.491
4	−0.350	14.647	−0.657	0.257	3.336
5	−0.749	1.286	−0.958	0.351	6.152

We thus find that the loss function V has its smallest value for $k = 4$. To find the accuracy of the model parameters we proceed as follows:

An estimate of Fisher's information matrix is obtained from the matrix of second partial derivatives (Ref. 13, Lemma 2),

$$\hat{I} = \lambda^{-2} V_{\theta\theta}.$$

It is further shown in Ref. 13 (Theorem 4) that if $V_{\theta\theta}$ is nonsingular the estimate is asymptotically normal (θ_0, I^{-1})

and we thus have the following estimate of the covariance of the asymptotic distribution:

$$\hat{I}^{-1} = \lambda^2 V_{\theta\theta}^{-1} = \begin{bmatrix} 0.006 & 0.042 & 0.004 \\ 0.042 & 2.202 & 0.020 \\ 0.004 & 0.020 & 0.008 \end{bmatrix}.$$

Summarizing, we thus find the following numerical values for the best first-order model; where the computations are based on 100 pairs of input-output data:

$k = 4$ $\qquad$ $c = -0.66 \pm 0.09$

$a = -0.35 \pm 0.08$ $\qquad$ $\lambda = 0.257 \pm 0.017$

$b = 14.6 \pm 1.5$ $\qquad$ $V = 3.34 \pm 0.44.$

Proceeding to a second-order model the identification algorithm gives the following results, based, again, on 100 pairs of input-output data:

$k = 3$ $\qquad$ $c_1 = -0.73 \pm 0.18$

$a_1 = -0.46 \pm 0.14$ $\qquad$ $c_2 = 0.12 \pm 0.16$

$a_2 = 0.04 \pm 0.12$ $\qquad$ $\lambda = 0.249 \pm 0.017$

$b_0 = 3.4 \pm 1.6$ $\qquad$ $V = 3.15 \pm 0.43$

$b_1 = 12.3 \pm 2.2$

The matrix of second partial derivatives at the minimum is:

$$V_{\theta\theta} = \begin{bmatrix} 22.47 & 13.82 & -0.08 & 0.36 & -7.61 & -1.87 \\ 13.82 & 22.47 & -0.17 & -0.08 & -4.94 & -7.59 \\ -0.08 & -0.17 & 0.04 & 0.02 & 0.05 & -0.05 \\ 0.36 & -0.08 & 0.02 & 0.04 & 0.06 & 0.06 \\ -7.61 & -4.94 & 0.05 & 0.06 & 11.06 & 6.60 \\ -1.87 & -7.59 & -0.05 & 0.06 & 6.60 & 10.56 \end{bmatrix}.$$

It now follows from Ref. 13 (Theorem 4) that the parameter estimates for a large number of input-output pairs is asymptotically normal $N(\theta_0, \lambda^2 V_{\theta\theta}^{-1})$. Assuming that asymptotic theory can be applied we can now solve various statistical problems. We will, for example, test the hypothesis that the model is of first order; i.e., our null hypothesis is

$$H_0 : (a_2^\circ = b_\circ^\circ = c_2^\circ = 0).$$

Using the asymptotic theory we find that the statistic

$$\xi = \frac{V_2 - V_1}{V_2} \cdot \frac{N - 6}{3}$$

has an $F(3, N - 6)$ distribution under the null hypothesis. The symbol V_2 denotes the minimal value of the loss function for the second-order model; V_1, the minimal value

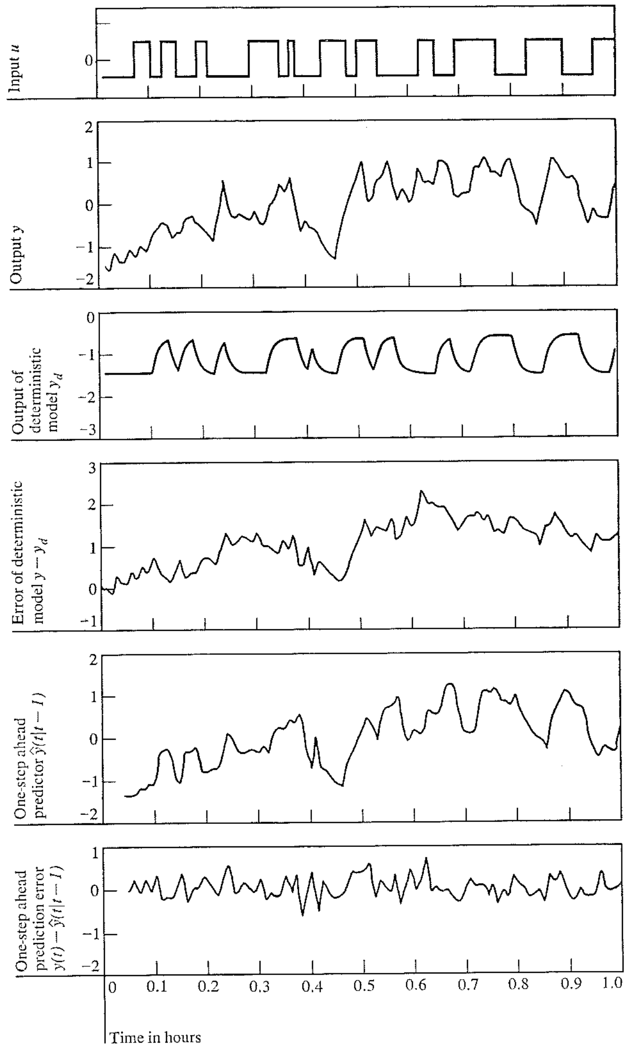


Figure 4 Illustration of the results of the identification of first-order model relating dry basis weight to thick stock flow. Notice in particular the relative magnitudes of the output of the deterministic model, the error of the deterministic model and the error of the one-step ahead predictor. Also notice the trend in the error of the deterministic model.

for the first-order model and N, the number of input-output pairs. In this particular case we have $\xi = 1.9$. At a risk level of 10% we have $F(3, 96) = 2.7$ and the null-hypothesis, that the system is of first order, thus has to be accepted.

The results of the identification procedure are illustrated in Fig. 4. In this figure we show

- the input u
- the output y
- the deterministic output y_d defined by

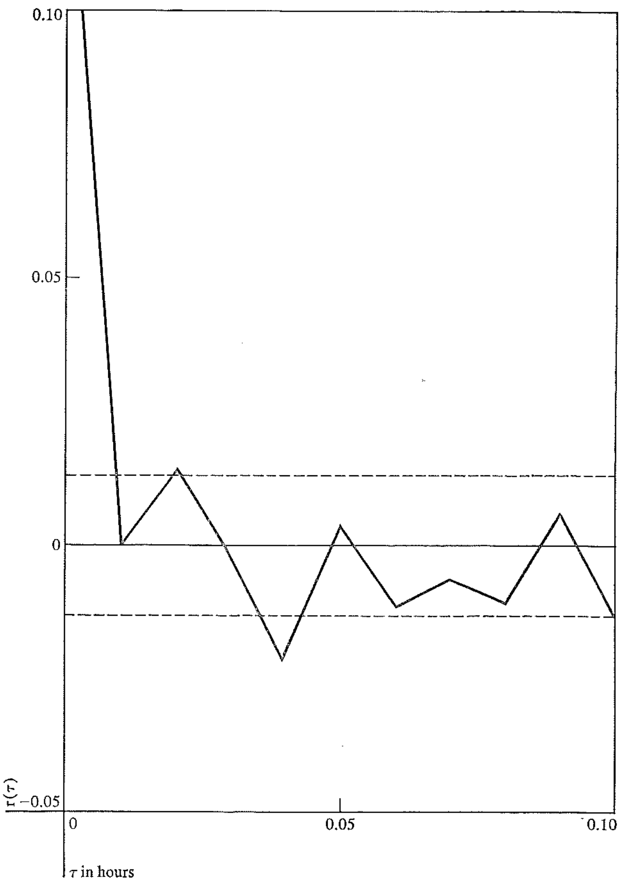
$$y_d(t) = \frac{b_0 + b_1 z^{-1} + \cdots + b_{n-1} z^{-n+1}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} u(t)$$

- the error of the deterministic model
  $e_d(t) = y(t) - y_d(t)$
- the one-step-ahead predictor $\hat{y}(t|t-1)$ of $y(t)$
- the one-step-ahead predictor error $y(t) - \hat{y}(t|t-1)$.

Figure 4 illustrates the properties of the identification procedure. The deterministic output $y_d(t)$ shows how much of the output $y(t)$ can be explained by the input $u(t)$. The error $e_d(t)$ thus represents the part of the output that is caused by the disturbances. Notice in particular the drifting character of the error e_d. The one-step ahead predictor illustrates how well the output can be predicted one-step ahead. Recall that the model was in fact constructed so as to minimize the sum of squares of the one-step ahead prediction error.

The identification procedure was based on the assumption that the residuals were normal and uncorrelated. Having performed the identification and calculated the one-step ahead prediction errors, $\hat{e}(t)$, we thus have the possibility of checking this assumption. In Fig. 5 we show

Figure 5 Sample covariance function for the residuals $\epsilon(t)$ of the first-order model for basis weight. According to the assumptions made in the identification theory[12, 13] $r(\tau)$ should equal zero when $\tau \neq 0$. The dashed line gives the one sigma limit for $r(\tau)$, $\tau \neq 0$.



the correlation function of the one-step ahead prediction errors.

Example 2—Model relating wet basis weight to thick stock flow

As our second illustration of the numerical identification procedure we will now use the data of Fig. 3 to find a model relating wet basis weight to thick stock flow. In this case we find that the minimum value of the loss function for the first-order case occurs at $k = 4$ and the coefficients of the best first-order model are:

$$k = 4 \qquad c_1 = -0.62 \pm 0.10$$
$$a_1 = -0.38 \pm 0.05 \qquad \lambda = 0.364 \pm 0.025$$
$$b_1 = 27.1 \pm 2.1 \qquad V = 6.60 \pm 0.94.$$

Similarly, the best second-order model is given by the coefficients:

$$k = 3 \qquad c_1 = -0.82 \pm 0.14$$
$$a_1 = -0.64 \pm 0.11 \qquad c_2 = 0.21 \pm 0.14$$
$$a_2 = 0.22 \pm 0.09 \qquad \lambda = 0.335 \pm 0.024$$
$$b_0 = 6.4 \pm 2.0 \qquad V = 5.73 \pm 0.80.$$
$$b_1 = 20.2 \pm 3.0$$

The matrix of second-order partial derivatives of the minimal point is

$$V_{\theta\theta} = \begin{bmatrix} 79.24 & 53.37 & -0.13 & 0.76 & -12.68 & -0.13 \\ 53.37 & 79.12 & -0.40 & -0.13 & -5.93 & -11.44 \\ -0.13 & -0.40 & 0.04 & 0.02 & 0.06 & -0.07 \\ 0.76 & -0.13 & 0.02 & 0.04 & 0.10 & 0.10 \\ -12.68 & -5.93 & 0.06 & 0.10 & 17.64 & 7.83 \\ -0.13 & -11.44 & -0.07 & 0.10 & 7.83 & 15.12 \end{bmatrix}.$$

We now test the null hypothesis that the system is of first order; i.e.,

$$H : (a_2^\circ = b_o^\circ = c_2^\circ = 0)$$

Using the asymptotic results, we find $\xi = 4.8$ and the hypothesis thus has to be rejected. Increasing the order to three does not give any significant improvements in the loss function.

Hence if we consider dry basis weight as the output of the system, we find that the model is of first order, but if we consider wet basis weight as the output, the model is of second order. This also shows up very clearly in Fig. 6 where we illustrate the results of the identification of the models for wet basis weight. There is a physical explanation for this difference in behaviour. As mentioned previously

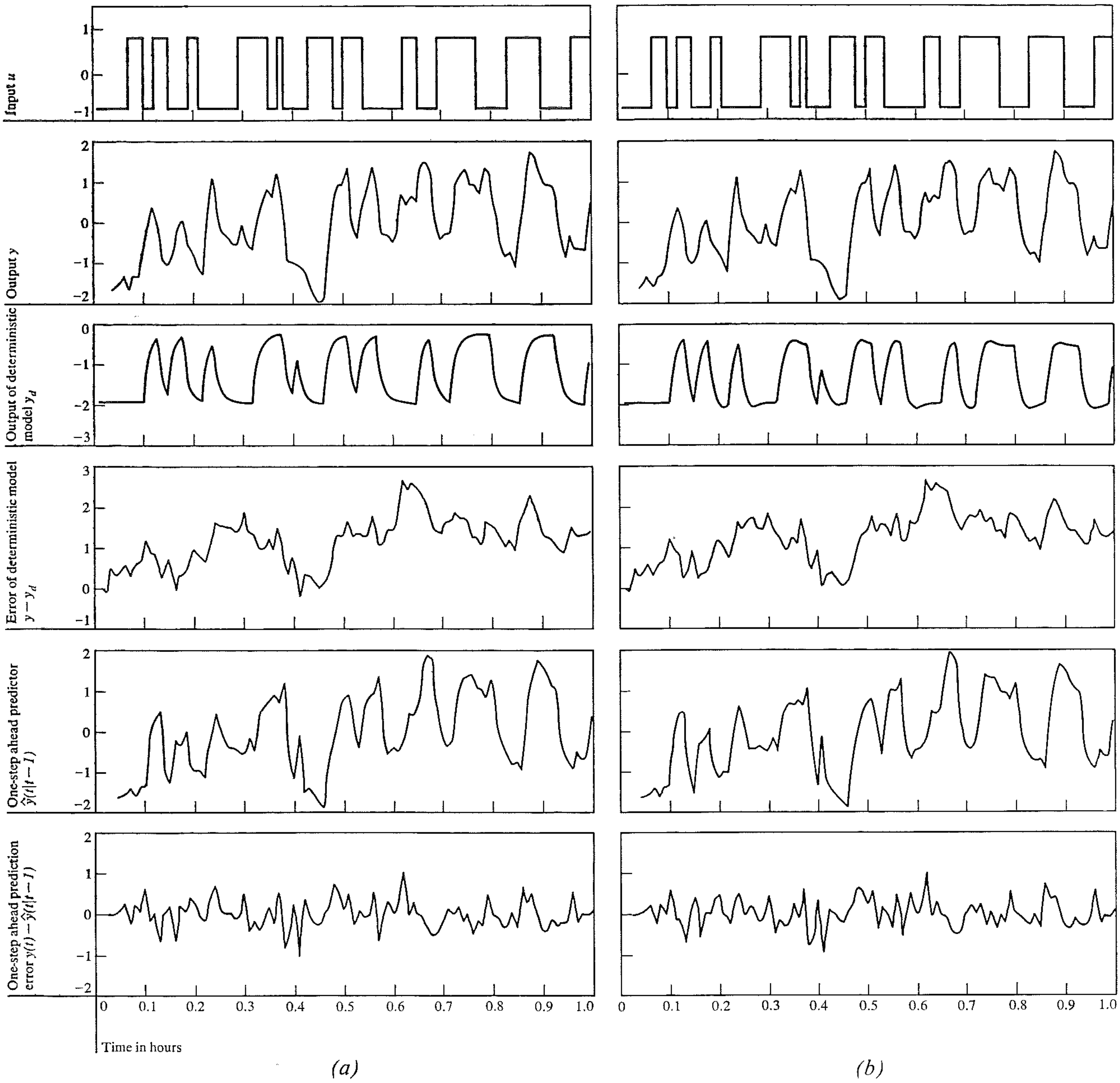


Figure 6 Illustration of the results of the identification of models for wet basis weight. A first-order model is shown in (a) and a second-order model in (b). Notice in particular the differences between the outputs of the deterministic models for first- and second-order systems.

and as can be seen from Fig. 3, a change in thick stock flow will influence dry basis weight as well as moisture content. After an increase in thick stock flow, we find that both dry basis weight and moisture content will increase. The increase in moisture content will then be eliminated by the moisture control feedback loop which controls the set point of the fourth drying section by feedback from the moisture gauge.

These two effects will explain the overshoot in the response of the wet basis weight. It is also clear from this discussion that the response of the wet basis weight will be influenced by the settings of the moisture control loop. This fact is another argument for using dry basis weight as the control variable, when the basis weight loop is considered as a single-input, single-output system.

6. Practical experiences with on-line basis weight control

We shall now summarize some of the practical results achieved with on-line basis weight control. The experi-

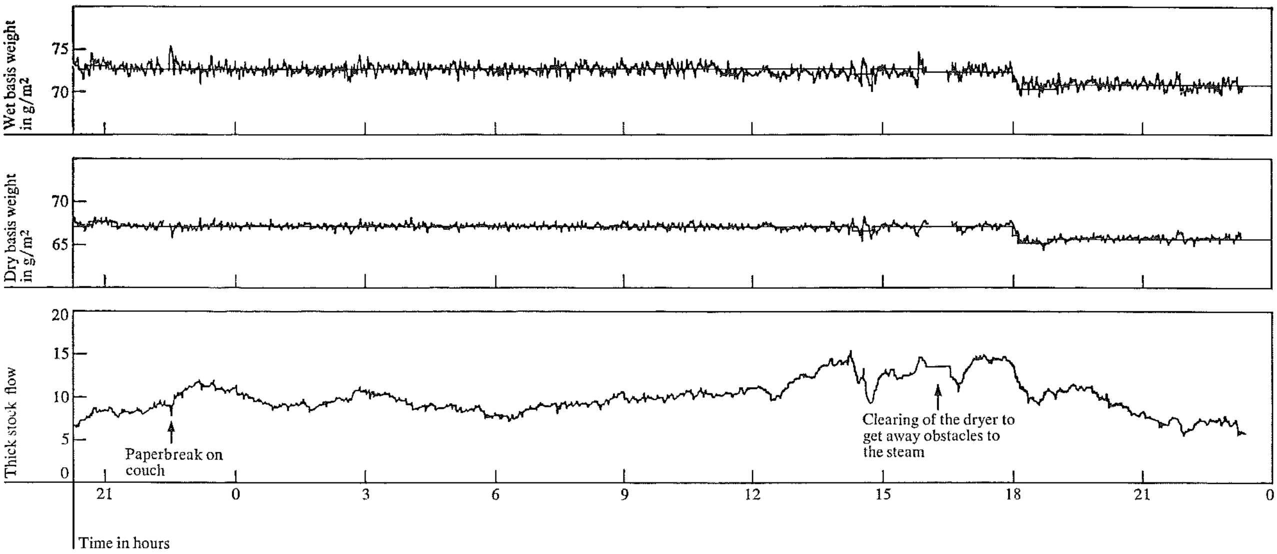


Figure 7 Results of test run with on-line control of basis weight.

mental program that was carried out had a dual purpose: to arrive at control strategies for the particular application at hand and to test the general procedure developed for steady-state control. This dual purpose led us to continue some experiments, even though these particular loops have been working satisfactorily. Several control schemes have been investigated. We have chosen thick stock flow as well as thick stock concentration as control variables. We have regulated both wet and dry basis weight. In the first experiments the concentration of the thick stock was chosen as the control variable. This was later changed to thick stock flow mainly for two reasons. We found that the basis weight responds faster to changes in the set point of the thick stock flow regulator than to changes in the thick stock concentration. We also found that the dynamics of the concentration regulator changed with operating conditions, thereby introducing variations in the dynamics of the control loop.

In general it is very difficult to evaluate the performance of the control loops in practice and in particular to compare different control laws. The main reason for this is that there are variations in the disturbance level. This implies that in order to evaluate the different control loops we need test periods of considerable length.

It is also very difficult to judge the improvements unless reference values are available. In the case of basis weight we had the results of the feasibility study. In all cases studied before the control computer was installed, standard deviation of basis weight was greater or equal to 1.3 g/m^2 and this value was therefore chosen as a conservative reference value. In the feasibility study the target value for basis weight fluctuations was set to 0.7 g/m^2. In actual operation we can now consistently achieve standard deviations of 0.5 g/m^2 wet basis weight and 0.3 g/m^2 dry basis weight.

Basis weight was controlled sucessfully on-line on April 28, 1965 for a test period of 10 hours. The first experiments showed that it was indeed possible to obtain the variances predicted from the results of the process identification. We could also show that the deviations for the controlled system were moving averages of the appropriate order. The basis weight control loop has been subject to extensive investigations and has been in continuous operation since the beginning of 1966.

Two types of experiments have been performed. In one, the control loop is permitted to operate in the normal way for several weeks. Some data are collected at comparatively long sampling intervals (0.1 hour). The results are not analyzed extensively and the performance of the control system is evaluated on the basis of the maximum deviations of test laboratory data, inspection of strip-chart recorders, and the judgement of machine tenders.

The other type of experiment is a controlled experiment extending over periods of 30 to 100 hours. Important process variables are logged at a sampling interval of 0.01 hour and analyzed. When analyzing the data, we compute covariance functions of the controlled variables and test whether they are moving averages of appropriate order (cf. Remark 3 of the Theorem of Section 3). Variances are checked against reference values. In some cases we also identify dynamic models, calculate minimum variance control strategies and update the parameters of the control algorithms if required.

In Fig. 7 we give a sample covering 24 hours of operation of the basis weight control loop. In the diagram we show wet basis weight, dry basis weight (the controlled output)

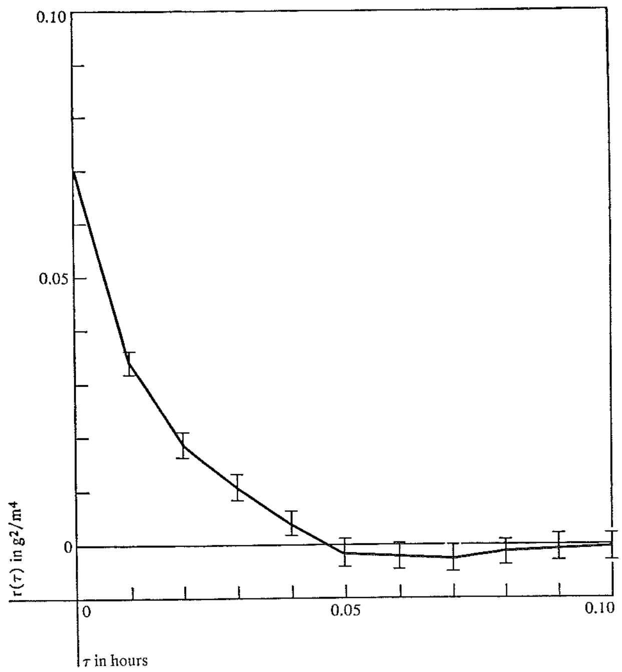


Figure 8 Covariance function for fluctuations of dry basis weight in the time interval 23.00 to 12.00 of Fig. 7. Compare with Fig. 2 which shows the covariance function without computer control.

and thick stock flow (the control signal). The scale for the control signal, thick stock flow, is chosen as dry basis weight. The magnitude of the control signal will thus directly indicate how much of the fluctuations in dry basis weight are removed by the control law. The control signal will thus approximately show the disturbances in the output of the system. Notice the different characteristics of the disturbances at different times. The large disturbances occurring at times 14.30 and 18.00 are due to large fluctuations in thick stock consistency.

Also notice that there are two interrupts in the operation of the system, one paper break and one interrupt to clear the drying section. In these instances the basis weight control loop is automatically switched off and the control signal is kept constant until the disturbances are cured when the loops are automatically switched on again. Notice that a paper break does not introduce any serious disturbances. Also notice that there are some grade changes from which we can judge the response of the controlled system to step changes in the references values.

Moisture content was controlled by feedback from the moisture meter to the set point of the pressure regulator of the fourth drying section. The standard deviation of moisture content was 0.4%. In Fig. 8 we show the covariance function of dry basis weight in the time interval. As is to be expected from the Theorem of Section 3 this is the covariance function of a moving average of fourth order.

We have also made experiments to verify that the high frequency fluctuations in moisture content and basis weight have the same characteristics. This was one essential assumption made in Section 3. If this was true, the variance in dry basis weight would be independent of dry or wet basis weight control. In the table below we give standard deviations recorded during a 30 hour test, where alternatively wet and dry basis weight was controlled.

	Standard deviation	
	Wet basis weight	*Dry basis weight*
Wet basis weight controlled	0.50	0.32
Dry basis weight controlled	0.52	0.28

Acknowledgment

The author would like to express his gratitude to several of his colleagues for valuable comments and stimulating discussions, in particular to A. Ekström and to R. W. Koepcke. The identification procedure was developed jointly with T. Bohlin. S. Wensmark carried out a large amount of the programming required for this work. Special thanks are also due to N. Vogt of the Billerud Company who participated in the first experiments with on-line control.

References

1. K. J. Åström, "Control Problems in Papermaking," *Proceedings of the IBM Scientific Computing Symposium on Control Theory and Applications*, October 19–21, 1964.
2. A. Ekström, "Integrated Computer Control of a Paper, Machine—System Summary," *Proceedings of the Billerud-IBM Symposium* (June, 1966).
3. K. J. Åström, and T. Bohlin, "Integrated Computer Control of a Kraft Paper Mill—New Methods for Control Strategy Design in Operation," *Proceedings of the Billerud-IBM Symposium*, (June 1966).
4. T. Bohlin, "Paper machine identification for purposes of computer control," *International Congress on Automation and Instrumentation in the Paper, Rubber and Plastics Industries*, Antwerp (October, 1966).
5. K. J. Åström, R. W. Koepcke and F. Tung, "On the Control of Linear Discrete Dynamic Systems with Quadratic Loss," Report RJ–222, IBM Research Laboratory, San Jose, Cal. (1962).
6. G. E. P. Box and G. M. Jenkins, "Some Statistical Aspects of Adaptive Optimization and Control," *Journ. Royal Statist. Soc., B* 24, 297–343 (1962).
7. G. E. P. Box and G. M. Jenkins, "Further Contributions to Adaptive Quality Control: Simultaneous Estimation of Dynamics: Non-Zero Costs," Technical Report No. 19, Dept. of Statistics, University of Wisconsin (1963).
8. J. J. Florentin, "Partial Observability and Optimal Control," *Journ. Electronics Control* 13, 263–279 (1962).
9. T. L. Gunckel and G. F. Franklin, "A General Solution fcr Linear Sampled Data Control," *Trans. ASME Ser. D (J. Basic Eng.)* 85, 197–203 (1963).

10. P. D. Joseph and J. T. Tou, "On Linear Control Theory," *AIEE Trans. Applications in Industry*, 193–196 (1961).
11. R. E. Kalman, "New Methods and Results in Linear Prediction and Filtering Theory," *Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability*, ed. J. L. Bogdanoff and F. Kozin, New York, Wiley, 1963.
12. K. A. Åström, T. Bohlin and S. Wensmark, "Automatic Construction of Linear Stochastic Dynamic Models for Stationary Industrial Processes with Random Disturbances using Operating Records," Technical Paper TP 18.150, IBM Nordic Laboratory (1965).
13. K. J. Åström and T. Bohlin, "Numerical Identification of Linear Dynamic Systems from Normal Operating Records," *Proceedings of the IFAC Conference on Self-Adaptive Control Systems*, Teddington, (1965).
14. K. J. Åström, "Notes on the Regulation Problem," IBM Nordic Laboratory (1965).
15. J. A. Smethurst, "Direct Computer Control of Production Changes on a Simulated Quality Paper Machine," *Proc. IFAC Congress*, London (1966).
16. R. E. Kalman, "Design of a Self-Optimizing Control System," *Trans. ASME Ser. D* (*J. Basic Eng.*) 80, 468–478 (1958).
17. L. A. Zadeh and C. A. Desoer, "Linear System Theory, The State Space Approach," New York, McGraw–Hill, 1963.
18. R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," *J. SIAM Control, Ser. A* 1, 152–192 (1963).
19. J. L. Doob, "Stochastic Processes," New York, Wiley, 1953.
20. A. E. Beecker, "Dynamic Modeling Techniques in the Paper Industry," *TAPPI* 46, 117–120 (1963).

Received October 7, 1966