Estimation of Cross-Directional Properties: Scanning vs. Stationary Sensors

Matthew L. Tyler and Manfred Morari
Chemical Engineering, 210-41, California Institute of Technology, Pasadena, CA 91125

Periodic time-varying Kalman filter equations for problems involving scanning sensors are solved using "lifting" techniques common for multirate systems. The solution to this problem is used to compare the performance of scanning sensors vs. stationary sensors in the estimation of cross-directional properties. Furthermore, controller performance is examined when the outputs from the Kalman filter are used as inputs to a state feedback control law. Although adding sensors may significantly enhance the estimates of cross-directional properties, feedback of these improved estimates may translate to lower levels of improvement in cross-directional variations.

Introduction

Recent attention has been focused on cross-directional control of processes such as paper manufacturing and coating (McFarlin, 1983; Boyle, 1978; Richards, 1982; Wilhelm and Fjeld, 1983; Bergh and MacGregor, 1987; Laughlin, 1988; Laughlin et al., 1993; Braatz et al., 1992). The objective of these control strategies is to maintain some property such as basis weight or coating thickness uniform across a sheet of paper. Several control strategies for this problem have been reported in the literature. These strategies rely on a measurement of the property across the cross direction. Process measurements, however, are typically made by scanning sensors. These sensors move back and forth across the paper sheet as the paper moves in the machine direction. Thus, the cross-directional variations are not measured directly. In many applications, it is sufficient to assume that the fluctuations in the cross direction occur on a much slower time scale than in the machine direction. Under these assumptions and using physically motivated models, estimation methods for the cross-directional moisture content in paper have been reported by Wang et al. (1993) and Dumont et al. (1993). This article considers the problem of cross-directional estimation when an input-output model is employed and the dynamics of variations in the cross direction cannot be neglected. We show how results from optimal control of multirate systems can be applied to solve the optimal estimation problem of a scanning sensor. Using this solution, we develop guidelines for determining when adding sensors will improve the estimation and control of cross-directional properties. We will show that this decision depends on parameters that affect the estimation problem, such as the system dimension, the dominant time constant, the sampling rate, the ratio of process noise to measurement noise, and correlation between process disturbances, as well as parameters that affect control performance, such as process delays and robustness considerations.

Throughout this article, we assume the reader has a certain familiarity with stochastic state estimation and Kalman filtering. A detailed treatment of this subject can be found in several textbooks (for example, Kwakernaak and Sivan, 1972).

Model

Consider a discrete time, linear time-invariant (LTI) process described by an input-output model with the following state space representation:

$$x[k+1] = Ax[k] + Bu[k] + Gw[k],$$
$$\hat{y}[k] = Cx[k] + \hat{v}[k], \tag{1}$$

where $x \in \Re^l$, $u \in \Re^m$, y, $\hat{v} \in \Re^n$, $w \in \Re^p$, and A, B, C, and G are constant matrices of the appropriate dimensions. Here we interpret the variable u as a known signal such as control input, whereas w and $\hat{v}$ are stochastic variables, representing process and measurement noise, respectively, whose distributions are described by

$$E\left\{\begin{pmatrix} w \\ \hat{v} \end{pmatrix}\begin{pmatrix} w^T & \hat{v}^T \end{pmatrix}\right\} = \begin{bmatrix} Q & 0 \\ 0 & \hat{R} \end{bmatrix}, \tag{2}$$

Correspondence concerning this article should be addressed to M. Morari.
Research by M. L. Tyler was supported by Fanny and John Hertz Foundation.

where E denotes the expectation operator. The variable $\hat{y}$ denotes some property of interest at specified locations along the cross direction. We model the action of the scanning sensor as a linear, periodic time-varying (PTV) operator $S[k]: \mathbb{R}^n \rightarrow \mathbb{R}^s$, with period N. When combined with the above system, the scanning sensor gives the following PTV system:

$$x[k+1] = Ax[k] + Bu[k] + Gw[k],$$
$$y[k] = C[k]x[k] + v[k], \qquad (3)$$

with

$$y[k] = S[k]\hat{y}[k],$$
$$v[k] = S[k]\hat{v}[k],$$
$$R[k] = S[k]^T RS[k] = E\{v[k]v[k]^T\},$$
$$C[k] = S[k]C. \qquad (4)$$

We now consider the problem of calculating an estimate $\hat{x}$ of the system state x from the measured output y which is optimal in the sense that $\|x - \hat{x}\|_2$ is minimized. The solution to this problem is well known and is given by a time-varying Kalman filter (Kwakernaak and Sivan, 1972). The state estimate $\hat{x}$ is calculated in a two-step procedure:

$$\hat{x}[k+1|k] = A\hat{x}[k|k] + Bu[k],$$
$$\hat{x}[k|k] = \hat{x}[k|k-1] + K[k](y[k] - C[k]\hat{x}[k|k-1]), \qquad (5)$$

where $\hat{x}[i|j]$ denotes the estimate of $x[i]$ given measurements y up to and including time j. The filter gain $K[k]$ is given by

$$K[k] = \Sigma[k]C^T[k](C[k]\Sigma[k]C^T[k] + R[k])^{-1}, \qquad (6)$$

where $\Sigma[k]$ is the covariance of $\hat{x}[k|k-1] - x[k]$ and may be calculated as the solution to the following Riccati equation

$$\Sigma[k+1] = A\Sigma[k]A^T - A\Sigma[k]C^T[k](C[k]\Sigma[k]C^T[k] + R[k])^{-1}C[k]\Sigma[k]A^T + GQG^T. \qquad (7)$$

Because of the periodicity of $C[k]$ and $R[k]$, the solution to Eq. 7 will be periodic as well. The solution may be obtained by iterating on Eq. 7 until it converges, as suggested by Bergh and MacGregor (1987). However, a more efficient solution technique follows from using methods developed for multirate systems (Lee et al., 1991; Amit, 1980; Lee and Morari, 1991). These methods and their application to this problem will be discussed in the next section.

Solving PTV Riccati Equation via Lifting

For multirate systems, Amit (1980) established the following approach to solving PTV Kalman filter equations. Consider a PTV system. By viewing the output to consist of all the measurements made during one period, the system can be "lifted" to form an LTI system. The state space equations for the lifted system can be obtained by augmenting the input vector to include all inputs during one period. For example, the lifted version of the system (Eq. 3) would be given by

$$X[j+1] = A_l X[j] + B_l U[j] + G_l W[j],$$
$$Y[j] = C_l X[j] + D_u U[j] + D_w W[j] + V[j], \qquad (8)$$

where

$$X[j] = x[jN], \quad A_l = A^N, \quad B_l = [A^{N-1}B, A^{N-2}B, \ldots, B],$$
$$G_l = [A^{N-1}G, A^{N-2}G, \ldots, G],$$

$$Y[j] = \begin{bmatrix} y[jN] \\ y[jN+1] \\ \vdots \\ y[jN+N-1] \end{bmatrix}, \quad C_l = \begin{bmatrix} C[0] \\ C[1]A \\ \vdots \\ C[N-1]A^{N-1} \end{bmatrix},$$

$$U[j] = \begin{bmatrix} u[jN] \\ u[jN+1] \\ \vdots \\ u[jN+N-1] \end{bmatrix}, \quad W[j] = \begin{bmatrix} w[jN] \\ w[jN+1] \\ \vdots \\ w[jN+N-1] \end{bmatrix},$$

$$V[j] = \begin{bmatrix} v[jN] \\ v[jN+1] \\ \vdots \\ v[jN+N-1] \end{bmatrix},$$

$$D_u = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ C[1]B & 0 & 0 & \cdots & 0 \\ C[2]AB & C[2]B & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ C[N-1]A^{N-2}B & C[N-1]A^{N-3}B & \cdots & C[N-1]B & 0 \end{bmatrix},$$

$$D_w = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ C[1]G & 0 & 0 & \cdots & 0 \\ C[2]AG & C[2]G & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ C[N-1]A^{N-2}G & C[N-1]A^{N-3}G & \cdots & C[N-1]G & 0 \end{bmatrix}.$$

Making the substitution $V_l = D_w W + V$, we consider the system

$$X[j+1] = A_l X[j] + B_l U[j] + G_l W[j],$$
$$Y[j] = C_l X[j] + D_u U[j] + V_l[j]. \quad (9)$$

The states of the lifted system evolve as the Nth iterate of the original states, and the lifted system has Nm inputs, Ns outputs, a process disturbance of dimension Np, and a measurement disturbance of dimension Ns. The covariance of the stochastic variables is given by:

$$E\left\{\begin{pmatrix} G_l W \\ V_l \end{pmatrix}\left((G_l W)^T \quad V_l^T\right)\right\} = \begin{bmatrix} Q_l & T_l \\ T_l^T & R_l \end{bmatrix}, \quad (10)$$

where

$$Q_l = G_l \begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & Q & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & Q \end{bmatrix} G_l^T,$$

$$R_l = D_w \begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & Q & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & Q \end{bmatrix} D_w^T + \begin{bmatrix} R & 0 & \cdots & 0 \\ 0 & R & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & R \end{bmatrix},$$

$$T_l = G_l \begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & Q & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & Q \end{bmatrix} D_w^T. \quad (11)$$

Note that although the process and measurement noise signals were uncorrelated for the original system, the lifted system contains a nonzero cross correlation T_l.

The following theorem is due to Amit (1980).

Theorem

Let $\Psi = A_l - T_l(R_l)^{-1}C_l$, and consider the Riccati equation

$$\Sigma_l - \Psi\Sigma_l\Psi^T + \Psi\Sigma_l C_l^T(C_l\Sigma_l C_l^T + R_l)^{-1}C_l\Sigma_l\Psi^T + T_l R_l^{-1} T_l^T - Q_l = 0. \quad (12)$$

If Σ_l satisfies Eq. 12, then $\Sigma[0] = \Sigma_l$ satisfies Eq. 7 for $k = 0$.

This theorem provides an efficient method for calculating the periodic Kalman filter. First, $\Sigma[0]$ is obtained by solving Eq. 12, and then each $\Sigma[k]$ for $1 \le k < N$ is calculated successively via Eq. 7. The estimate obtained can be used as the input to any state feedback control law. By using the filter output, it is unnecessary to wait for a complete scan before implementing the control action since the feedback calculation may use the estimate after each measurement. For control objectives that do not vary in time, such as minimization of $E\{x^T R_1 x\}$ where R_1 is a constant matrix, the optimal state feedback would not be time-varying due to the separation nature of the optimal solution. Therefore, all time-varying characteristics of the controller are contained in the estimator.

Estimation Error

The goal of this section is to develop effectiveness factors that will indicate when estimation errors and variations in controlled variables can be decreased by adding sensors. Let $e[k] = x[k] - \hat{x}[k|k-1]$ be the one-step-ahead estimation error. We consider the estimation performance measure $p_e[k] = e[k]^T W e[k]$, where W is a constant weighting matrix. For $W = I$, $p_e[k]$ is the square of the Euclidean norm of $e[k]$. The expected value of $p_e[k]$ is given by (Kwakernaak and Sivan, 1972)

$$E\{p_e[k]\} = E\{e[k]^T W e[k]\} = \text{trace}(\Sigma[k]W), \quad (13)$$

where $\Sigma[k]$ solves Eq. 7. For a PTV system with period N, $\Sigma[k] = \Sigma[k+N]$, and the expected value of $p_e[k]$ may be averaged over k, yielding

$$E\{p_e\} = \frac{1}{N}\sum_{i=0}^{N-1} E\{p_e[i]\} = \frac{1}{N}\sum_{i=0}^{N-1} \text{trace}(\Sigma[i]W). \quad (14)$$

If n-fixed sensors were used in lieu of the scanning sensor, $p_e[k]$ would be independent of k, and its expectation value would be given by trace $(\Sigma_0 W)$, where Σ_0 is the solution to the algebraic Riccati equation obtained by replacing $C[k]$ by C and $R[k]$ by R in Eq. 7. We define an estimation effectiveness factor η_e as the ratio of the average value of p_e using a scanning sensor to its value using n fixed sensors:

$$\eta_e = \frac{E\{p_e|\text{scanning sensors}\}}{E\{p_e|\text{fixed sensors}\}} = \frac{1}{N}\frac{\Sigma_{k=0}^{N-1}\text{trace}(\Sigma[k]W)}{\text{trace}(\Sigma_0 W)}. \quad (15)$$

The quantity η_e measures the potential improvement in one-step-ahead estimation that may be achieved by adding sensors. For $\eta_e \approx 1$, little improvement can be obtained by adding sensors, whereas for $\eta_e \gg 1$, considerable improvement could be achieved.

Now consider the case where the control variables u are determined by a constant state estimate feedback law $u[k] = -F\hat{x}[k|k-1]$. We consider the control performance measure $p_c[k] = x[k]^T R_1 x[k]$. For example, with $R_1 = C^T C$, $p_c[k]$ is the square of the Euclidean norm of $\hat{y}[k]$. In the case of scanning sensors, the closed-loop system is periodic with period N, so a meaningful measure is obtained as the average over one period of the expectation value of $p_c[k]$:

$$E\{p_c\} = \frac{1}{N}\sum_{i=0}^{N-1} E\{p_c[i]\} = \frac{1}{N} E\left\{\sum_{i=0}^{N-1} x[i]^T R_1 x[i]\right\}. \quad (16)$$

Since $x[k] = e[k] + \hat{x}[k|k-1]$, and the Kalman filter has the property that $e[k]$ and $\hat{x}[k|k-1]$ are statistically independent, this measure is equivalent to

$$\frac{1}{N}\text{trace}\left(\sum_{i=0}^{N-1} R_1(\Sigma[i]+\Sigma^*[i])\right), \tag{17}$$

where $\Sigma^*[k]=E\{\hat{x}^T[k|k-1]\hat{x}[k|k-1]\}$. By combining the feedback law, the definition of $e[k]$, and Eq. 5, $\hat{x}[k|k-1]$ can be seen to evolve as

$$\hat{x}[k+1|k]=(A-BF)\hat{x}[k|k-1] + AK[k](C[k]e[k]+v[k]). \tag{18}$$

Then by using the statistical independence of $\hat{x}[k|k-1]$, $e[k]$, and $v[k]$, the evolution of $\Sigma^*[k]$ follows:

$$\Sigma^*[k+1]=(A-BF)\Sigma^*[k](A-BF)^T + AK[k](C[k]\Sigma[k]C^T[k]+R[k])K^T[k]A^T. \tag{19}$$

Σ^* will also be periodic with period N. It can be calculated from Eq. 19, which can easily be transformed into a linear algebraic equation in Σ^*. Alternatively, it is shown by Kwakernaak and Sivan (1972) how Eq. 17 may be rewritten as:

$$\frac{1}{N}\text{trace}\left(\sum_{i=0}^{N-1} R_1\Sigma[i]+PAK[i](C[i]\Sigma[i]C^T[i] + R[i])K^T[i]A^T\right) \tag{20}$$

where P satisfies (In the case where the objective in Eq. 16 includes terms $u^T[i]R_2u[i]$ in the sum, the results are easily altered by adding the term $F^TR_2F\Sigma^*[i]$ to the sum in Eq. 17 and replacing R_1 in Eq. 21 by $R_1+F^TR_2F$. In the case where F is the optimal controller as in Eq. 31, Eq. 21 becomes the Riccati equation in X.)

$$P=(A-BF)^TP(A-BF)+R_1. \tag{21}$$

When n fixed sensors are used, $E\{p_c[k]\}$ is independent of k, and is given by

$$\text{trace}(R_1\Sigma_0+PAK_0(C\Sigma_0C^T+R)K_0^TA), \tag{22}$$

where $K_0=\Sigma_0C^T(C\Sigma_0C^T+R)^{-1}$.

We are now in a position to compare control improvement by adding sensors. Define a control efficiency factor η_c as the ratio of the average value of $p_c[k]$ using a scanner sensor to its value using n fixed sensors:

$$\eta_c=\frac{E\{p_c|\text{scanning sensor}\}}{E\{p_c|\text{fixed sensors}\}},$$

$$=\frac{1}{N}\frac{\text{trace}(\Sigma_{i=0}^{N-1}R_1\Sigma[i]+PAK[i](C[i]\Sigma[i]C^T[i]+R[i])K^T[i]A^T)}{\text{trace}(R_1\Sigma_0+PAK_0(C\Sigma_0C^T+R)K_0^TA).} \tag{23}$$

For $\eta_c\approx 1$, adding sensors will not reduce the average of the measure p_c, whereas for $\eta_c\gg 1$, substantial performance improvements may be realized.

Example

In this section, we consider estimation and control using both scanning sensors and stationary sensors at each measurement location. We consider a model that can be described by the transfer function equations

$$x(s)=G_1(s)u(s)+G_2(s)w(s), \tag{24}$$

where we consider $u(s)$ to be known control inputs, and $w(s)$ unmeasured disturbances.

Estimation

For constructing the PTV state estimator, we need only consider G_2. We will examine the case where G_2 is first order, that is

$$G_2(s)=\frac{\tau_2}{\tau_2 s+1}M_2, \tag{25}$$

where M_2 is a constant matrix that reflects the interactions between x and the disturbances w. We consider a scanning sensor that measures one of n variables every T time units and discretize the transfer function, yielding

$$G_2(z)=\frac{z^{-1}}{1-z^{-1}e^{-(T/\tau_2)}}M_2. \tag{26}$$

Introducing a scanning sensor, we rewrite the transfer function in state space form as:

$$x[k+1]=ax[k]+M_2w[k],$$

$$y[k]=C[k]x[k]+v[k], \tag{27}$$

where $a = e^{-(T/\tau_2)}$ and $C[k]$ has the form

$$\begin{aligned} C[0] &= (1, 0, \ldots, 0, 0) \\ C[1] &= (0, 1, \ldots, 0, 0) \\ &\vdots \\ C[n-2] &= (0, 0, \ldots, 1, 0) \\ C[n-1] &= (0, 0, \ldots, 0, 1) \\ C[n] &= (0, 0, \ldots, 1, 0) \\ &\vdots \\ C[2n-3] &= (0, 1, \ldots, 0, 0) \\ C[k+2n-2] &= C[k]. \end{aligned} \tag{28}$$

Also, let $R = r$, $Q = qI_m$, where $r, q \in \Re$, and M_2 is the Toeplitz matrix given by $M_2(i, j) = \rho^{|i-j|}$ (for $\rho = 0$, $M_2 = I_m$). The quantity η_e was calculated for this model for a system with 12 measurements and with values of $T/\tau_2 = 0$ and 2^m for $m = -4, -3, \ldots, 1$. For this model form, η_e depends only on the ratio q/r, and this quantity was assigned the values 0.1, 1, and 10. ρ was allowed to range from 0 to 1 by increments of 0.1. The measure η_e is symmetric in $\rho[\eta_e(\rho) = \eta_e(-\rho)]$. The results are shown in Figure 1.

In the limit as $q/r \to 0$, $\eta_e \to 1$. This is expected since large measurement noise ($q/r \approx 0$) implies inaccurate measurements. In this situation, more confidence is given to the model than to the measurements, and little improvement could be obtained by adding sensors. An increasing value of q/r suggests that the measurements become more accurate than the model. In this situation, η_e becomes large. In the limiting case that $q/r \to \infty$, and uncorrelated process disturbances ($\rho \to 0$) $K[k] \to C^T[k]$ for the model in Eq. 27. From Eq. 5, it is easily seen that this $K[k]$ corresponds to updating the unmeasured states using the model and the measured states with the measurement.

Figure 1 also shows that when the process disturbances are uncorrelated ($\rho = 0$), η_e is larger than for the case where the disturbances are identical throughout the cross direction ($\rho = 1$). For the model in Eq. 27, $\rho = 0$ yields a $K[k]$ which uses the measurement to update only the most recently measured variable [$K[k] = \alpha C^T[k]$, $\alpha \in (0, 1)$], whereas for $\rho = 1$, $K[k]$ updates each variable in the same manner,

$$K[k] = \alpha \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

In the case of an adhesive die coater (Braatz et al., 1992), the latter case may correspond to the physical situation in which the disturbances are caused primarily by variations in the flow of adhesive to the die, in which case cross-directional control may not be necessary, whereas the former case corresponds to the measurements being too far apart to feel the effects of neighboring positions. A value of ρ between 0 and 1 corresponds to partially correlated disturbances which vary across the machine direction and whose cause may be due to imperfections in the die, the roller, or the feed paper.

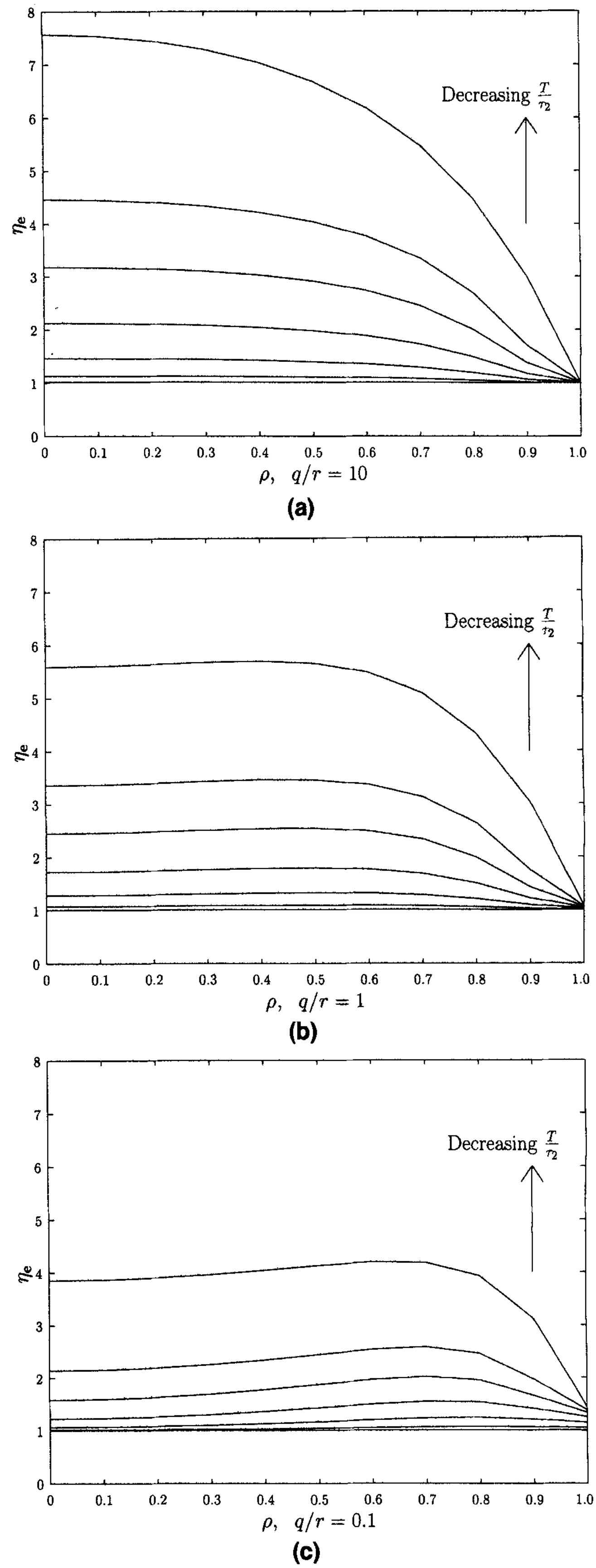


Figure 1. η_e ($n = 12$) as a function of ρ and q/r, for $T/\tau_2 = 0$, 1/16, 1/8, 1/4, 1/2, 1, and 2.

The time constant to sampling rate ratio T/τ_2 has a strong influence on η_e. Disturbances that die out quickly in comparison to the sampling rate ($T/\tau_2 \to \infty$) are not easily estimated, and adding sensors leads to little improvement ($\eta_e \approx 1$). At the other extreme, a sequence of steplike disturbances ($T/\tau_2 \to 0$) can be estimated easily, and adding sensors significantly reduces the estimation error (η_e large).

As the dimension of the system increases, one would expect that the estimate obtained from using but one sensor will become inadequate. In Figure 2, we show the dependence of η_e on system size for a value of $\rho = 0.5$ and $q/r = 10$. As can be seen, as the system size increases, η_e increases also. When one sensor is inadequate, one may not want to add sensors at each actuator location for economic or other extraneous considerations. Suppose, instead of adding a sensor for each variable, only one additional sensor is added. The sensors are configured such that the first sensor scans the variable, $1, \ldots, N/2$ and the second sensor scans variables $N/2+1, \ldots, N$. The above analysis can be used to determine that the optimal configuration would consist of the sensors in phase: sensor 1 scans variable k, while sensor 2 is scanning variable $k + N/2$. In the case of a system with 12 measurements, $\rho = 0.5$, $q/r = 10$, a single sensor yields a value of $\eta_e = 6.69$, whereas a two-sensor configuration produces $\eta_e = 2.99$, an improvement of more than a factor of 2.

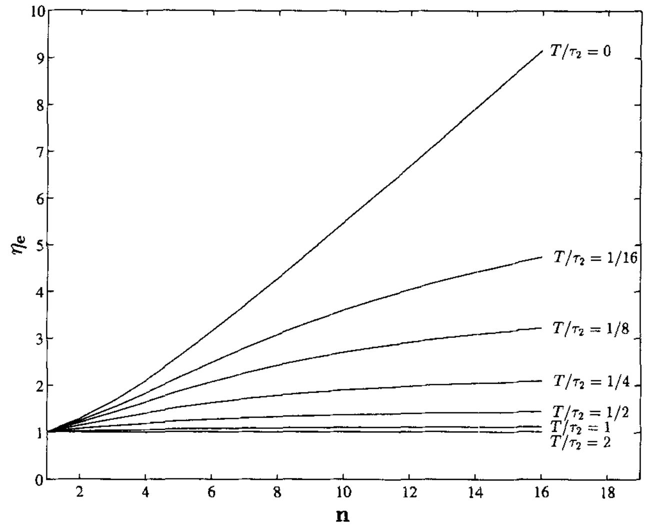


Figure 2. η_e as a function of T/τ_2 and system size n for $\rho = 0.5$ and $q/r = 10$.

Closed-loop performance

We now consider the closed-loop performance of the system described by

$$x(s) = \frac{\tau_1 e^{-\theta s}}{\tau_1 s + 1} M_1 u(s) + \frac{1}{s} M_2 w(s). \tag{29}$$

M_2 is as before, with $\rho = 0.5$ and $M_1 = M_2$. We choose the step model for the disturbance, as the previous analysis has shown this form has the largest potential for improving the estimation error by adding sensors. By using a sampling time of T and assuming an integer value of $\theta/T = d$, the system may be written

$$\begin{aligned} x_1[k+1] &= \beta x_1[k] + M_1 \Delta u[k-d], \\ x_2[k+1] &= x_2[k] + M_1 \Delta u[k-d-1] + M_2 w[k], \\ \hat{y}[k] &= x_1[k] + x_2[k], \end{aligned} \tag{30}$$

where $\beta = e^{-(T/\tau_1)}$. In this description, x_2 represents the process disturbances and the effect of all previous control moves. The representation can easily be transformed to state space by letting $\Delta u[k]$ be the input, and $\xi[k] = [x_1^T[k], x_2^T[k], \Delta u^T[k-d-1], \ldots, \Delta u^T[k-1]]^T$ be the states. The state feedback law which minimizes the objective $E\{\hat{y}^T R_1 \hat{y} + \Delta u^T R_2 \Delta u\}$ is given by

$$\Delta u[k] = -F\xi[k]$$

$$F = (R_2 + B^T X B)^{-1} B^T X A$$

$$A^T X A - X - A^T X B (R_2 + B^T X B)^{-1} B^T X A + C^T R_1 C = 0, \tag{31}$$

where A, B, C are the state-space matrices for the system in Eq. 30 and are given by

$$A = \begin{bmatrix} \beta I & 0 & 0 & M_1 & 0 & \ldots & 0 \\ 0 & I & M_1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & I & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & I & \ldots & 0 \\ \vdots & & & & & \ddots & \\ 0 & 0 & 0 & 0 & 0 & \ldots & I \\ 0 & 0 & 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix},$$

$$C = [I \;\; I \;\; 0 \;\; 0 \;\; 0 \;\; \ldots \;\; 0]. \tag{32}$$

For this example, we let $R_1 = r_1 I_n$ and $R_2 = r_2 I_m$, where r_1 and r_2 are scalars. The time parameter β has only a weak effect on η_c; however, r_2 and the delay d strongly affect η_c. Therefore, even when adding sensors may give a substantial improvement in estimation error, if the system delay is significant, or robustness considerations require a large value of r_2, obtaining a better cross-directional estimate may not significantly enhance closed-loop system performance. For example, Figure 3a depicts the effect of process delay and tuning parameter r_2 on the efficiency factor η_c for a system with 12 measured variables, $\rho = 0.5$, $q/r = 10$, and $a = 0.5$. For this example, $\eta_e = 6.69$; however, for $r_2 = 1$, $\eta_c = 3.78$ for delay $d = 0$ and drops to a value of 1.49 for $d = 10$. If two scanning sensors are used, η_c assumes the values of 1.97 and 1.17 for $d = 0$ and 10, respectively.

Figure 3b shows the effects of the noise parameters q/r and ρ on η_c for $d = 0$ and $r_2 = 0.01$. As expected, the most significant improvements from stationary sensors can be obtained when the measurements are relatively noise-free, and the disturbances are uncorrelated (q/r large and ρ small).

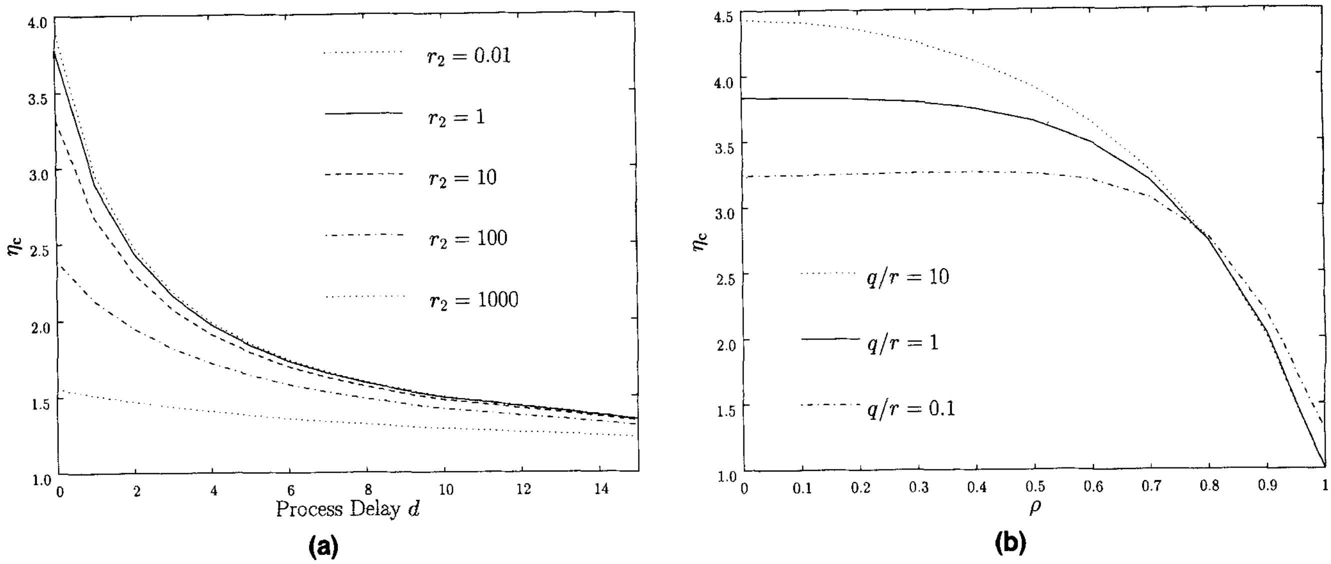


Figure 3. a. η_c as a function of r_2 and d, $\rho = 0.5$; b. η_c as a function of q/r and ρ, $d = 0$, $r_2 = 0.01$.

Robustness to errors in disturbance statistics

As the properties of the estimator depend strongly on the parameter ρ, we would like to consider the performance of the closed-loop scheme with respect to this parameter. In this discussion, we will assume that disturbances enter the system as a sequence of random steps and that disturbances occur at time instances far enough apart so that the closed-loop system completely rejects the previous disturbance before a new disturbance enters. We also assume that measurement noise can be neglected. With the time-varying Kalman filter and constant state feedback, the closed-loop system can be described by

$$\begin{bmatrix} \xi[k+1] \\ \hat{\xi}[k+1|k] \end{bmatrix} = \begin{bmatrix} A & -BF \\ K[k]C[k] & A - AK[k]C[k] - BF \end{bmatrix} \times \begin{bmatrix} \xi[k] \\ \hat{\xi}[k|k-1] \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} w[k],$$

$$= \bar{A}[k] \begin{bmatrix} \xi[k] \\ \hat{\xi}[k|k-1] \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} w[k],$$

$$\hat{y}[k] = C\xi[k]. \tag{33}$$

Here, $\hat{y}[k]$ denotes the vector of cross-directional properties, including those at positions not measured at time k. We now consider the first $s+1$ outputs from a disturbance introduced at $k = 0$,

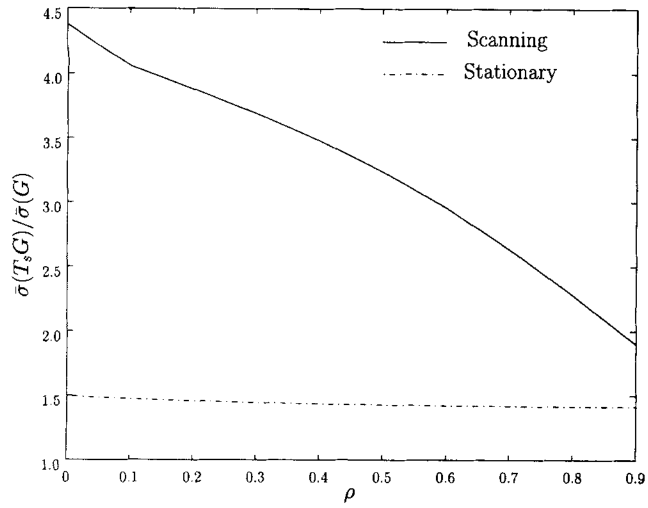
$$\mathbf{Y}_s[0] = \begin{bmatrix} \hat{y}[0] \\ \hat{y}[1] \\ \vdots \\ \hat{y}[s] \end{bmatrix} = \begin{bmatrix} C \\ C\bar{A}[0] \\ \vdots \\ C\Pi_{j=0}^{s-1}\bar{A}[j] \end{bmatrix} G w_0 = T_s G w_0. \tag{34}$$

We consider the measure

$$\max_{\|Gw_0\|_2 = 1} \|\mathbf{Y}_s[0]\|_2 = \frac{\bar{\sigma}(T_s G)}{\bar{\sigma}(G)}, \tag{35}$$

where $\bar{\sigma}$ denotes the maximum singular value. The quantity $\|\mathbf{Y}_s[0]\|_2^2$ is the sum of squares of the norm of $\hat{y}$ over the first $s+1$ time steps after the disturbance is introduced. For s large enough and an asymptotically stable closed loop, $\|\mathbf{Y}_s[0]\|_2$ approaches the temporal two norm of $\|Y[k]\|$.

We consider two cases for the matrix G. In the first case, $G = [0, M_2, 0, \ldots, 0]^T$, that is, the correct structure assumed in deriving the correct filter parameters $K[k]$. In the second case, $G = [0, I_m, 0, \ldots, 0]^T$, that is, the disturbances are equally likely to enter in any direction. With a scanning sensor, the worst-case disturbance enters in the direction which takes the longest for the filter to detect and is therefore quite pessimistic. For example, when $\rho = 0$, the worst-case disturbance would be isolated at the position which maximizes the time between introducing the disturbance and scanning the disturbed position. Thus, if the scanner is located at position 2 and traveling toward position 3, the worst-case disturbance is a unit deviation at position 1. Clearly, the measure in Eq. 35 depends on the location of the scanning sensor at $k = 0$. For this example, we assumed the scanner to be located in the center of the sheet. Figure 4 compares the maximum gain for the scanning sensor and the stationary sensor cases for the case where $K[k]$ was calculated using $q/r = 10$, 12 measurement positions, and $a = 0.5$, $d = 0$. Note that when the disturbances are highly correlated and this information is correctly included in the filter equations, performance is better than for low levels of correlation; however, if high correlations are used for the filter calculation and are not physically justified, the performance deteriorates. Thus it is better to underestimate the extent of disturbance correlation for this model. In addition, it can clearly be seen from Figure 4 that adding sensors increases robustness to uncertainty in disturbance directions, as expected.



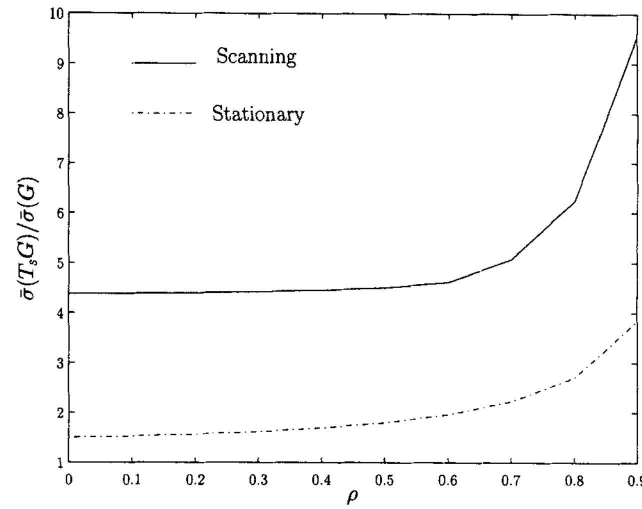


Figure 4. $\bar{\sigma}(T_sG)/\bar{\sigma}(G)$ **as a function of** ρ**: a.** $G=[0, M_2, 0, \ldots, 0]^T$**; b.** $G=[0, I_m, 0, \ldots, 0]^T$**.**

Conclusions

Estimates of cross-directional properties from a scanning sensor can be obtained using a periodic, time-varying Kalman filter. The equations for this filter can be solved by "lifting" the PTV system to form an LTI system, and thus transforming the periodic Riccati difference equation governing the Kalman filter into an algebraic Riccati equation. The solution to the periodic Riccati equation can be obtained by stepping through the difference equation, with the solution to the algebraic Riccati equation as the starting point.

The periodic Kalman filter and the solution to its accompanying Riccati equation can be used to estimate improvement to estimation errors. The extent to which adding sensors can improve performance of the estimation scheme depends on the sampling time, the evolution characteristics of the disturbance, and the accuracy of the model and the measurements. Additional sensors will decrease the estimation error most significantly when the disturbances are steplike, and the measurements are substantially more accurate than the model ($q/r \gg 1$).

When the estimates obtained from the Kalman filter are used as feedback for a control scheme, the solution to the Kalman filter problem also provides a measure of the improvement of control objectives that can be expected by adding sensors. Substantial improvement in the cross-directional estimates may translate to much less improvement in the cross-directional variations. In particular, when robustness considerations require the controller to be detuned, or large process delays are inherent in the system, little improvement is observed.

The pilot adhesive coater described by Braatz et al. (1992), modified so that control action is taken after each measurement rather than after each complete scan, provides a typical situation for the application of these results. This system has 12 cross-directional measurements. Due to the measurement instrumentation, the sampling time T is inherently large. In this case, only steplike disturbances need to be considered, as transient disturbances will die out between sampling instances. For very large T, the process delay d will approach 0, corresponding to only a measurement delay. Due to the sluggish plant behavior, robustness considerations will not be important, allowing a small value of the parameter r_2. If we consider disturbance interaction of $\rho = 0.5$, $\eta_c = 3.78$, suggesting that the root-mean-square (rms) cross-directional deviations can be reduced by nearly a factor of 2 by using 12 stationary sensors. If two scanning sensors are used, $\eta_c = 1.98$, indicating rms deviations approximately 40% higher than when 12 sensors are used.

Acknowledgment

We are grateful to Prof. J. B. Rawlings for various helpful discussions during the course of this work.

Literature Cited

Amit, N., "Optimal Control of Multirate Digital Control Systems," PhD Thesis, Stanford Univ. (1980).

Bergh, L. G., and J. F. MacGregor, "Spatial Control of Sheet and Film Forming Processes," *Can. J. of Chem. Eng.*, **65**, 148 (1987).

Boyle, T. J., "Practical Algorithms for Cross-Directional Control," *Tappi J.*, **61**, 77 (Jan., 1978).

Braatz, R. D., M. L. Tyler, M. Morari, F. R. Pranckh, and L. Sartor, "Identification and Cross-Directional Control of Coating Processes," *AIChE J.*, **38**, 1329 (1992).

Dumont, G. A., I. M. Jonsson, M. S. Davies, F. T. Ordubadi, K. Natarajan, C. Lindeborg, and E. M. Heaven, "Estimation of Moisture Variations on Paper Machines," *IEEE Trans. on Cont. Sys. Technol.*, **1**, 101 (June, 1993).

Kwakernaak, H., and R. Sivan, *Linear Optimal Control Systems*, Wiley, New York (1972).

Laughlin, D. L., "Control System Design for Robust Performance Despite Model Parameter Uncertainties: Application to Cross-Directional Response Control in Paper Manufacturing," PhD Thesis, Cal. Inst. of Technol., Pasadena (1988).

Laughlin, D. L., M. Morari, and R. D. Braatz, "Robust Performance of Cross-Directional Basis-Weight Control in Paper Machines," *Automatica*, **29**, 1395 (1993).

Lee, J. H., M. S. Gelormino, and M. Morari, "Model Predictive Control of Multi-rate Sampled-Data Systems: a State-Space Approach," *Int. J. of Cont.*, **55**, 153 (1991).

Lee, J. H., and M. Morari, "Robust Inferential Control of Multi-rate Sampled-Data Systems," *Chem. Eng. Sci.*, **47**, 865 (1991).
McFarlin, D. A., "Control of Cross-Machine Sheet Properties on Paper Machines," *Proc. Int. Pulp and Paper Process Cont. Symp.*, p. 49, Vancouver, BC, Canada (1983).
Richards, G. A., "Cross Direction Weight Control," *Japan Pulp & Paper*, p. 41 (Nov., 1982).
Wang, X. G., G. A. Dumont, and M. S. Davies, "Estimation in Paper Machine Control," *IEEE Cont. Sys.*, **13**(8), 34 (Aug., 1993).
Wilhelm, R. G., and M. Fjeld, "Control Algorithms for Cross-Directional Control: the State of the Art," *Proc. IFAC Conf. on Instrument. and Automat. in Paper, Rubber, Plastics, and Polymer. Ind.* (*PRP*-5), p. 139 (1983).

Manuscript received Feb. 7, 1994, and revision received May 6, 1994.