ESTIMATORS FOR ILL-CONDITIONED PLANTS: HIGH-PURITY DISTILLATION.

Thor Mejdell and Sigurd Skogestad

Chemical Engineering, Norwegian Institute of Technology (NTH),
N-7034 Trondheim, Norway.

Abstract. Temperatures and flows are often used as secondary measurements to estimate the product compositions (outputs) in distillation columns. The problem is characterized by strong collinearity (correlation) between the temperature measurements, and often between the effects of the inputs on the outputs. In a linear study three different estimator methods, the Kalman-Bucy Filter, Brosilow's inferential estimator, and Principal Component Regression (PCR) are tested for performance with mu-analysis. It is found that use of input flow measurement has a damaging effect on the estimator performance for this ill-conditioned plant (with high RGA-elements). This is the main reason why the Brosilow inferential estimator is found to perform poorly. It is found that the static PCR-estimator performs remarkable well compared with the dynamic Kalman filter. Contrary to some claims in the literature, it is found that the performance of the estimate generally is improved by adding temperature measurements.

Keywords. Estimation; Distillation columns; Large-scale systems; Data reduction; Kalman filters; Inferential control; Multivariate Calibration; Input uncertainty; Robust control; Temperature control.

1 Introduction

This paper addresses estimation of unmeasured process outputs based on multiple secondary measurements. The application chosen here is the use of temperature and flow measurements to estimate the product compositions in a distillation column. This is an interesting application which features: i) a large number of strongly coupled measurements, and ii) several disturbances and inputs with similar effects on the outputs.

Problem definition. The objective is to obtain the best estimate $\hat{y}$ of the outputs (product compositions in our application) using all available information, $\tilde{\theta}$. In terms of deviation variables the linear estimator may be written

$$\hat{y}(s) = K(s)\tilde{\theta}(s) \tag{1}$$

This estimate should be obtained based on a description of the process (nominal model and expected uncertainty), the expected noise and disturbances, and a more precise definition of what we mean by "best". In the general case $\tilde{\theta}$ should include all measured dependent variables (primary measurements, y, and secondary measurements, θ), and all known independent variables (manipulated inputs, u, and measured disturbances, d). In this paper we usually have $\tilde{\theta} = \theta$, that is, the estimate is based on only secondary measurements (temperatures in our application). The reason is that we assume no primary measurements, no measured disturbances, and we shall show for our case that the additional information contained in u is of limited value. This estimation problem is usually called "inferential" estimation in process control litterature.

Use of separate estimator. A control scheme for a distillation column based on an estimator is shown in Fig.1. Note that we are implicitly assuming that the controller should be separated into two parts: one estimator which condenses all the measurements into a few estimated outputs, and a "small" (in terms of number of inputs) controller which uses these estimates for feedback control. The motivation for doing this is reliability, design simplicity and robustness.

In this paper we consider three different approaches to the problem: i) The Kalman-Bucy Filter, ii) Brosilow's Inferential Control Method, and iii) Principal Component Regression (PCR). In the last two cases we shall base the analysis on the steady-state, and use a constant gain matrix K.

Use all available measurements ? The statement in the problem definition above that the best estimate should be based on *all* available measurements is not as obvious as one should think. Actually, a large number of authors (eg. Joseph and Brosilow, 1978, Morari and Stephanoplous, 1980, Patke et al., 1982, Yu and Luyben, 1986, Moore et al., 1987, Keller and Bonvin 1987) have suggested that one should only use a few of the *temperature* measurements to avoid the poorly conditioned problem of obtaining information from the strongly correlated temperatures. For example, our example column has 41 temperature measurements. That is, we need to determine 41 parameters in K for each output if all temperatures are used. However, the temperatures are of course strongly coupled and the 41 parameters must also be strongly coupled. In fact, our distillation column has only three degrees of freedom at steady state. This implies that, at least for the linear case with small perturbations from

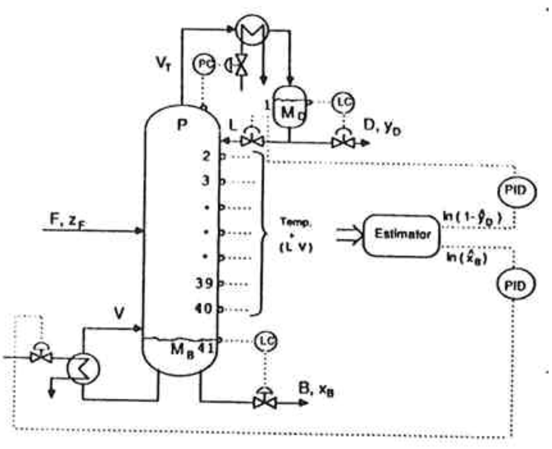

Figure 1: Control scheme based on LV configuration.

the nominal operating point, we may determine at most three of these 41 parameters independently. This points out the need for a robust way of obtaining the matrix K which avoids this overparameterization.

One such method is the Principal Component Regression (PCR), where the available data are smoothened by obtaining a smaller number, k, of "latent variables", t, which are less coupled and contain most of the original information. These are subsequently used for estimation. In the linear case these latent variables may be written $t = P_t\theta$, where P_t is the projection matrix. The estimator then becomes $\hat{y} = K_t t$ where K_t is a "small" matrix with k parameters for each output (typically $k = 3$ in our examples), and the overparametrization in the regression step is avoided.

Another "method", but certainly not the optimal one, is to delete measurements in θ, and use, for example, only three temperatures. This approach is implicit in some of the papers mentioned above.

2 Distillation Column Application

As an example we use the distillation column A studied by Skogestad and Morari, 1988. The column separates a binary mixture with relative volatility 1.5, and has 40 theoretical stages, including the reboiler, plus a total condenser. Column data are given in Table 1. The liquid holdups are assumed constant, that is, the flow dynamics are neglected. This gives rise to a 41th order linear model in terms of the mole fraction of the light component on each tray. For our binary mixture with constant pressure there is a one-to-one relationship between the mole fraction and the temperature on each tray, and the model becomes

$$y(s) = G_u(s)u(s) + G_d(s)d(s) \tag{2}$$

$$\theta(s) = F_u(s)u(s) + F_d(s)d(s) \tag{3}$$

where $u = [L, V]^T, d = [z_F, F]^T, y = [y_D, x_B]^T$ and θ = temperatures on all trays. The two dominant time constants of the column are 194 min and 15 min. At steady-state (Skogestad et al., 1988, Skogestad and Morari, 1988)

$$G_u(0) = \begin{pmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{pmatrix} \tag{4}$$

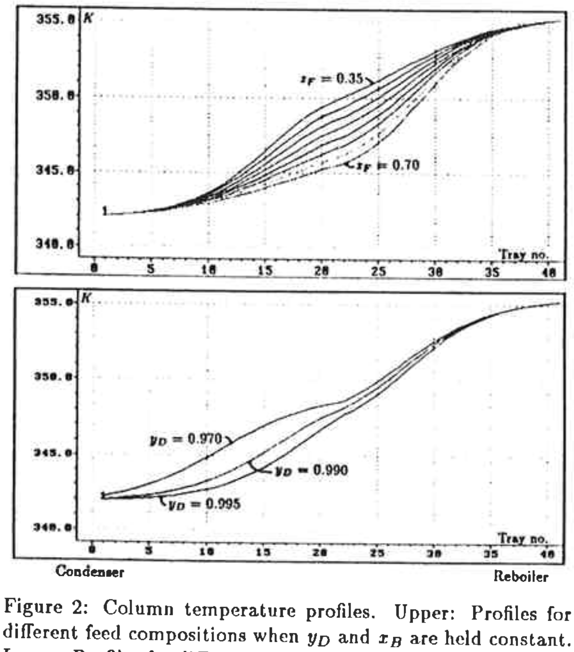

Figure 2: Column temperature profiles. Upper: Profiles for different feed compositions when y_D and x_B are held constant. Lower: Profiles for different top product compositions when z_F and x_B are held constant.

α	N	N_F	z_F	y_D	x_B	D/F	L/F
1.5	40	21	0.50	0.99	0.01	0.500	2.706

- Feed is liquid.
- Const. molar flows.
- Ideal VLE
- Pressure 1 atm.
- Holdup on each tray; $M_i/F = 0.5$ min

Antoine parameters:	A	B	C
Light component	15.8366	2697.55	−48.78
Heavy component	15.4311	2697.55	−48.78

Table 1: Data for distillation column example.

Here the gain matrix is scaled such that an output of magnitude 1 corresponds to 0.01 mole fraction units, and the inputs are scaled relative to a feed rate of $F = 1$. The condition number of this matrix is 141.7 and the diagonal RGA elements are 35.5. This plant is consequently illconditioned (Skogestad and Morari, 1987). The other matrices are also ill-conditioned at steady state.

The difference in boiling points of the two pure components is 13 °C. In Figure 2 some temperature profiles for the column are displayed. We note that variations in temperature are small towards the ends of the columns, and that changes in feed composition have a large effect on the temperatures inside the column even though the product compositions are constant. This demonstrates the limitations of single temperature control where a temperature measurement at the end of the column will not be sensitive enough, while a measurement inside the column (e.g. tray 10) will be biased by changes in feed composition.

3 Estimators

3.1 Kalman filter.

In the Kalman filter scheme (Kalman and Bucy, 1961) a dynamic state space model is used in parallel with the process, and the deviation between the outputs from the process and the model is used as feedback to the model through a filter gain K_f.

The linear state space model for the process is

$$\dot{x} = Ax + Bu + Ev \tag{5}$$

$$y = Cx \tag{6}$$

$$\theta = C_\theta x + w \tag{7}$$

Here x is the state vector, u the manipulated inputs, y the primary outputs to be estimated, θ the secondary measurements, v the process noise (disturbances), and w the measurement noise. v and w are assumed to be white noise processes with covariance matrices $\mathcal{V}$ and $\mathcal{W}$.

Minimizing the expected variance of $\theta - \hat{\theta}$ yields the estimated states

$$\dot{\hat{x}} = A\hat{x} + Bu + K_f(\theta - C_\theta\hat{x}) \tag{8}$$

$$= (A - K_f C_\theta)\hat{x} + Bu + K_f\theta \tag{9}$$

where filter gain K_f is

$$K_f = \mathcal{X} C_\theta^T \mathcal{W}^{-1} \tag{10}$$

Here $\mathcal{X}$, the covariance matrix of $\hat{x}$, is found from the matrix Riccati equation

$$\dot{\mathcal{X}} = A\mathcal{X} + \mathcal{X}A^T - \mathcal{X}C_\theta^T\mathcal{W}^{-1}C_\theta\mathcal{X} + E\mathcal{V}E^T \tag{11}$$

We use constant filter gains which give $\dot{\mathcal{X}} = 0$, and Eq. (11) is reduced to an algebraic equation. The overall Kalman estimator then becomes

$$\hat{y}(s) = C(sI - A + K_f C_\theta)^{-1}(K_f\theta(s) + Bu(s)) \tag{12}$$

The covariance matrix of the measurement noise $\mathcal{W}$ was set to $0.04I$ for our example column (I is the identity matrix). This corresponds to 0.2 °C noise on each temperature. The process noise (disturbance) is here $v^T = [L, V, F, z_F]$ (reflux, boilup, feedrate and feed composition). Its covariance matrix, $\mathcal{V}$, was assumed diagonal and was varied in order to tune the filter. Four different values of the variance on L and V were selected (Table 2) and the corresponding filter gain matrices are denoted K1 to K4. The assumption of white noise process disturbances is somewhat unrealistic in a distillation column, and we might add an integrator and use a non-stationary disturbance $d = \frac{1}{s}v$. However, the estimator is not expected to be significantly improved by such changes, although it would remove some of the steady-state offset which is apparent in later μ-plots.

3.2 Brosilow estimator.

The following linear steady-state model of the column in terms of deviation variables is used in the Brosilow estimator (Weber and Brosilow, 1972, Joseph and Brosliow, 1978):

$$y = G_d d + G_u u \quad (13)$$

$$\theta = F_d d + F_u u \quad (14)$$

Here d denotes the disturbances. The matrices above are of course related to those used in the state space description in the Kalman filter. For example, $G_u = -CA^{-1}B$ and for the case $v = d$ we have $F_d = -C_\theta A^{-1}E$. Using (14) the estimated disturbances become

$$\hat{d} = F_d^\dagger(\theta - F_u u) \quad (15)$$

where the pseudoinverse $F_d^\dagger$ is the optimal inverse in the general least square sense. The inversion in (15) may be impossible, or at least numerically ill-conditioned, when there are collinearities among the variables. To avoid some of these problems one should in general case obtain the pseudoinverse, $F_d^\dagger$, from a SVD of F_d by deleting directions with singular values equal to zero (eg., see Strang, 1980, p. 142).

Combining (13) and (15) yields the Brosliow estimator

$$\hat{y} = K_B\theta + (G_u - K_B F_u)u \quad (16)$$

where

$$K_B = G_d F_d^\dagger \quad (17)$$

In the example column $d^T = [z_F, F]$ and $u^T = [L, V]$. The matrices F_d,F_u,G_d and G_u were found by linearizing the model at the nominal operating point.

A modified estimator $K_{B_{mod}}$ was formed by *not* using information about the manipulated inputs u, and instead using $d'^T = [L, V, z_F]$ as the disturbances to be inferred. The estimator then becomes $\hat{y} = K_{Bmod}\theta$ where

$$K_{Bmod} = G'F'^\dagger \quad (18)$$

and F' and G' are the process matrices formed by these three variables. In the linear case with no numerical errors in the matrices G' and F', this estimator is identical to the PCR-estimator.

3.3 PCR estimator.

We want to estimate p outputs (y) from q known variables (θ). The problem is then to obtain the matrix K in

$$\hat{y} = K\theta \quad (19)$$

To this end obtain n "calibration" runs corresponding values of y and θ, and place these as *rows* in the matrices $Y^{n\times p}$ and $\Theta^{n\times q}$, respectively. If the estimator was perfect we would have

$$Y = \Theta K^T \quad (20)$$

Case		$\mathcal{V}$				
		L	V	F	z_F	
K1	diag{	200	200	0.01	0.01	}
K2	diag{	0.10	0.10	0.01	0.01	}
K3	diag{	0.01	0.01	0.01	0.01	}
K4	diag{	0.0	0.0	0.01	0.01	}

Table 2: Process disturbance covariance matrix of Kalman filter gains. In all cases $\mathcal{W} = 0.04I$

The general least square solution is (e.g. Strang, 1981, p 139)

$$K_{GLS} = Y^T[\Theta^T]^\dagger \quad (21)$$

In addition to minimizing $(y - \hat{y})^2$ this solution minimizes the norm of K. The pseudo inverse is obtained from a singular value decomposition (SVD) of Θ. Using standard notation from the statistics literature (see eg. Wold et. al.), the SVD of Θ is written

$$\Theta = t_1 p_1^T + t_2 p_2^T + \cdots + t_m p_m^T \quad (22)$$

where $m \leq min(n, q)$ is the rank of Θ. Here p_1 is the eigenvector corresponding to the largest eigenvalue of $\Theta^T\Theta$, (the square of the largest singular value of Θ), and p_2 is the eigenvector corresponding to the second largest eigenvalue, and so on. The loading vectors (p's) give the directions of the principal components, while the scores (t's) give the magnitude. If all m terms in (22) are retained we obtain the generalized pseudoinverse in (21). However, in PCR we select only those first k principal components that can be distinguished from the measurement noise. Let the matrices $P^{q\times k}$ and $T^{n\times k}$ include only these k most important directions. Define the new latent variables as $t = P^T\theta$. Note that $P^T = P^{-1}$ since P is orthonormal. The least square solution to $y = K_t t$ becomes $K_t = Y^T T[T^T T]^{-1}$, and the overall estimator gain matrix becomes

$$K_{PCR} = Y^T[\Theta_k^T]^\dagger = Y^T T[T^T T]^{-1}P^T \quad (23)$$

In the general case θ may be replaced by $\tilde{\theta}$ which includes also the inputs and measured disturbances.

The calibration sets for the example column were obtained from a *linear* steady state column model. A factorial design method was used to select 16 different runs around the operating point. (The reason for using more runs than strictly necessary, was to better study the effect of measurement noise and to get better statistical information.) The specified variables were chosen as the outputs y_D and x_B and the feed composition z_F. (Since the column conditions in the simulation model are independent of the load, it is not necessary to simulate different feed rates). With this approach we may freely vary the *outputs* (y_D and x_B), and are thus able to span all directions in the output space. This is different from the Brosilow approach, which is based on an open-loop model in terms of the *inputs* (L, V, F, z_F), and where the output space will not be properly spanned for ill-conditioned plants with strongly coupled outputs.

When stated random noise of magnitude 0.1 °C was added on all temperatures in the calibration sets, but the default is no noise. The temperature data were reduced to the desired number of principal components and K_{PCR} was computed from (23).

4 Analysis of the Estimators.

4.1 Evaluation criteria

Open-loop evaluation (OL). One obvious criteria for evaluating the different estimators is the estimation error $e_{OL} = y - \hat{y}$. This is the difference between the real (y) and the estimated output ($\hat{y}$). The system is assumed to operate under feedback, since this is closer to a real situation than a pure open loop test where it may "drift away". The term "open loop" is still used since the controller uses the actual y, that is, there is no feedback from the estimate $\hat{y}$. We use single-loop PID controllers since this is the most common choice in practice. The tunings yield optimal robust performance (minimize μ) when the estimate is exact. To make our results less dependent on the controller used, we shall usually consider the *nominal performance* in this test, i.e., without any uncertainty. This makes the comparison independent of the robust stability requirement of the system which depends strongly on the controller.

Closed-loop evaluation (CL). The main objective of the estimator is to replace the primary measurement of y, that is, use the estimate $\hat{y}$ for feedback control. The error of interest to be minimized, is then the control error $e_{CL} = y - y_s$ Here y_s is the setpoint (reference signal). We consider robust performance of e_{CL} in this case, i.e. uncertainty is included. We use the same controller as for the open-loop comparison, that is, a PID controller tuned optimally for perfect estimates. Using the same PID controller for all estimators will bias the comparison somewhat, as the optimal controller in each case will depend on the estimator used.

4.2 μ-analysis.

Our tool is the Structural Singular Value (μ) analysis (Doyle, 1982). In this framework we rearrange our system to fit the general form shown in Fig. 3. Here the interconnection matrix N includes the plant, the controller, the estimator and the weights. $\tilde{d}$ denotes external inputs (disturbances, noise and setpoint changes), and e is the "error" we want to keep small. We have a separate Δ-block loop to represent the model uncertainty. In the μ-analysis we evaluate the maximum amplification from $\tilde{d}$ to e at each frequency. Weights are used to scale the signals, $\tilde{d}$ and e, and the uncertainty Δ to be less than 1.

μ expresses the worst-case error at a given frequency, and the performance requirement for the error e_{OL} or e_{CL} is satisfied if μ is less than one at all frequencies. Nominal performance (with $\Delta = 0$) for the estimation error e_{OL} is satisfied iff $\mu(N_{22}) < 1$, and robust performance (for all allowed Δ's) for the control error e_{CL} is satisfied iff $\mu(N) < 1$. In the paper we plot μ as a function of frequency, and estimators with small μ-values are prefered. The weights used for external inputs, performance and uncertainty are given in Mejdell and Skogestad (1989). Noise was generated by adding a constant vector of random values with normal distribution and a standard deviation of 0.2 $°C$ to all 41 temperatures. The noise is icluded in the μ-analysis only if stated.

5 Results.

5.1 Comparison of Kalman filter and static PCR estimator.

In Fig. 4 we compare the μ-plots of the Kalman and PCR estimators, using 41 temperatures. The first thing to note is how well the simple static estimator $\hat{y} = K_{PCR}\theta$ performs. The main reason is that the dynamic responses of the temperatures θ and the compositions y are very similar. This will be the case for most distillation columns, at least for sections of the column, but may of course not be the case for other applications.

In the Open Loop analysis the Kalman filter is significantly better at higher frequencies. This is due to the dynamics included in this estimator. On the other hand, the "Closed Loop" test shows that the estimators will perform about equally well when used for feedback, and also as well as using perfect measurements. Actually, for some frequencies, the PCR estimator is even better than using perfect measurements. The reason is that the temperatures in the middle of the column generally change slightly faster than at the ends, and the steady state estimator will therefore have a small inherent "feedforward" effect. The simulation responses in Figure 5 confirm that the PCR-estimate is almost equal to the true value. One exception is for feed composition disturbances, where it shows a small inverse response.

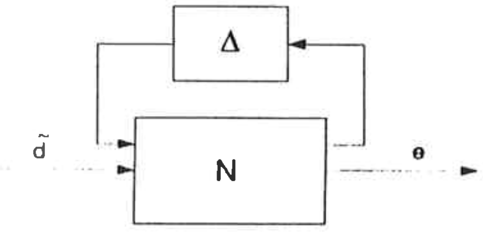

Figure 3: General structure for studying any linear control problem.

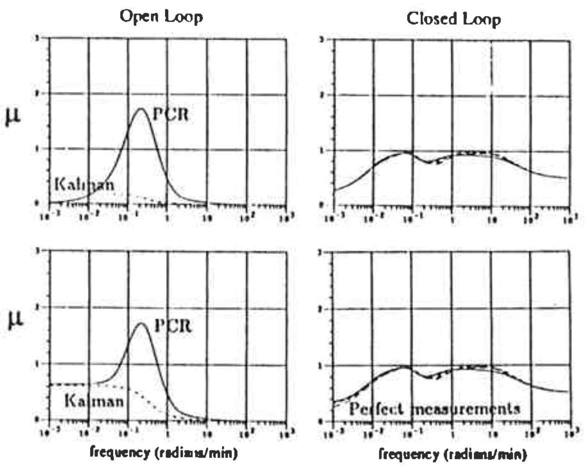

Figure 4: Comparison of Kalman (K1) and PCR estimator with 41 temperatures. Upper: without noise in μ-analysis, Lower: with noise.

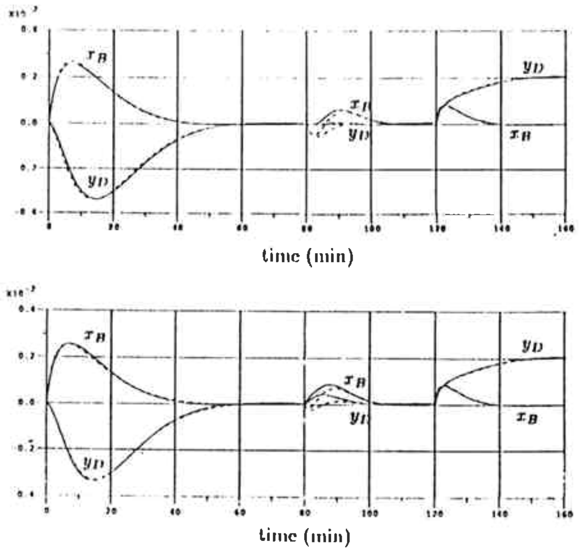

Figure 5: Comparison of output $y(t)$ (solid) and PCR-estimate $\hat{y}(t)$ (dotted line). Responses under feedback control are shown for a 20 % increase in feedrate at t=0, a 20 % increase in feed composition at t=80 min, and a setpointchange in y_D at t=120 min. Upper: y used for feedback control, Lower: $\hat{y}$ used for feedback control.

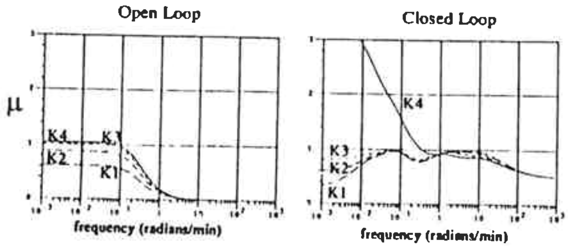

Figure 6: μ-plots for different Kalman filter gains (Table 2). With 41 temperatures and noise included in μ-analysis. Upper left: Nominal estimation error. Upper right: Robust control error. Lower left: Robust estimation error.

The PCR estimator in this paper uses only temperatures, but we did also evaluate the effect of adding inputs. However, the improvement in estimator performance was very small even at steady state. Furthermore, the dynamic behaviour of a static estimator is much worse when inputs are used.

5.2 Different Kalman filters and use of inputs in estimator.

Figure 6 shows μ-plots for the Kalman filters obtained using the four different levels of process noise on L and V in Table 2. The best Kalman filter, K1, is the one that was compared with PCR above. The remarkable thing with this best estimator is the very large assumed variance on the inputs u (L and V). In effect, this variance is so large that the transfer function from u to $\hat{y}$ in Eq.(12) is approximately zero, that is, the estimator does not use the information about the input signals.

The worst Kalman filter, K4, assumes disturbances (noise) of magnitude 0.1 for F and z_F, but assumes no disturbances on the inputs. This estimator performs reasonably well in the μ-test when there is no uncertainty (left part in Figure 6). However, it is extremely poor when input uncertainty is added (right).

5.3 Brosilow estimator.

The Brosilow inferential estimator for the system with different numbers of measurements is shown in Figure 7. It clearly demonstrates that the estimator as originally proposed performs poorly, and its performance does not improve with increasing number of measurements. The "Open-loop" test shows that the estimator nominally works well at very low frequencies ($\omega < 0.001$ min^{-1}). The poor dynamic performance (intermediate frequencies) is due to the fact that the estimator uses the input signals u (L and V) as shown in Eq. (16); the dynamic behaviour of u and the compositions y are very different and using a constant matrix $G_u - K_B F_u$ does not work well. This problem could have been corrected using a low-pass filter on the inputs with a large time constant, e.g., 194 minutes (that is, add dynamics to G_u and F_u). However, even this estimator would not perform well in practice, as the "Closed-loop" test shows that the robust performance is poor even at low frequencies. This is due to the input uncertainty, that is, the actual values of L and V are different from what the estimator thinks they are.

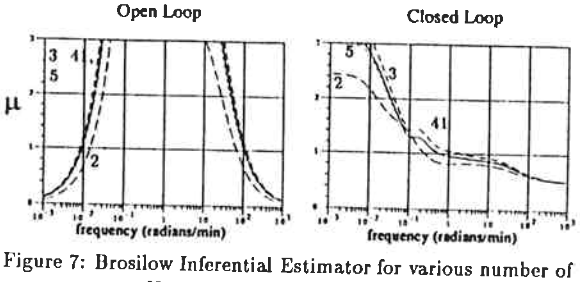

Figure 7: Brosilow Inferential Estimator for various number of measurements. No noise.

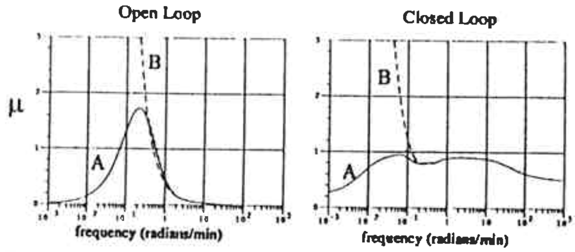

Figure 8: Modified Brosilow estimator based on temperatures only with 41 temperatures and no noise. A: Perfect model. B: Model when 1% random error is added to process matrix elements.

We therefore conclude that using the measured input signals u (which are inaccurate) does not improve the estimate. A better approach then seems to be to regard the inputs L and V as unknown disturbance together with z_F. This gives rise to the modified estimator $\hat{y} = K_{Bmod}d'$ where $d' = [L, V, z_F]^T$. This estimator performs much better as seen from curve A in Figure 8. The estimated values of the variables L, V and z_F may not be correct, but this error is not important as long as the estimate $\hat{y}$ is accurate. However, using L, V and z_F as disturbance variables has very poor numerical properties because G' and F' are ill-conditioned. For example, curve B in Figure 8 shows the drastic deterioration in performance caused by adding 1% random error to the elements of the matrices G' and F'.

5.4 Number of measurements.

The μ-plots in Figure 9 for the PCR estimator shows the effect of using varying numbers of measurements. Note that the noise in this case is put on the temperatures in *the calibration set* and not in the μ-analysis. Fig 9 demonstrates that adding temperature measurements improves the estimates and the control performance. The main difference is between two and three measurements. These results also applies for the Kalman filter. Another benefit of using many temperatures is that the performance becomes much less sesitive for measurement locations, which may be very important when dealing with 2 or 3 temperatures.

6 Discussion

Kalman filter.

Model uncertainty is not included explicitly when obtaining the Kalman filter and it may require physically unrealistic values of the noise weights, V and W, in order to obtain the best estimator when uncertainty is included. This is illustrated by the large value needed for noise (disturbances) on the inputs in order to obtain the best Kalman filter, K1. Otherwise, the Kalman Filter performed well in the μ-tests and was undoubtedly the best estimator in the open loop μ-test. The main reason is its inherent dynamic model. Furthermore, because of the weights, it is flexible, and it may be tuned to perform well for ill-conditioned plants a well. As mentioned above this

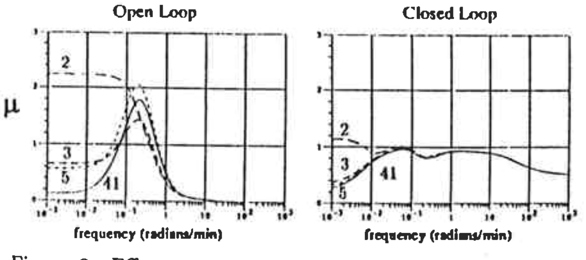

Figure 9: Effect on μ of number of temperatures for PCR-estimator. The temperatures in the calibration set are corrupted with 0.1 °C noise.

is done by adding (artificial) large noise (disturbances) on the inputs to the process.

Brosilow estimator.

As discussed above the Brosilow Inferential estimator as originally proposed suffers from four main weaknesses:

W1. *For ill-conditioned plants with large RGA-values input error causes poor estimates when the estimator uses information about the manipulated inputs u.*

W2. *Even for plants which are not ill-conditioned, the dynamic behaviour of a static estimator which directly uses inputs is often poor. The reason is the dynamic "lag" which usually exists between the inputs u and the outputs y.*

W3. *It does not handle collinearity among the variables in an appropriate way.*

W4. *For ill-conditioned plants (with large condition numbers in G or F) the use of inputs and disturbances as latent variables is numerically ill-conceived.*

Weakness W1 has already been discussed in general terms. Weaknesses W1 and W2 may be corrected using the "modified" Brosilow estimator, and also W3 may be corrected using an appropriate pseudoinverse of F_d. That is, instead of using only selected measurements as proposed by Joseph and Brosilow (1978), one should rather delete small *directions* in F_d using the singular value decomposition.

The key idea of the Brosilow Estimator is to first estimate the *independent* inputs that caused the observed outputs, and weakness W4 can not be corrected. Brosilows approach may be satisfactory in some cases, but not for ill-conditioned plants. For our example column the condition numbers of G' are F' are 165 and 321. This explains the sensitivity to small errors in the matrix elements in Fig.8 (it should be stressed that this numerical sensitivity is different from the sensitivity to input errors in W1).

However, even for ill-conditioned plants there may still be a rather simple direct relationship between various *dependent* variables, for example, between temperatures and composition in a distillation column, and a simple regression estimator, like PCR, between these variables may work well.

PCR estimator

The PCR-estimator does not have the same weaknesses as the Brosilow estimator. First, the estimator used here does not use the input values, and does not suffer from uncertainty with respect to their exact value (W1) and poor dynamic performance (W2). Second, and more important, its numerical properties are much better. The matrix to invert in PCR, the score matrix T in Eq. (23), is generally much better conditioned than F' used by the modified Brosilow estimator. For example, for our column the condition number of T is 4.7, whereas the condition number of F' is 321. To get a well-conditioned T one must ensure that excitations of the weak directions are included in the calibration set. To ensure such excitations, one should use data from the column with feedback (that is, with specified outputs), for example, by specifying the product compositions together with the feed composition in an factorial design. One should *not* use open loop data, like step responses etc., which will excite only the strong directions (The gain matrices in Brosilow's scheme will typically result from such excitations).

Obtaining and implementing the estimators.

Both the Kalman filter and the Brosilow estimator require a linear open-loop model. On the other hand, the PCR approach only deals with the data. This is an advantage, especially when experimental data are used, but also when we do have a good model, as in this paper, since we save a significant effort in obtaining the linear model matrices.

As for implementation, the static Brosilow and PCR estimators are of course much simpler than the dynamic Kalman filter.

NOMENCLATURE.

d - disturbances ($= [z_F, F]^T$ in most cases)
D - distillate flow rate
$\bar{d}$ - external inputs in μ-analysis
F - feed flow rate
F_u, F_d, F' - Gain matrices from inputs (u and d) to secondary measurements (θ)
G_u, G_d, G' - Gain matrices from inputs to primary outputs.
L - reflux flow rate
PCR - Principal Component Regression.
u - manipulated inputs ($= [L, V]^T$)
$\mathcal{V}$ - process noise covariance matrix.
V - boilup rate from reboiler
$\mathcal{W}$ - measurement noise covariance matrix.
w_i - input uncertainty weight
w_p - performance weight
x_B - mole fraction of light component in bottom product
y - primary output vector $= [y_D, x_B]^T$
$\hat{y}$ - estimated primary outputs
y_D - mole fraction of light component in distillate
z_F - mole fraction of light component in feed

Greek symbols

α - relative volatility
Δ - uncertainty block
μ - Structural Singular Value
θ - secondary measurements (temperature vector)
$\bar{\theta}$ - vector of all available information
Θ - data matrix of θ

REFERENCES

Doyle, J. C., 1982, "Analysis of Feedback Systems with Structured Uncertainties", *IEEE Proc.*, **129**, D, 242–250.

Joseph, B., C. B. Brosilow, 1978, "Inferential Control of Processes. Part I. Steady State Analysis and Design.",*AIChE Journal*, **24**, 485–492.

Kalman, R. E., and R. S. Bucy, 1961, "New results in Linear Filtering and Prediction Theory", ASME, *J. Basic Eng.*, **83**, 95–108.

Keller, J.P, and D. Bonvin, 1987, "Selection of inputs for the purpose of model reduction and Controller Design", Presented at the 10^{th} *World Congr. on Aut. Cont. (IFAC)*, Munich.

Mejdell and Skogestad, 1989, "Estimate of process outputs from multiple secondary measurements",*ACC–1989*, 2112–2121.

Moore, C., Hackney, J. and Canter, D., 1987, "Selecting Sensor Location and Type for Multivariable Processes" *Shell Proc. Contr. Workshop*, Butterworth Publishers, Boston.

Morari, M. and Stephanopoulos, G., 1980, "Optimal Selection of Secondary Measurements within the Framework of State Estimation in the Presence of Persistent Unknown Disturbances", *AIChE Journal*, **26**, 247–259.

Patke, N. G., P. B. Deshpande and A. C. Chou, 1982, "Evaluation of Inferential and Parallel Cascade Schemes for Distillation Control", *Ind. Eng. Chem. Process Des. Dev.* **21**, 266–272.

Skogestad, S. and M. Morari, 1987 "Implication of Large RGA-Elements on Control Performance", *Ind. & Eng. Chem. Res.*, **26**, 11, 2121–2330.

Skogestad, S. and M. Morari, 1988, "LV-Control of a High-Purity Distillation Column", *Chem. Eng. Sci.*, **43**, 1, 33–48.

Skogestad, S., M. Morari and J.C. Doyle, 1988, "Robust Control of Ill-Conditioned Plants: High-Purity Distillation", *IEEE Automatic Control*, **33**, 12, 1092–1105.

Strang, G., 1980, "Linear Algebra and its Applications", Harcourt Brace Jovanovich, New York.

Weber, R. and Brosilow, C., 1972, "The Use of Secondary Measurements to Improve Control", *AIChE Journal*, **18**, 614–623

Wold, S., Espensen, K. and Geladi, P., 1987, "Principal Component Analysis", *Chemometrics and Intel Lab Syst*, **2**, 37–52.

Yu, C. C. and W. L. Luyben, 1987 "Control of Multicomponent Distillation Columns using Rigorous Composition Estimators." *I. Chem. E. Symp. Ser.*,**104**, 29–69.