Presented at ACC'89, Pittsburgh, June 21-23. 1989
Session: Large-Scale Systems

Estimate of process outputs from multiple secondary measurements

Thor Mejdell

Sigurd Skogestad

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

Abstract

Temperature and flow measurements are used to estimate the product compositions in a distillation column. The problem is characterized by strong colinearity (correlation) between the temperature measurements. Contrary to some claims in the literature, it is found using a Kalman-Bucy Filter that the goodness of the estimate, even when used for feedback control, is improved by adding temperature measurements. This does not apply to Brosilows inferential estimator which in its original form is very sensitive to colinearity in the measurements. It is important to use only those directions in the measurement space which are excited by the independent variables (inputs and disturbances). The Partial Least Square Regression (PLS) method used in statistics adresses this explicitly. In the paper we use the PLS method to gain insight into the directions of the temperature space.

1 Introduction

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

The use of temperature measurements for feedback control of distillation columns is quite extensively discussed in the chemical engineering literature (eg., Nisenfeld and Seeman [1], p. 85-95). Temperatures are not used because they are of interest themselves, but because composition measurements are often expensive and unreliable and often not available, while temperature measurements are inexpensive and reliable. One problem is that temperature is a true indicator of the tray composition only if the mixture is binary and at constant pressure. However, this problem may be partly overcome by using several measurements.

Measurement selection. Most columns have temperature sensors located at about every fifth tray in the column, that is, a typical column may have 5-10 temperature measurements. In industry all these measurements are rarely used. Usually each composition measurement is replaced by a single temperature measurement and used for single-loop feedback control. The main problem is then to find a suitable location for this temperature. According to Nisenfeld and Seeman [1] the most important issues are, i) that the temperature should be sensitive to changes in the composition, and ii) that the correlation between temperature and composition should be insensitive to disturbances in feed composition and in flows. Since the products are often very pure the first criteria favors placing the temperature sensor away from the products. The second criteria favors placing the sensors close to the product. Another factor favoring placing the sensor away from the column ends is nonlinearity caused by changes in operating conditions. However, note that nonlinear effects may alternatively be counteracted by using the logarithm of the impurity for feedback instead of its absolute value.

In this paper, measurement location is not an important issue. The reason is that we use several (typical five or more) temperature measurements and then estimate the product compositions. In this case the exact location is far less important than when single temperature measurements are used.

The estimation problem. The objective is to obtain the best estimate of the product compositions using all available measurements. This estimate should be obtained based on a description of the process (nominal model and expected uncertainty), the expected disturbances, and a more precise definition of what we mean by "best". It is assumed that the estimated outputs are to be used for feedback control.

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" controller which uses these estimates for feedback control (Fig. 1). In general, this solution is suboptimal

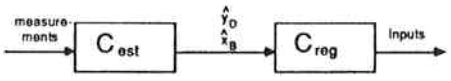


Figure 1. Controller block $C = C_{reg}C_{est}$ split in separate blocks for estimation and control.

compared to using one big controller which directly uses all available measurements. The reason is of course that some information is lost as the original measurements are condensed into the fewer estimated variables. In some cases it may be shown that no information is lost and this is then referred to a separation principle. In particular, this may be the case if all the states of the system are estimated since the states contain all information about the system at any given point in time. However, in this paper we will not use all states for feedback control and the separation principle does not apply. However, since we in our case are estimating the actual controlled outputs, we may postulate that the performance loss caused by the separation is not a major problem. However, this point remains to be proved rigorously.

The motivation for separating the controller in two parts in our case is reliability, design simplicity and robustness. In particular this is the case if the "small" controller is based on single loops. Reliability then means that we can still control one composition even if the other loop fails. Design simplicity means that the control system is build up of smaller pieces which each perform a specific task which is relatively easy to understand. By robustness we mean robust in a broad sense, both with respect to model-plant mismatch, but also robustness with respect to the accuracy of the problem definition itself. This last point may seem somewhat philosophical, but is often of great practical value. For example, suppose you are to design a multivariable controller for a very ill-conditioned plant (with large RGA- values [2]). To avoid "stupid" controllers such as decouplers you have to add specifications such as uncertainty or limitations on the input signals. However, if the controller structure is limited to be diagonal (decentralized control) then there is no possibility for doing stupid things (simply because the controller structure is so limited) and your result in terms of optimal tunings is only weakly on the specifications mentioned above. The same applies if one wants to design a "big" controller using all measurements: to get a meaningful solution from an optimization procedure you have to be extremely careful with the problem formulation, that is, the solution is not robust with respect to the problem formulation.

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 Brosolow [3], Morari and Stephanoplous [4], Yu and Luyben [5], Moore et al.[6], Keller and Bonvin [7] have suggested that one should only use a few of the temperature measurements to avoid the badly conditioned problem of obtaining information from the strongly correlated temperatures. The reasons for this are again related to the somewhat philosophical points raised above about sensitivity to the problem descriptions. By not using all the measurements the solution space is limited and you do not need to define the problem carefully in order to avoid "stupid" solutions. One of the main objectives of this paper is to study some of the available methods for designing estimators to find out if they are sensitive to the kind of colinearity which we intuitively expect may cause control problems. In particular, we expect the Brosilow Inferential Estimator [8] to be sensitive since the users of this method strongly argue against using too many measurements.

We consider three different approaches to the estimation problem: i) The Kalman-Bucy Filter, ii) Brosilows Inferential Control Method, and iii) Partial Least Square Regression.

Kalman estimator. The traditional "optimal" approach of modelling disturbances and noise as stochastic processes and minimizing a quadratic error function gives rise to the Kalman-Bucy filter. This estimator contains a full model of the plant, and the states are updated by using constant gain feedback from the measurements. This algorithm does not seem to have been widely used in industry for estimation of compositions, but this may partly be caused by the need for computing power which previously made it impractical.

Brosilow estimator. In process control, Weber and Brosilow [8] introduced a different method for use of secondary measurements. The main idea is that disturbances in process control applications tend to vary slowly compared to the process dynamics, and are certainly not well modelled by stochastic disturbances. Brosilow proposed to use the secondary measurements to estimate the disturbances. These estimated disturbances are then assumed to be constant in the future and the disturbance estimates are used in a sort of feed-forward scheme to counteract their expected on the outputs. We shall only use the estimator part of Brosilows scheme and not the feedback part. Brosilows scheme has a strong intuitive appeal and seems to have found some use in industry. However, the method appears to be very sensitive to the number of measurements if these are colinear (eg., Joseph and Brosilow [3].

To understand what the Brosilow estimator does consider the simplified steady-state estimation problem with constant manipulated variables: Given a set of (secondary) measurements θ, obtain an estimate $\hat{y}$ of the the outputs y,

$$\hat{y} = K\theta \tag{1}$$

where K is the gain matrix of the estimator. To obtain K Brosilow considers the model

$$\theta = F_d d, \quad y = G_d d \tag{2}$$

where d contains all independent variables (disturbances), and then obtains

$$K = G_d F_d^{\dagger} \tag{3}$$

where the latter denotes the pseudo inverse of F_d. The pseudo inverse arises as the best estimate in the least square sense.

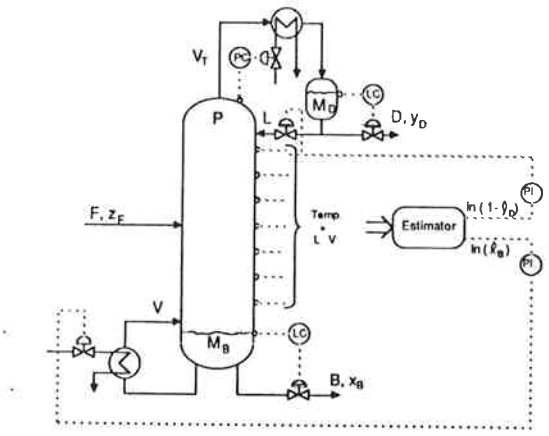


Figure 2. Control scheme based on LV configuration.

PLS estimator. A more direct way of obtaining K, without going the way through estimating disturbances, is inspired by the approach used by analytical chemists in their "multivariable calibration problem" [9]. In this case one first obtains sets of corresponding values of Y and Θ (The matrices Y and Θ contain corresponding values of y and θ obtained from "calibrations" or simulations). A SVD on Θ is performed to delete directions in Θ which are only weakly excited and then one computes

$$K = \Theta^{\dagger} Y. \tag{4}$$

Using this approach one avoids explicitly modelling the effect of the independent variables on θ and y (although typical variations should be included in the calibration set). The approach also has an intuitive appeal to engineers as one seems to skip the modelling step. Physical insight may be obtained by studying the directions of the matrix Θ using SVD.

Analysis of estimators. We shall compare these estimators on the basis of robust performance. This involves computing the structured singular value μ [10]. The performance requirement is satisfied for the worst case if μ is less than one at all frequencies. The composition estimates are assumed to be used as inputs to two PID-controllers using the LV-configuration as shown in Fig. 2. The PID-controllers were tuned such that robust performance was achieved when the estimate is exact. The estimators are compared under the following two conditions:

1. Open-loop estimation. In this case feedback from the true compositions is used, and we compare based on the estimation error $y - \hat{y}$ (Fig. 4).

2. Closed-loop estimation. In this case the estimates are used for feedback control and we compare based on the control error $y - y_s$, where y_S is the setpoint for y (Fig. 5).

Example column. As an example column we use column A studied by Skogestad and Morari [11]. This is a column with 41 stages, including reboiler and 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. Because the mixture is binary and pressure is assumed constant, there is an exact relation between this composition and the temperature on the tray. We compute the temperature using Raoults law and Antoine parameters for pure component vapor pressures given in Table 1.

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

- Feed is liquid.
- Constant molar flows.
- Constant pressure 1 atm.
- Holdup on each tray; $M_i/F = 0.5$ min

Antoines parameters:

Light component			Heavy component		
15.8366	2697.55	-48.78	15.4311	2697.55	-48.78

Table 1. Data for distillation column example.

2 Estimation Methods

2.1 Brosilow estimator.

The following linear steady-state model of the column in terms of deviation variables is used

$$y = G_d d + G_u u \tag{5}$$

$$\theta = F_d d + F_u u \tag{6}$$

Here the dependent variables are the outputs y (compositions) and secondary measurements θ (temperatures). The independent variables are the disturbances d (F, z_F), and manipulated inputs u (L,V). Introduce new variables y_d and θ_d for the effect of disturbances on y and θ.

$$y_d = y - G_u u = G_d d \tag{7}$$

$$\theta_d = \theta - F_u u = F_d d \tag{8}$$

The disturbances are estimated from the measurements θ using a least square estimator (pseudoinverse of F_d). The estimated effect of these disturbances on the outputs are then

$$\hat{y}_d = K_B \theta_d \tag{9}$$

$$K_B = G_d F_d^T (F_d F_d^T)^{-1} \tag{10}$$

and the estimate of the outputs becomes

$$\hat{y} = (G_u - K_B F_u) u + K_B \theta \tag{11}$$

In this paper we use the estimate in eq. 11 for feedback control using two PID controllers. The estimator applies also in the dynamic case provided the dynamics of disturbances and inputs are equal. This is reasonable to some extent for distillation columns since the dynamics are dominated by a time constant τ_1. Brosilow uses the estimate in eq. 11 combined with a kind of "feedforward" scheme to keep $\hat{y} = 0$. This gives rise to the inferential controller, but this is not used here.

2.2 Partial Least Square Regression.

We want to estimate p outputs (compositions y) from q known variables (temperatures θ, inputs u). Denote the known variables as the extended measurement vector θ_e. The problem is then to obtain the matrix K in

$$y = K \theta_e \tag{12}$$

To this end obtain n "calibration" sets of corresponding values of y and θ_e. In this paper they are obtained from the nonlinear steady-state column model. Place the n calibration sets of y and θ_e as rows in the matrices $Y^{n \times p}$ and $\Theta^{n \times q}$, respectively. We then have

$$Y = \Theta K^T \tag{13}$$

The ordinary least square solution for K is:

$$K_{LS} = Y^T \Theta [\Theta^T \Theta]^{-1} \tag{14}$$

The matrix $\Theta^T \Theta$ is n times the sample covariance matrix of the measurements θ_e, and may be singular or nearly singular if strong colinearity in the temperatures exists. This will generally be the case in a column with measurements located close to each other.

To avoid this difficulty transform the measurement variables to their principal components (singular value decomposition)

$$\Theta = t_1 p_1^T + t_2 p_2^T + \cdots + t_m p_m^T \tag{15}$$

where $m = min(n, p)$. Here p_1 is the eigenvector corresponding to the largest eigenvalue of $\Theta^T\Theta$, (or equivalently the square of the largest singular value), and p_2 is the eigenvalue corresponding to the second largest eigenvalue and so on. The loading vectors (p's) give the directions of the principal components (factors) , while the scores (t's) give the magnitude. Select only the principal components that can be distinguished from the measurement noise, i.e. the first k components, and let the matrices $P^{q\times k}$ and $T^{n\times k}$ include only these k most important directions. The least square solution to $y = K_t t$ becomes

$$K_t = Y^T T[T^T T]^{-1} \tag{16}$$

and since $t = P^T\theta_e$ (P is orthogonal), the K in eq. 12 becomes

$$K = Y^T T[T^T T]^{-1} P^T \tag{17}$$

Another version of this method is to also take into account the directions in Y when finding the approximate pseudo inverse of Θ (partial least square, PLS). The main difference is that one computes the largest eigenvalues from $\Theta^T Y Y^T \Theta$ in stead of $\Theta^T\Theta$. This takes in account the directions in Θ which have the largest covariance with y, and thus ensure that these directions are not deleted.

2.3 Kalman filter.

This is the standard estimator used in stochastic optimal control. A dynamic model is used in parallel with the process itself, and the deviation between the output from the process and that of the model is used as feedback to the model through a constant filter gain K_f. The linear state space model is

$$\dot{x} = Ax + Bu + Ev \tag{18}$$
$$y = Cx \tag{19}$$
$$\theta = C_\theta x + w \tag{20}$$

Here x is the state vector, v the process noise (eg. to represent disturbances), w the measurement noise, and the other variables u, y and θ are as defined before. v and w are supposed to be white noise processes.

The estimator then becomes

$$\dot{\hat{x}} = A\hat{x} + Bu + K_f(\theta - C_\theta\hat{x}) \tag{21}$$
$$= (A - K_f C_\theta)\hat{x} + Bu + K_f\theta \tag{22}$$

The calculation of the filter gain K_f is based on modelling the covariance matrices of v and w, $\mathcal{V}$ and $\mathcal{W}$.

$$K_f = \mathcal{X} C_\theta^T \mathcal{W}^{-1} \tag{23}$$

where $\mathcal{X}$, the covariance matrix of x, is found from the matrix Riccatti equation

$$\dot{\mathcal{X}} = A\mathcal{X} + \mathcal{X}A^T - \mathcal{X}C_\theta^T\mathcal{W}^{-1}C_\theta\mathcal{X} + E\mathcal{V}E^T \tag{24}$$

3 Estimators for the example column.

In this section we describe how the different estimators were obtained for the example column with 41 stages.

3.1 Brosilow estimator.

The matrices F_d,F_u,G_d and G_u in equations 5 and 6 are found by using the steady state values of a linearized model in the operating point.

3.2 PLS-estimator.

The Partial Least Square estimator was found by selecting 16 different sets of the outputs x_B and y_D and the feed composition z_F (disturbance) and use a nonlinear column model to caculate the steady state temperature profile. Note that it is not necessary to simulate different feed rates using this approach. The 16 values are listed in Table 2. The data was spread with equal distances around z_F, but with logarithmic equal distances for x_B and $(1 - y_D)$. This is to give a better balanced calibration set for this kind of nonlinear high-purity column. The operating point data from the first run in the table was used as basis for making deviation variables of the remaining 15 runs. The temperature data was then reduced to 3 factors (directions), and K_{PLS} was computed.

z_f	y_d	x_b	z_f	y_d	x_b
0.5000	0.9900	0.0100	0.4875	0.9962	0.0189
0.5375	0.9913	0.0262	0.4750	0.9956	0.0087
0.4250	0.9738	0.0151	0.5625	0.9934	0.0115
0.5250	0.9700	0.0132	0.4625	0.9772	0.0300
0.4125	0.9801	0.0058	0.4375	0.9950	0.0038
0.6000	0.9849	0.0044	0.4500	0.9924	0.0173
0.5125	0.9942	0.0066	0.5750	0.9868	0.0228
0.5500	0.9827	0.0076	0.5875	0.9885	0.0050

Table 2. Data to simulate stationary temperature profile. See also Table 1.

3.3 Kalman estimator.

The Kalman filter gain was found by assuming the covariance matrix of the measurement noise $\mathcal{W} = 0.04I$. I is the identity matrix. This indicates that we model the measurement uncertainty to be 0.2 °C. For the process noise v, defined as $v^T = (LVFz_F)$ (reflux, boilup, feedrate and feed composition), we have used four different covariance matrices $\mathcal{V}$. This gave rise to the four different filter gains in table 3 denoted K1 to K4. In K4 we have assumed a standard deviation of 10% on all input variables. The three others have larger values to compensate for model uncertainty and unknown disturbances. The difference between them is how they model the disturbance on the inputs.

The assumption of white noise disturbances is of course rather doubtful, but the estimator is not expected to be very sensitive to this aproximation. Nevertheless, it is clear from the above more or less arbitrary choices of weights for $\mathcal{V}$ that it is not at all clear how to set up the problem to get a reasonable estimator. This is of course one the main disadvantages with this approach.

Gain	ν
K1	0.4 diag{0.5 0.5 1.0 1.0}
K2	0.4 diag{1.0 1.0 1.0 1.0}
K3	0.4 diag{0.0 0.0 1.0 1.0}
K4	0.01 diag{1.0 1.0 1.0 1.0}

Table 3. Process noise covariance matrix for different Kalman filter gains.

3.4 Number of measurement and their locations

The estimation methods above were applied to different numbers and locations of temperature measurements. Three cases were considered: i) all 41 temperature measurements, ii) five measurements with two options (**A** and **B**) for their location, and iii) two measurements on trays 10 and 30 (only for the Brosilow estimator). In case ii) we used the loading plot from the PLS-calibration, see fig. 11, to select the best location for set **A**. This gave trays 10,15,22,29, and 33 (reboiler is no. 41). Set **B** corresponds to a more arbitrary equal distance selection using trays 1,12,21,30 and 41.

4 Mu-analysis of the Estimators.

The objective is to evaluate the different estimation methods described above. In this section we set up our problem and derive criteria for the evaluation.

4.1 μ-analysis.

Our tool is the Structural Singular Value (μ) analysis. In this framework we rearrange our system to fit the general form shown in fig. 3. Here M denotes the generalized nominal plant including the plant and the weights, d is disturbances and setpoint changes, e is the "error" we want to keep small. We have one Δ-block loop, which represent the model uncertainty, and one controller loop . The μ analysis is to evaluate the maximum amplification from d to e at each frequency. To make any sense all values, i.e d, e and Δ, must be scaled properly, that is, scaled to be less than 1. The requirements for the system, defined by the weights (scaling), is satisfied if μ is less than one at all frequencies. For more exact definitions of μ see Doyle [10].

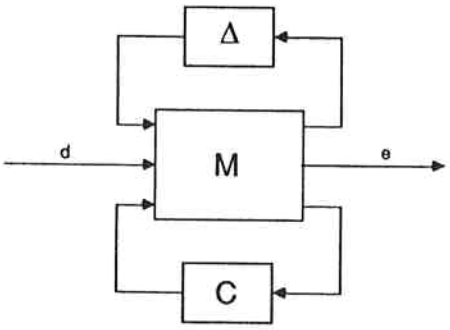


Figure 3. General structure for studying any linear control problem.

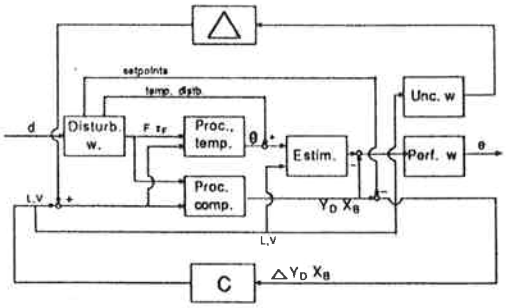


Figure 4. Block diagram for "open-loop" test.

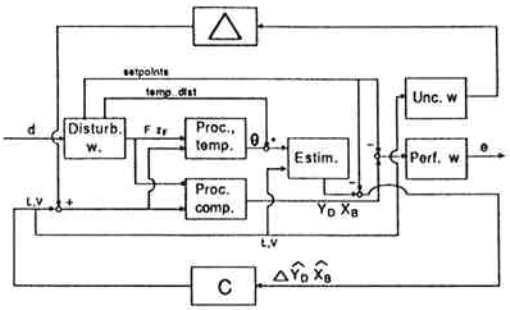


Figure 5. Block diagram for "closed-loop" test.

4.2 Evaluation criteria

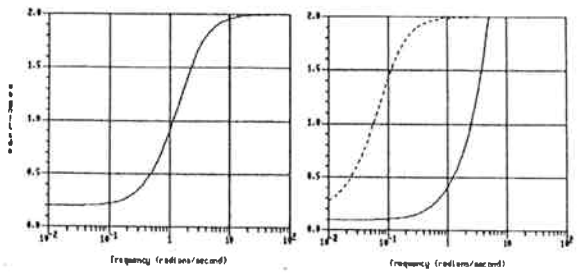
One obvious criteria for evaluating the different estimators is their ability to follow the true composition value. The "error" e in fig. 3 then becomes the difference between the real value (y) and the estimated one ($\hat{y}$). We let the system have feedback, since this is more close to a real situasion than a pure open loop test, but the controller will use the actual y, and not the estimate as input. Our test will then be independent of the controller used. We call this evaluation test "Open Loop" because the estimator is not connected to the feedback system. See fig. 4.

The ultimate goal is however the performance of the *total* feedback system, i.e it's capability of holding the specifications. The "error" to be minimized is now the difference between y and y_{set}.(Fig. 5). This "Closed loop" test implies that we have to choose a suitable controller that is fixed for all estimators. For reasons outlined in the introduction, we choose a PID controller. The PID-parameters was found by a μ-optimization of the system in fig. 5 without estimator. One main drawback of this test is however that it is not independent of the controller chosen.

4.3 Uncertainty weights.

The most important source of uncertainty is assumed to be on the inputs L and V. We shall use the same uncertainty weigtht as Skogestad and Morari [11], which is given by

$$w_I(s) = 0.2\,\frac{5s+1}{0.5s+1} \qquad (25)$$



a b

Figure 6. a) Uncertainty weight. b) Inverse performance weights. Solid line: Open loop, dotted line: Closed loop.

The weight is shown graphically in fig 6a. In the low frequency range it allows for a 20% uncertainty in flow *changes* (L and V are deviation variables), due to the inaccuracy of valve settings. The uncertainty increases at higher frequencies, reaching a value of 100% at about $\omega = 1$ min^{-1}. The increase at high frequencies will take care of neglected flow dynamics in the model, and it will allows for a time delay of about 1 min between L and V and the outputs y_D and x_B.

4.4 Performance weights.

In the "Open Loop" test we use the following performance weight

$$w_p = \frac{10}{4s+1} \tag{26}$$

which is shown in fig 6b. This weight will require less than 10% deviation of the estimates $(1 - \hat{y_D})$ and $\hat{x_B}$ at steady-state ($\omega \leq 0.1$ min^{-1}). At higher ω the weight increases to 1 at $\omega = 2.5$ min^{-1}. This allows a deviation $y - \hat{y}$ greater than 100% at frequencies above 2.5 min^{-1}. This is not unresonable, since the two dominant time constants of the column are 194 min and 15 min. [11]

In the "Closed Loop" test we have chosen a performance weight

$$w_p = 5\frac{10s+1}{100s+1} \tag{27}$$

This impies that the deviation of $y - y_{set}$ should be within 20% at steady state, i.e we tolerate a deviation of the product composition of about 0.2 mole%. Our feedback system should be effective up to about $\omega = 0.05$ min^{-1} and the amplification at high frequencies should never exceed 2.

4.5 Disturbances.

The external inputs to the systems (the d's) consists of both setpoints and ordinary disturbances. They may have any frequencies, but are normalized specifying their maximum values.

The maximum setpoint changes are set to 100% of x_B and $(1-y_D)$. Since the operating point is 0.01 and 0.99 this implies that the x_{Bset} may vary from 0 to 0.02, and y_{Dset} from 0.98 to 1.0. The disturbances in the feedrate F are set to 20% and in the feed composition to 10%, i.e. it may vary from 0.4 to 0.6 in mole fraction.

For some tests we also used disturbances on the temperature measurements. The disturbances were then set to 0.2 °C on all temperatures.

5 Results.

5.1 Insights in the colinearity using PLS.

The elements in the matrix K for the calibration with 41 temperatures, is plotted in fig. 7 below. The first 2 values are the streams L and V, and the others (3 - 43) are the temperatures on each stage. The reboiler is no 43. In fig. 8 we see how the variances $\Theta^T\Theta$ are explained with increasing number of factors (principal components). After 3 factors almost all (98%) of the variation in θ_e is explained.

Fig. 9 shows the same thing for the output y. Here three factors explain 94 % and four 97 % of the variance in Y. Using 3 or 4 factors depends on the magnitude of the noise in the

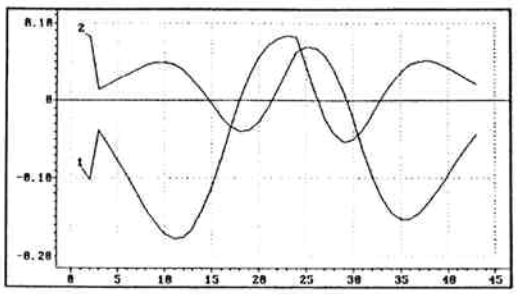


Figure 7. Elements in matrix K. Abscissa: θ_e, 1) elements for y_D, 2) elements for x_B.

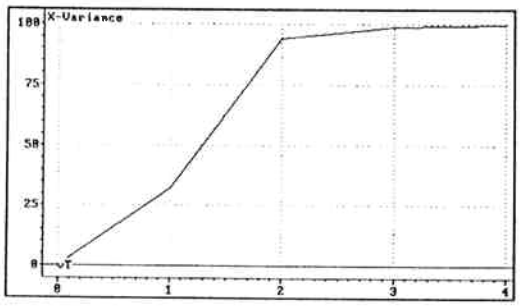


Figure 8. Explained X-variance (%). Abscissa: number of factors.

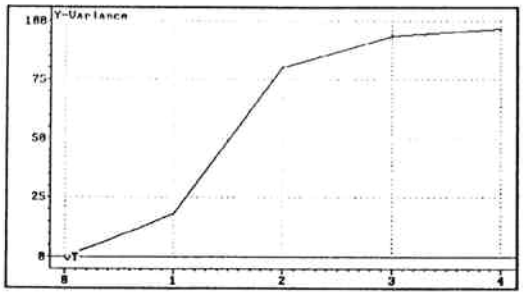


Figure 9. Explained Y-variance (%). Abscissa: number of factors.

temperatures. Fig. 10 displays the variance in Y when the same set of data was corrupted with 0.2 °C of white noise. There is no longer any improvement from factor 3 to factor 4. This indicates that there are hardly more than 3 different directions in the dataspace of the pressure compensated temperatures and input streams. In fig. 11 the loading vetors are displaid, i.e how the different measurements are summed up to make the factors. It shows that there are some temperatures that are more important than others. We used this information to select the measurement set **A**. From the figure we also see that the temperatures near the product streams contain little information about the composition. (They will only be useful for detecting pressure variations).

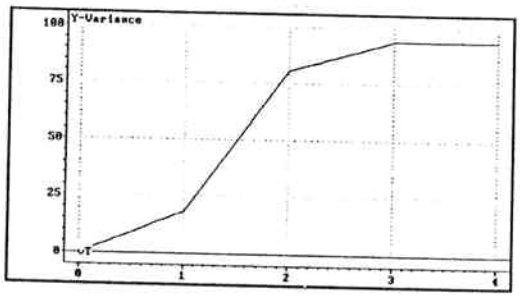


Figure 10. Explained Y-variance (%). Data corrupted with 0.2 °C noise. Abscissa: number of factors.

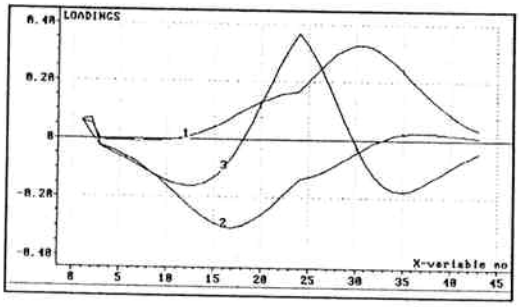


Figure 11. Loading plot. Abscissa: θ_e, curve identifier: factor number.

5.2 Comparison of estimators using 41 Temperatures.

In fig. 12 we compare the μ-plots for the Kalman and the PLS estimators. The Kalman filter gain K1 is given in table 3. For closed loop both estimators hold the specifications ($\mu \leq 1$), but the Kalman estimator is better at steady state. It is as good as with perfect measurements. In the "Open loop" test this difference is still larger. Contrary to PLS, the Kalman estimator is dynamic, and has a model available, so this is not surprising.

Fig. 13 shows the estimators when the column no longer is operating at $x_B = 0.01$, $y_D = 0.99$ and $z_F = 0.5$ (Operating point α), but at 0.03, 0.97 and 0.45 respectively (Operating point β). Although none of them reach the specifications, PLS is better than Kalman. The new operating point is within the calibration set of PLS, and being there it is not that sensitive to model changes as Kalman, which is developed around the first operating point.

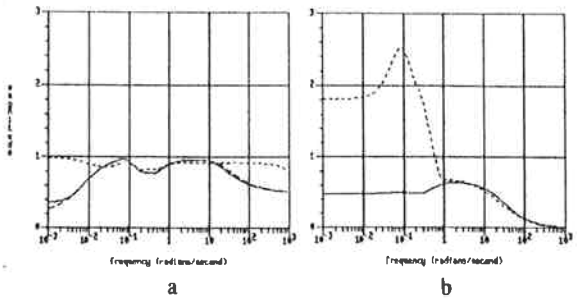


Figure 12. μ-plots for Kalman and PLS- estimators. Normal operating point α. a) Closed loop b) Open loop. Solid line: Kalman, short dotted line: PLS, long dotted line: No estimator.

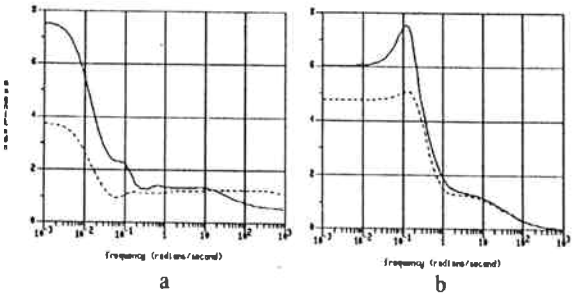


Figure 13. μ-plots for Kalman and PLS- estimators. Operation point β. a) Closed loop b) Open loop. Solid line: PLS, dotted line: Kalman.

5.3 Comparison of estimators using 5 Temperatures.

Fig. 14 shows the Brosilow estimator for 41, 5 and 2 temperatures. It demonstrates clearly that the estimator cannot handle many temperatures. Only the estimator with 2 temperatures (min for 2 points control) is in the neighborhood of the specifications. The set **B** is "better" than **A** because it has more distance between the measurements, and this gives rise to less colinearity.

In contrast to the Brosilow estimator, the PLS and the Kalman estimators are improved with increasing number of measurements (fig. 15). The temperature set selected with PLS, i.e. **A**, is slightly better. Nevertheless, the performance seems not to be very sensitive to the selection when operating with 5 temperatures.

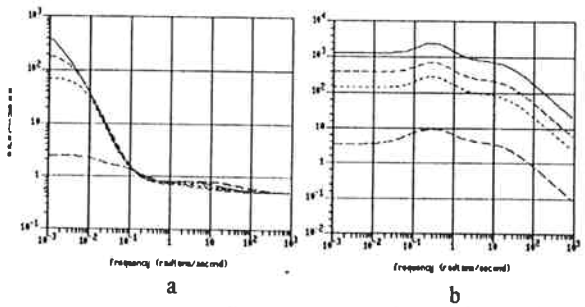


Figure 14. μ-plots for Brosilow estimator. a) Closed loop b) Open loop. Solid line: 41 temp, upper dotted line: 5 temp. selection B, short dotted line: 5 temp. selection A, lower dotted line: 2 temp.

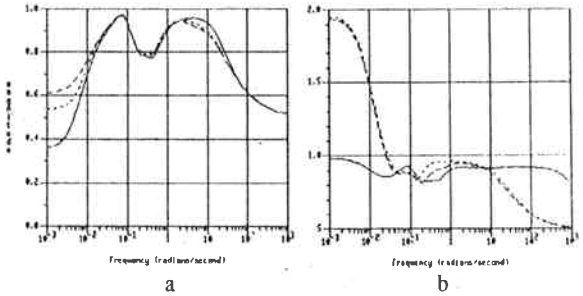


Figure 15. μ-plots for Kalman and PLS estimators, closed loop. a) Kalman, b) PLS. Solid line: 41 temp., short dotted line: 5 temp selection A, long dotted line: selection B.

5.4 Different Kalman Filters Gains.

The performance of the different Kalman filter gains K1, K2 and K3 from table 3 are shown in fig. 16. for 41 temperatures. The difference in performance is caused by the difference in modelled variances of the disturbances on L and V. K1 and K2 has almost equal performance. On the other hand, K3, with zero variance on the inputs, is significantly worse. This demonstrates the importance of modelling the input streams.

Fig. 17 shows the difference between the estimators when the relationship between $\mathcal{V}$ and $\mathcal{W}$ is changed. Since $\mathcal{V}$ and $\mathcal{W}$ are diagonal matrices it is the relationship between them and not their magnitude that is important. Compared to K1, K4 depend less on the measurements and more on the model. The figure shows that K1 is the best when both the measurements and the model are perfect. If the operating point is changed from α to β one should expect that K1 is still better than K4. In fig. 18 we see that this actually is the case, but the differences are quite small.

On the other hand, if measurement disturbances are present, one should expect that the estimator which depends less on the measurements are best. In fig. 19 this is shown. Here the improvement is quite clear.

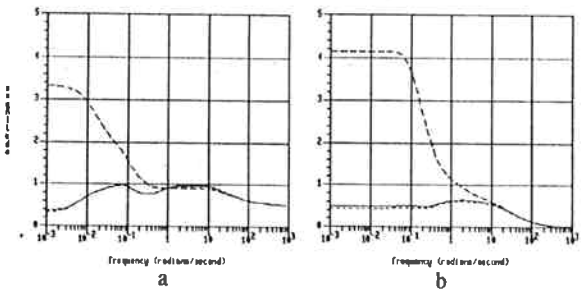


Figure 16. μ-plots for Kalman estimators. a) Closed loop b) Open loop. Solid line: K1. Short dotted line: K2. Long dotted line: K3.

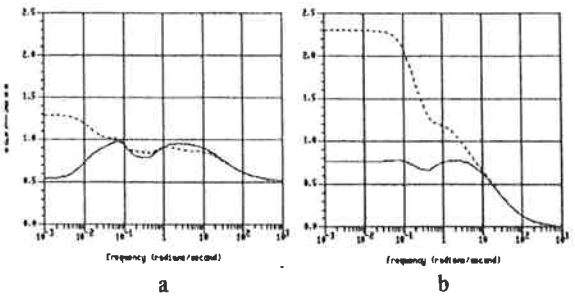


Figure 17. μ-plots for Kalman estimators, 5 temperatures. a) Closed loop b) Open loop. Solid line: K1, dotted line: K4.

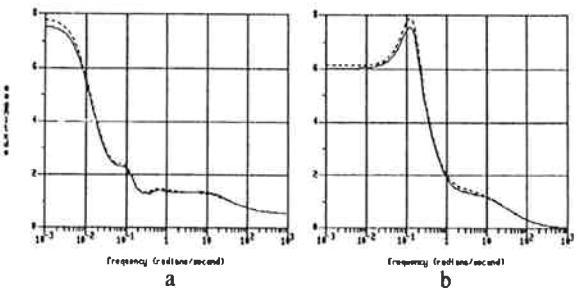


Figure 18. μ-plots for Kalman estimators. Operating point β, 41 temperatures. a) Closed loop b) Open loop. Solid line: K1. Dotted line: K4.

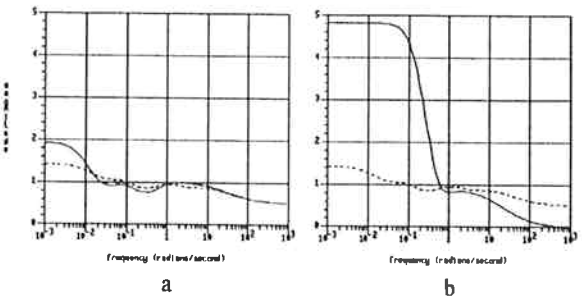


Figure 19. μ-plots for Kalman estimators. Disturbances on the temperatures. a) Closed loop b) Open loop. Solid line: K1, dotted line: K4.

6 Discussion

Kalman-Bucy Filter. For most test cases the Kalman-Bucy Filter appears to be the estimator with the best performance. The main reason is its inherent dynamics, which is a great advantage unless the disturbances is very slow compared to the process. In addition to using all the measurements, it takes advantage of the information incorporated in a mathematical model. However, this gives a tuning problem. Depending of the relative goodness of the model and the measurements, one have to tune the filter gains to make them useful. Although the μ-analysis have proved to be useful for tuning these gains, the tuning is not a straight forward process.

The results show also the importance of including the uncertainty of the manipulated variables in the process noise description, and not lump them together with the unmanipulated ones, which is normally done.

One apparent drawback of the estimator is it's dependence of the operating point. This is due to the strong nonlinearity in the system. For compensating this one may use the extended Kalman filter with on-line updating and calculations of the filtergains. For distillation columns this will rapidly give a heavy computer load, caused by the dimension of the state space involved. For example, for a n-component destillation the state space description will have a dimension n-1 times the number of temperatures used. An alternative way to counteract for the nonlinearity is to express the operating conditions in logarithmic compositions [3, 11], also for the estimator, that is, estimate $\ln x_B$ and $\ln(1 - y_D)$ instead of x_B and y_D. We will then involve scaled gains in the state space model. The transformation to logarithmic compositions will be a useful improvement for the other estimators too.

The PLS estimator. The PLS regression is only using steady state information and this is perhaps its main drawback. To introduce dynamics into the method is not trivial, but may be a fruitful area for future research.

Nevertheless the method provide insight into the estimation problem. It gives information about the main dimensions in the temperature space and the location of the most influential measurements. For example, the variance plot shows that using tree factors, the temperatures can explain 94% of the changes in the compositions y. This will be a measure of the maximum prediction ability of the temperatures, or how well conditioned the estimation problem is.

PLS estimator will be an interesting alternative to Kalman Filter in those cases where the disturbances are slow, and a static estimator will be sufficient. Especially, if the process also is difficult to model, or the modelling becomes very rigorous, as will be the case for some multicomponent distillations columns. This "black box" method has the advantage that one may skip the modelling part and use the data directly. It also has an inherent pressure compensation, if different pressure levels are included in the calibration data set.

We have shown that the estimation improve when all available temperature measurements are used, and there are also other reasons for using them all. The sensetivity due to the location becomes less, and the estimator becomes more robust with respect to measurements faults. With PLS it is easy to include some features to detect erronenous temperatures because of the great redundancy in the measurements.

One main difficulty with PLS is to obtain good "calibration" data. Often it is difficult to make precise online steady states measurements, and one must rely on simulations. Even more important is the selection of a optimal "calibration" set, that is, how to ensure that all important directions are excited in the calibration set. Also the way of updating the calibration model is important for the performance. However, all this issues are of a general nature, and a detailed discussion will be beyond the scope of this paper.

Brosilow . The results demonstrates clearly that the Brosilow estimator is not useful to estimate compositions when there is colinearity in the temperatures. (Too many temperatures with nearly the same information). When using the Brosilow method we are simply trying to get more information from the temperatures θ than there actually is. This was pointed out by Brosilow himself, and his recommendation was to reduce the number of measurements [8].

Improved Brosilow Learning from the PLS-regression, one may avoid this problem by including only a limited number of independent directions (certain linear combinations of the temperatures) . One should delete *directions* in θ with little information, but not delete measurements as such. This could most easily be done by computing the pseudo inverse from the singular value (principal components) decomposition (SVD) of F_d and deleting directions corresponding to small singular values in F_d. Avoiding the problem of colinearity, the Brosilow estimator has the advantage of making sure that all the input directions are well excited in the estimator. The drawback is that one have to model all the matrices in the equations 5 and 6.

7 Conclusions

The objective of this paper has been to evaluate some estimator algorithms for estimating the product composition in a distillation column from reflux, boilup, and temperature measurements. The product estimates was supposed to be used as input to two PID controllers.

We have found the μ-analysis very suitable for this task, because we then get a picture of the performance for the different disturbance frequencies without having to perform a large number of simulations.

Although the open-loop test in most cases yield somewhat greater differences between the estimators, the same trends were also found in the closed loop test.

The analysis showed that the Brosilow estimator is impossible to use with many coupled temperatures. The best way to avoid this problem is to delete the directions corresponding to small singular values when using pseudo inverses.

The Kalman-Bucy Filter was found to be the estimator with best performance. The drawbacks are complexity, heavy computer loads, and tuning requirements.

PLS may be a good alternative if a static estimator is sufficient. The method has been useful in giving insight in the estimation problem.

Acknowledgement. The authors wish to thank David Di Ruscio at the Divition of Engineering Cybernetics, NTH for help in calculating the Kalman filter gains, and prof. Terje Hertzberg for many useful comments. Parts of this work has been financed by The Royal Norwegian Council of Scientific and Industrial Research.

NOMENCLATURE

C_θ - measurement matrix from states to temperatures.
d - normalized disturbances
e - normalized "error"
F - feed rate
k - number of factors used in calibration
K - estimator constant
K_f - Kalman filter gain
L - reflux flow rate
M - generalized nominal plant
n - number of calibration set
N - number of theoretical trays
N_F - number of feed tray
p - loading vector (direction of principal component)
q_F - fraction liquid inn feed
t - principal component, score
T - matrix of scores

u - manipulated inputs $(VL)^T$
v - process noise vector
$\mathcal{V}$ - process noise covariance matrix.
V - boilup from reboiler
w - measurement noise vector
w_i - input uncertainty weight
w_p - performance weight
$\mathcal{W}$ - measurement noise covariance matrix.
x - state space vector
x_B - mole fraction of light component in bottom product
$\mathcal{X}$ - state vector covariance matrix.
y - output vector $(y_D x_B)^T$
$\hat{y}$ - estimate of compositions
y_D - mole fraction of light component in distillate
Y - data matrix of y
z_F - mole fraction of light component in feed

Greek symbols

α - relativ volatility
Δ - uncertainty block
μ - Structural Singular Value
ω - frequency (min^{-1})
τ_1 - dominant time constant for external flows
τ_2 - time constant for internal flows
θ - temperature vector
θ_e - extended measurement vector
Θ - data matrix of $theta_e$

References

[1] Nisenfeld, A. E. and R. C. Seeman, "Distillation Columns", *ISA Monograph Series* **2**, 1981.

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

[3] Brosilow, C. B. and Joseph, B., "Inferential Control of Processes",*AIChE Journal*, **24**, 485-509, 1978.

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

[5] Yu, C. and Luyben, W. L., "Control of Multicomponent Distillation Columns using Rigorous Composition Estimators", *I. Chem. Eng. Symp. Ser.*, **No 104**, A29, 1986.

[6] Moore, C., Hackney, J. and Canter, D., "Selecting Sensor Location and Type for Multivariable Processes" *Shell Proc. Contr. Workshop*, 1986.

[7] Keller, J.P, and D. Bonvin, "Selection of inputs for the purpose of model reduction and Controller Design", *Proceedings IFAC Congress, Munich*, 1987.

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

[9] Höskuldsson, A., "PLS Regression Methods", *Jour. of Chemometrics*, **2**, 211-228, 1988.

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

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