

# Estimation of Distillation Compositions from Multiple Temperature Measurements using PLS Regression.

Thor Mejdell

Sigurd Skogestad

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

17-22 November 1991

AIChE 1991 Annual Meeting in Los Angeles  
Paper 148d.

## Abstract

The paper addresses the use of temperature measurements to estimate product compositions in distillation columns. A simple linear multivariate calibration procedure based on steady state data is used, which requires minimal modelling effort. It is found that these PCR- and PLS-estimators perform well, even for multicomponent mixtures, pressure variations, and nonlinearity caused by changes in operating conditions. The use of weighting functions, additional factors and logarithmic transformations improve the estimates and counteract nonlinearities, provided there is not too much noise on the temperatures. In the paper we also compare more generally regression methods based on SVD (generalized least square), PCR and PLS.

## 1 Introduction.

Reliable and accurate measurement of product compositions is one of the main difficulties in distillation column control. Most product analyzers, like gas chromatographs, suffer from large measurement delays and high investments and maintenance costs. The overall measurement delay is typically 10 to 20 minutes. This imposes severe limitations on achievable control performance. However, the reliability of the analyzers is perhaps their weakest point, and this also results in high maintenance costs in terms of manpower and expensive back up systems. One employee per every three GC analyzers is common in industry. The most popular alternative to analyzers is single temperature control, i.e., control of a given tray temperature. Temperature measurements are reliable and inexpensive, and have negligible measurement delays. However, they are not accurate indicators of product composition. Nevertheless, in most cases temperature control is preferred. For example, Kister (1990) recommends using temperature control unless the difference in boiling point between the key components is very small, or there are substantial economical benefits in keeping tight control of the product compositions.

This paper addresses two-product columns where we make a split between two defined key components, denoted the light key (LK) and heavy key (HK) component. We can make two

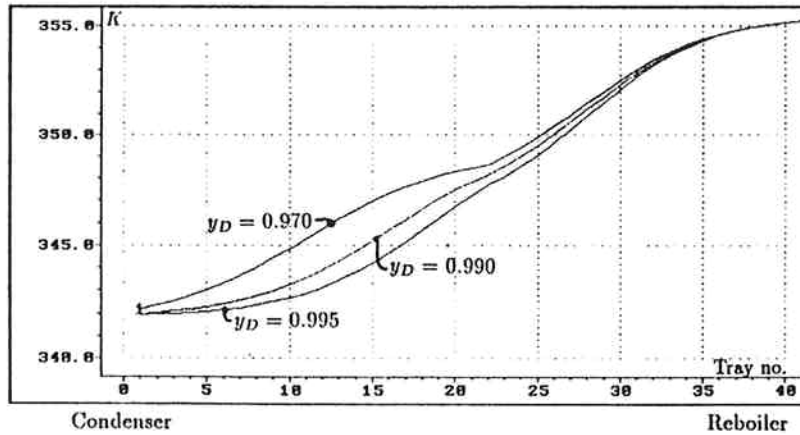


Figure 2: Effect of changes in top composition on the temperature profile. The feed and bottom compositions are constant.

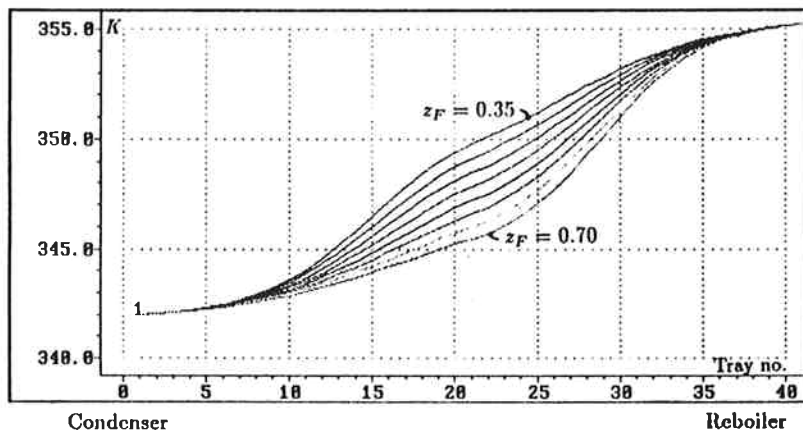


Figure 3: Effect of changes in feed composition on temperature profile. Top and bottom compositions are constant.

largest pressure change near the bottom. Varying liquid holdups and tray performances may give local pressure variations as well.

Some measures may be taken to counteract these problems. The high frequency noise may easily be filtered. Pressure variations may be compensated using pressure measurements, or using differential temperatures. The effect of non-key components may be reduced by locating the temperature measurement in regions of the column where their composition is nearly constant (Rademaker et.al., 1975).

However, some of these problems may not be corrected. In particular, keeping a temperature constant on a tray some distance away will not keep the product composition constant. For example, consider Figure 2 where it is shown that the temperatures in the lower part of the column are affected by changes in top composition, and Figure 3 where it is shown how the temperatures are affected by different feed compositions.

*Multiple temperatures.* One solution to these problems is to use multiple temperature measurements to infer the product composition. There has recently been reports from industry on successful implementations of somewhat *ad hoc* approaches. Whitehead and Parnis (1987) used a

A. Binary mixture:						Antoine parameters			
$i$	comp	$z_F$	$y_D$	$x_B$	$\alpha_{i,i+1}$	$T^b$ (K)	A	B	C
1	LK	0.500	0.99	0.01	1.5	341.9	15.83660	2697.55	-48.78
2	HK	0.500	0.01	0.99		355.4	15.43113	2697.55	-48.78

B. Multicomponent mixture:						Antoine parameters			
$i$	comp	$z_F$	$y_D$	$x_B$	$\alpha_{i,i+1}$	$T^b$ (K)	A	B	C
1	LNK	0.050	0.125	0.000	2.0	321.4	16.52975	2697.55	-48.78
2	LK	0.350	0.866	0.006	1.5	341.9	15.83660	2697.55	-48.78
3	HK	0.350	0.009	0.577	2.0	355.4	15.43113	2697.55	-48.78
4	HNK	0.250	0.000	0.417		381.6	14.73799	2697.55	-48.78

- $\ln P(\text{mmHg}) = A - B/(\theta(K) + C)$
- Feed is saturated liquid.
- Constant molar flows.
- Ideal VLE using Raoult's law
- Constant pressure, 1 atm.
- $N = 40, N_F = 21$  (theoretical trays)
- Holdup on all trays;  $M_i/F = 0.5$  min
- Flow dynamics are neglected.

Table 1: Data for distillation column example.

2. Multicomponent mixture consisting of one heavy and one light non-key in addition to the binary components in case 1.

The column in case 1 is "column A" (Skogestad and Morari, 1988b), which is the same example as used by Mejdell and Skogestad (1990). Data for the column and for the mixtures are given in Table 1. The difference in boiling point between the two pure key components is only 13 °C. This is approximately the lower limit for what is usually recommended when using single temperature control (Riggs, 1990). Typical temperatures profiles for the binary case are shown in Fig. 1 – 4.

## 2.2 The estimation problem.

Consider the case with binary mixture, constant pressure and feed and reflux as saturated liquid. Then specifying different values of feed composition  $z_F$ , distillate composition  $y_D$ , and bottom product  $x_B$ , yields unique steady state profiles of the 41 temperatures  $\theta$ . This may be expressed as

$$\theta^{41 \times 1} = f(z_F, y_D, x_B) \quad (1)$$

We want to find the inverse relation

$$\hat{y}^{2 \times 1} = \begin{pmatrix} \hat{y}_D \\ \hat{x}_B \end{pmatrix} = g(\theta^{41 \times 1}) \quad (2)$$

- Use unknown disturbances as latent variables. This procedure has been proposed by Brosilow and co-workers and employed in their inferential control, e.g. Weber and Brosilow (1972) and Joseph and Brosilow (1978).
- Use some geometric shape factors of the temperature profile as latent variables. This idea is used by Gilles and Retzbach (1980) and later by Marquardt (1989). Here the first factor is the location of the steepest temperature gradient.
- Use the principal components (PCR method) or the partial least square factors (PLS method) as latent variables. These methods have been introduced for distillation columns by Mejdell and Skogestad (1989), and are further outlined in the next section.

The simplest of the above methods is of course to reduce the number of measurements. However, this is not optimal, because additional measurements will

- improve the estimate because they contain more independent information (at least for few measurements)
- reduce the effect of measurement noise
- make the estimator less sensitive to measurement selection and changes in operating conditions
- better capture the effect of nonlinearities.

We have briefly investigated Marquardt's method and found it to perform very well with perfect measurements, but it was sensitive to measurement noise for our column. It might perform better for distillation columns with sharp profiles, for which it was originally developed.

The Brosilow estimator was studied in detail by Mejdell and Skogestad (1990). We found that it performed poorly for the binary distillation example. The reason is that for ill-conditioned plants, like distillation columns, the estimate is very sensitive to small errors. This sensitivity has indeed been pointed out also by Brosilow and co-workers.

Mejdell and Skogestad (1990) studied the static PCR estimator for the example column and found that it performed almost as well as an optimally tuned Kalman filter. In this paper we shall use the static regression estimators, PLS and PCR.

### 2.3 Multicomponent mixture.

The multicomponent mixture is obtained by extending the original binary with one light nonkey (LNK) and one heavy nonkey (HNK) component (See Table 1). The control objective for the separation is still the split with respect to the key components. Let on a given tray  $x'$  denotes the pseudo-binary molefraction (based on key components) of light key component, i.e.

$$x' = \frac{x_L}{x_L + x_H} \quad (6)$$

The product specification for both the binary and multicomponent case is  $y'_D = 0.99$  and  $x'_B = 0.01$ .

The composition profile for the multicomponent mixture at the nominal operating point is displayed in Figure 5. The off-key components are almost constant in the column except at the column ends and around the feed tray. The composition profile on a pseudobinary basis,  $x'$ , is compared with that of the binary example column,  $x$ , in Figure 6, and the corresponding temperature profiles are shown in Figure 7. The concentration profiles are almost identical, whereas the temperature profiles are quite different.

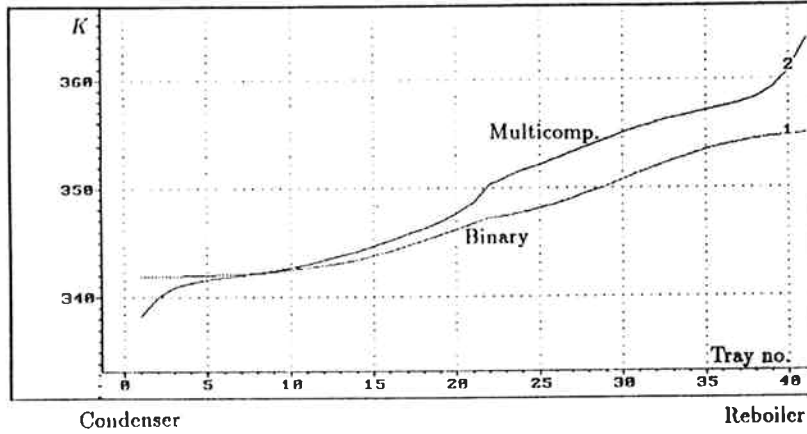


Figure 7: Temperature profiles for the binary and multicomponent cases.

## 2.4 Use of transformed variables.

The composition and temperature profiles are nonlinear functions of the operating variables. One way to deal with nonlinearity is to find other variables (factors) which can capture the nonlinearity. The Marquardt (1989) method is one such method. A simpler method is to use nonlinear transformations on each variable. Logarithmic transformation of the product compositions has been proposed by several authors (eg., Joseph et al., 1976, Skogestad and Morari, 1988a) as an effective way to linearize the dynamic and static *response* (with  $L, V, F, z_F$  etc. as independent variables). For binary mixtures

$$Y_D = \ln(1 - y_D); \quad X_B = \ln x_B \quad (7)$$

These transformations apply also for multicomponent mixtures if pseudobinary compositions are used. The composition *profile* (with tray number as an independent variable) may also be linearized using similar transformations. Ryskamp (1981) plotted the compositions on a probability scale. It is also common to use the logarithm of the separation parameter on each tray (eg., PROCESS, 1981)

$$X = \ln \frac{x_L}{x_H}, \quad (8)$$

Here  $x_L$  and  $x_H$  are mole fractions of light and heavy components on a given tray (the subscript  $i$  is omitted to simplify notation). In Appendix 1 and 2 we show that this transformation in addition to linearizing the *profile* also linearizes the *dynamic response*. Note that since most columns have  $y_D = x_{DL} \approx 1$  and  $(1 - x_B) = x_{BH} \approx 1$ , (7) may be viewed as a special case of (8).

Temperature is often a nearly linear function of composition. Based on Eq. 8 we therefore propose to use the following transformation to linearize the temperature response and profile (Appendix 1)

$$L_T = \ln\left(\frac{\theta - T_L^b}{T_H^b - \theta}\right) \quad (9)$$

$z_F$	$y_D$	$x_B$	$z_F$	$y_D$	$x_B$
0.5000	0.99000	0.01000	0.4000	0.97000	0.03000
0.5375	0.99130	0.02620	0.4000	0.97000	0.00333
0.4250	0.97380	0.01510	0.4000	0.99667	0.03000
0.5250	0.97000	0.01320	0.4000	0.99667	0.00333
0.4125	0.98010	0.00580	0.6000	0.97000	0.03000
0.6000	0.98490	0.00440	0.6000	0.97000	0.00333
0.5125	0.99420	0.00660	0.6000	0.99667	0.03000
0.5500	0.98270	0.00760	0.6000	0.99667	0.00333
0.4875	0.99620	0.01890	0.4500	0.98268	0.01732
0.4750	0.99560	0.00870	0.4500	0.98268	0.00577
0.5625	0.99340	0.01150	0.4500	0.99423	0.01732
0.4625	0.97720	0.03000	0.4500	0.99423	0.00577
0.4375	0.99500	0.00380	0.5500	0.98268	0.01732
0.4500	0.99240	0.01730	0.5500	0.98268	0.00577
0.5750	0.98680	0.02280	0.5500	0.99423	0.01732
0.5875	0.98850	0.00500	0.5500	0.99423	0.00577

Table 2: Specifications used in simulations to obtain static temperature profiles.

listed in Table 2. The data were spread with equal distances around  $z_F$ ,  $x_B$  and  $y_D$ . The first run in the left column is the nominal operating point, the other 15 runs were randomly chosen. The 16 runs in the right column were selected by a two composite design in four levels. The 32 runs correspond to maximum variations in column end temperatures of about 0.4 °C and of interior temperatures of about 5 °C.

Different versions of the 32 calibration set was made, which included

- Total pressure variations of  $\pm 0.1$  atm spread randomly on the various runs.
- Normal distributed random noise of magnitudes 0.1 or 0.2 °C on all temperatures.
- Non-key components. The pseudobinary compositions were as in the binary case (Table 2), but in addition the feed molefraction of light non-key component were varied between 0.025 and 0.075, and of heavy non-key between 0.15 and 0.35. In this case all temperatures were rounded to one decimal, which may be viewed as a (coloured) noise source.

## 2.6 Evaluation criteria.

The main criterion used to evaluate the performance of the estimators is the Mean Square Error of Prediction (MSEP) (Martens and Næs, 1989 p. 250), which is a measure of the expected error of *future* predictions. An estimate of MSEP was performed by a cross validation procedure as follows: The 32 calibration runs were divided into seven groups. Then the calibration procedure (i.e., finding the estimator  $K$ ) was performed seven times, each time with six groups used for calibration and one for testing. The mean square error of all test predictions was then obtained from

$$MSEP(k) = \frac{1}{32} \sum_{i=1}^{32} (\hat{y}_i(k) - y_i)^2 \quad (11)$$

Here  $k$  is the number of factors (or principal components) used in the calibration, and  $y_i$  is  $y_D$  or  $x_B$  for the  $i$ 'th test run. MSEP will generally *not* go to zero when the number of factors

### 3.2 Principal Component Regression. (PCR).

Write the SVD in Eq. (17) on the alternative form

$$\Theta = t_1 p_1^T + t_2 p_2^T + \dots + t_m p_m^T \quad (20)$$

Here  $t_i = u_i \sigma_i$  is the score vector (or latent variable) and  $p_i = v_i$  is the loading vector for principal component  $i$ . Keep only the  $k$  first terms which may be distinguished from measurement noise, and let the matrices  $P^{q \times k}$  and  $T^{n \times k}$  include only these  $k$  most important directions. Then  $\Theta \approx \Theta_k = TP^T$ . The latent variables for a given temperature vector,  $\theta$ , are then given by  $t^{k \times 1} = P^T \theta^{q \times 1}$ . The least square solution to  $Y = TK_t^T$  becomes  $K_t = Y^T T [T^T T]^{-1}$ . The condition number of  $T$ ,  $\gamma(T) = \sigma_1(T)/\sigma_k(T)$ , may be adjusted by selecting the number  $k$ . Since  $P^{-1} = P^T$  ( $P$  is orthonormal), the overall estimator gain matrix then becomes

$$K_{PCR} = Y^T (\Theta_k^\dagger)^T = Y^T T [T^T T]^{-1} P^T \quad (21)$$

### 3.3 Partial Least Square (PLS) Regression.

This is a variation of the PCR method which recently has become popular among analytical chemists. The latent variables are here determined in order to have the greatest covariance with the  $y$ -variables. It is an iterative process, which roughly (see Höskuldsson, 1988) may be described as follows:

For the centered  $\Theta$  and  $Y$  find:

1. the largest eigenvalue  $a_1$  and corresponding eigenvector  $w_1$  of  $\Theta^T Y Y^T \Theta$ . (Scale  $w_1$  to length one).
2. the scores  $t_1 = \Theta w_1$
3. the  $\Theta$ -loadings  $p_1 = \frac{\Theta^T t_1}{t_1^T t_1}$
4. the residual matrix  $E_1 = \Theta - t_1 p_1^T$
5. the  $Y$ -loadings  $q_1 = \frac{Y^T t_1}{t_1^T t_1}$
6. the  $Y$ -residual matrix  $F_1 = Y - t_1 q_1^T$

Then start from the top with the residual matrices  $E_1$  and  $F_1$  in stead of  $\Theta$  and  $Y$ , and continue until the matrix  $E_k^T F_k F_k^T E_k$  has only small eigenvalues left. For a more exact description of the algorithm and its different versions, see Martens and Næs, 1989.

The estimator based on  $k$  factors is

$$K_{PLS} = Q(P^T W)^{-1} W^T \quad (22)$$

where the matrices  $Q^{p \times k}$ ,  $P^{q \times k}$  and  $W^{q \times k}$  are formed by the vectors  $q$ ,  $p$  and  $w$  introduced above.

The main advantage of the PLS-algorithm compared to PCR, is that it selects the directions in  $\Theta$  which have the largest covariance with  $y$ , and thus ensures that these directions are treated first.

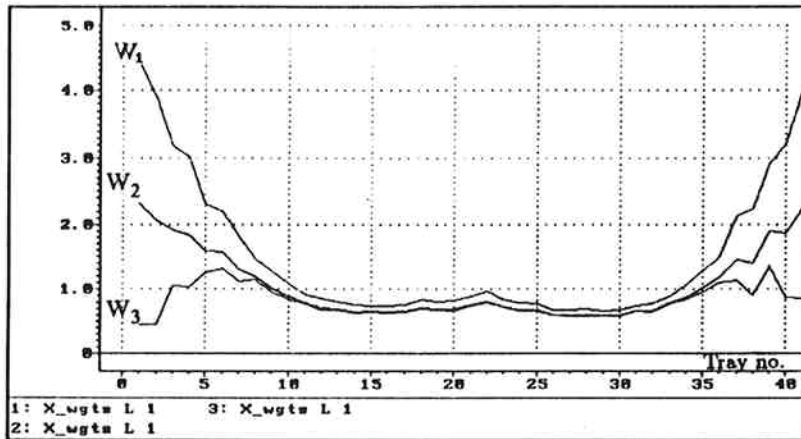


Figure 9: Weight functions  $W_1$ ,  $W_2$  and  $W_3$  for Case  $n_2$ . (Untransformed data with  $0.2^\circ\text{C}$  random noise.)

M	Multicomponent mixture.
P	Pressure variations of $\pm 0.1$ atm.
$L_Y$	Log. transf. comp. only
$L_T$	Log. transf. comp. and temp. (ref.: boiling temp.)
$L_\theta$	Log. transf. comp. and temp. (ref.: tray temp.)
$n_0$	no noise
$n_1$	$0.1^\circ\text{C}$ noise
$n_2$	$0.2^\circ\text{C}$ noise
$W_n$	Weight function n

Table 3: Estimator cases



No noise:							
	No. of factors						
	1	2	3	4	5	6	7
$n_0$	19.61	81.97	94.17	97.18	98.18	99.43	99.96
$n_0L_Y$	21.78	80.14	92.86	93.45	96.92	97.20	98.80
$n_0L_T$	19.72	96.38	<b>99.97</b>	99.98	99.99	100.00	100.00
$n_0L_\theta$	32.90	95.39	<b>99.95</b>	99.96	99.97	100.00	100.00
$n_0P$	-2.31	46.77	75.71	92.19	95.50	95.95	97.59
$n_0L_\theta P$	19.16	94.90	<b>99.94</b>	99.97	99.97	100.00	100.00
$Mn_0$	10.66	48.02	56.78	88.35	89.58	91.50	91.94
$Mn_0W_2$	11.14	48.44	57.40	87.76	90.61	91.84	92.25
$Mn_0L_Y$	18.52	48.88	52.59	83.66	84.10	87.84	87.43
$Mn_0L_\theta$	22.73	77.52	94.10	95.81	96.27	97.61	97.55
$Mn_0L_\theta W_2$	20.69	73.75	93.81	95.87	96.75	97.68	97.66
0.1°C noise							
	1	2	3	4	5	6	7
$n_1$	19.59	81.78	93.81	97.00	97.59	97.86	98.05
$n_1L_TW_3$	23.10	91.42	<b>98.89</b>	98.92	98.92	98.82	98.80
$n_1L_\theta W_3$	23.18	84.60	91.09	94.01	96.16	96.69	97.06
$n_1L_\theta W_3P$	26.89	91.05	98.25	98.78	99.04	99.08	99.04
0.2°C noise							
	1	2	3	4	5	6	7
$n_2$	18.87	81.10	93.78	94.92	94.49	94.42	94.47
$n_2W_1$	17.06	86.04	95.75	94.66	94.39	94.18	93.64
$n_2W_2$	20.40	84.65	95.66	95.36	94.74	94.70	94.12
$n_2W_3$	19.78	83.34	95.08	95.59	95.40	95.25	95.04
$n_2L_YW_3$	18.14	80.14	91.53	92.57	91.73	91.29	90.53
$n_2L_T$	12.50	78.41	81.90	84.04	87.59	90.63	91.16
$n_2L_TW_1$	22.65	86.65	95.44	96.25	96.70	97.22	97.33
$n_2L_TW_2$	22.78	87.66	96.43	96.99	97.30	97.64	97.59
$n_2L_TW_3$	23.04	89.07	<b>97.49</b>	97.68	97.86	98.14	98.05
$n_2L_\theta W_3$	26.41	84.53	86.59	87.43	90.07	89.10	88.52
$n_2L_\theta W_3P$	20.58	86.30	<b>98.01</b>	98.09	97.89	97.84	97.80

Table 4: EPV for different PLS estimators.

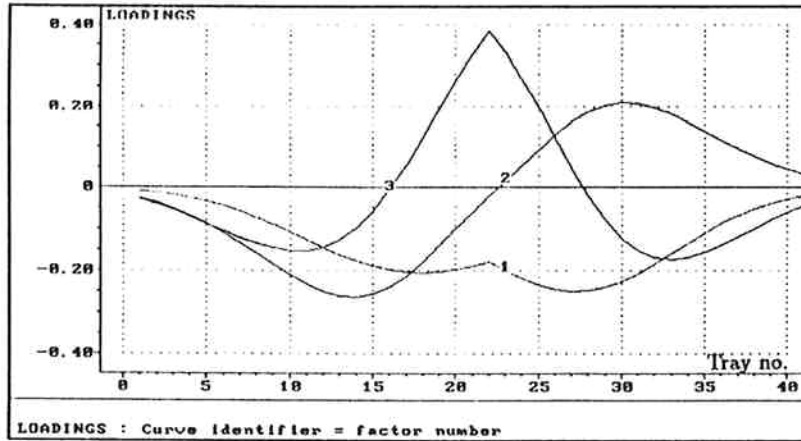


Figure 12: Loading plot (vector  $p_i$ ) of first three factors for estimator  $n_0$ . Curve identifier:  $i$ , factor number.

#### 4.5 Effect of pressure.

Figure 13 displays the loading plot for the case with total column pressure variations of 10% (case  $n_0P$  in Table 4). We see that the first factor, which has no predictive ability, mainly represents the pressure variation. To get good predictive properties we need at least 5 factors. The pressure variation may alternatively be taken care of by using differential temperatures.

#### 4.6 Use of logarithmic transformations.

The results in Table 4 seem to indicate that use of logarithmic transformed compositions, i.e.  $Y_D = \ln(1 - y_D)$  and  $X_B = \ln x_B$ , combined with untransformed temperatures ( $L_Y$ -estimators), generally has a negative effect on the estimate. However, the results are not quite comparable because the EPV is based on  $Y_D$  and  $X_B$  in stead of  $y_D$  and  $x_B$ . Other tests show that they are comparable in performance, but they have slightly different properties: The accuracy of the logarithmic estimator will be best in the pure region, and will never give estimates outside the region 0 – 1. Furthermore, for feedback control it may be an advantage to use  $Y_D$  and  $X_B$  because this makes the controller tunings less dependent on operating conditions (eg., Skogestad and Morari, 1988b).

However, estimator performance is significantly improved by using logarithmic transformations also on the temperatures. We see from Table 4, cases  $n_0L_\theta$  and  $n_0L_T$ , that with no noise EPV is close to 100% after only 3 factors. This is the case also when pressure variations are included, that is, these are automatically taken care of when transformed temperatures  $L_\theta$  are used.

#### 4.7 Effect of weights on temperatures.

We shall consider the case with logarithmic temperatures. In cases  $n_0L_T$  and  $n_0L_\theta$  with no noise there is no improvement of weighting the measurements, because the logarithmic transform will automatically weigh the temperatures similar to weight  $W_1$ . But when noise is added to the

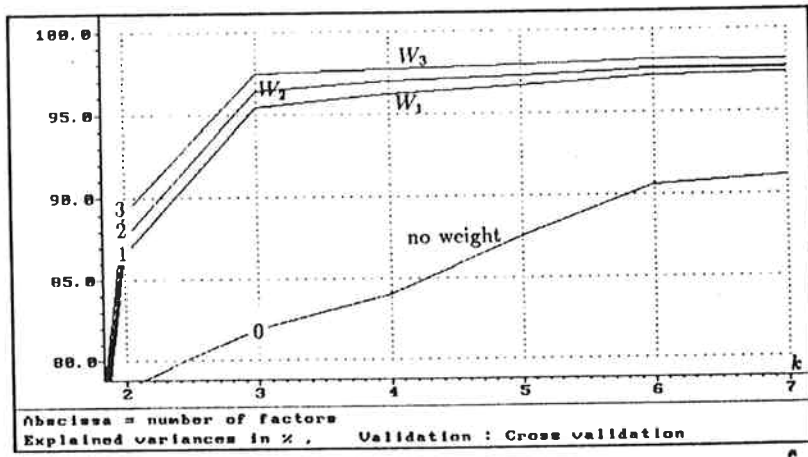


Figure 14: Effect of weights on EPV for estimator  $n_2L_TW_i$  with logarithmic temperatures and  $0.2^\circ\text{C}$  noise.

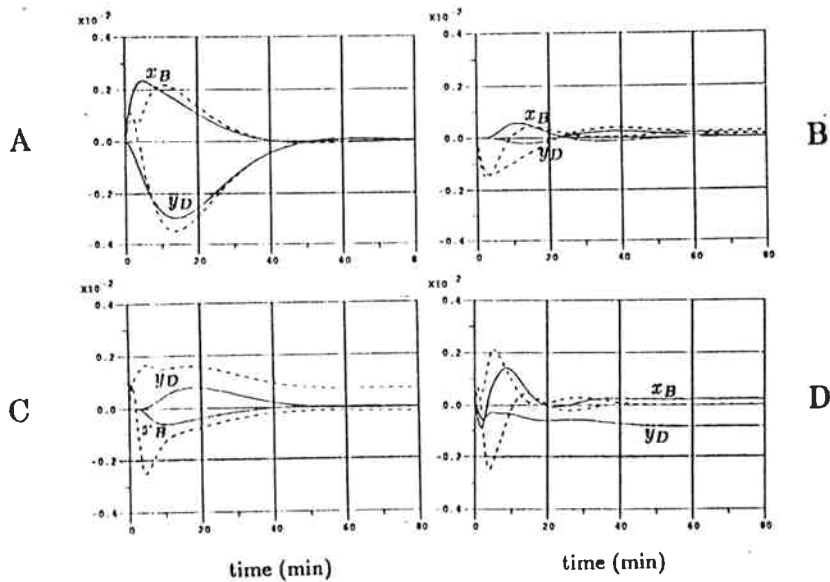


Figure 15: Closed loop responses of compositions  $y$  (solid line), and estimates  $\hat{y}$  (dotted line), for multicomponent case. Estimator is  $Mn_0L_\theta W_2$  with 4 factors. Results are shown for 20 % steps in A) feedrate, B) feed composition  $z'_f$ , and C, D) heavy non-key composition.  $y$  is used for feedback in cases A–C, whereas  $\hat{y}$  is used in case D.

Using a larger number of factors than the number of degrees of freedom is helpful when the noise level is not too high. The reason why this may help is that the product compositions have different nonlinear relations to different temperatures. These differences will appear as extra directions (factors) in the linear temperature space. However some of these directions may be too small to be distinguished from noise. A useful rule is to increase the number of factors until they no longer have any significant positive effect on the EPV. For our distillation column, the typical optimal number of factors is 3–5.

Weighting of variables is commonly used for dealing with different kind of measurements. This is to prevent that the measurements with the largest nominal changes dominate. In our distillation column only temperature measurements are used, so one might think that weighting is unnecessary. Nevertheless, weighting proved useful. The reason is that the temperature changes are very small at the column ends (in the calibration set their standard deviation is only 6% of the temperature with the largest variation), but even in the presence of noise they do contain useful information about the end compositions. By using weights we avoid that the PLS method discards the use of these temperatures. However, the noise should also be taken into account, and weight functions  $W_2$  and even more  $W_3$ , which include information about the noise, yield better results.

Compared to PCR, there is a kind of weighting inherent in the PLS-method (Höskuldsson, 1988). This follows since it searches for directions in the  $\Theta^T Y Y^T \Theta$ , instead of only  $\Theta^T \Theta$ . The weighting with the matrix  $Y Y^T$  gives temperatures which have the largest covariance with  $y$  larger weight when making the factors. The comparison between PLS and PCR shows that these “weightings” improved the estimates in some cases, although the difference was quite small.

The use of logarithmic transformations of temperatures was clearly the single method with the greatest effect. It appears to be a very powerful method to cope with the nonlinearity in distillation columns, and also automatically gives the temperature measurement at the ends a greater weight. However, when the temperatures are corrupted with noise, the noise will also be transformed, and have a relatively large effect on these end temperatures. Therefore, as seen from Figure 14, it is absolutely necessary to weight the transformed data to take this into account. As already mentioned, one should also take action to avoid very small or negative temperature differences before transforming the temperatures.

*Choice of reference temperature.* For multicomponent mixtures the reference temperatures should be located some distance away from the ends: 1) They should be in the section where the concentration of the off key component is almost constant (See Fig. 5). Much of the off-key component’s contribution to the temperature will then be cancelled. 2) They should be located as far out to the ends as possible to capture the nonlinearity. A method which combines these two criteria is to use the temperature with least variation in the calibration set.

*Measurement selection.* The results above are based on using *all* temperatures as measurements. This is of course not necessary. However, the number of measurement should at least be equal to the number of factors needed for prediction. For example, to capture 3 factors, we need at least 3-5 temperatures. The highest number applies to estimators which use differential temperatures, for example,  $L_\theta$ . Additional temperatures will mainly reduce the effect of measurement noise. As a simple method to select the location of temperature measurements, we recommend identifying the peak elements in the  $K$ -matrix for the weighted (scaled) variables. The number of peaks is usually the same as the number of factors. As an illustration, consider Figure 16 which displays the elements in  $K$  for the weighted untransformed temperatures for  $y_D$  using PLS-estimators  $n_0 W_2$ ,  $n_1 W_2$ ,  $n_2 W_2$  and  $n_0$  with 3 factors (the weight  $W_2$  is different in each case as it depends on the noise level). We see that when the noise is increased it seems better to locate the temperatures further from the end.

The results for multicomponent mixtures indicate that the estimator may perform well in a wide range of applications. Using section reference temperatures at locations with the smallest temperature variance, will make logarithmic transforms useful also here.

Besides being an efficient method of obtaining estimators, the standard multivariable calibration techniques yield added benefits, such as insight in the process, good statistical information about the prediction ability, and a method for sensor location.

### Acknowledgements.

Financial support from the Royal Norwegian Council of Scientific and Industrial Research (NTNF) is gratefully acknowledged. Most of the calculations have been performed by using the Unscrambler<sup>TM</sup> software package provided by H. Martens.

### NOMENCLATURE (also see Table 4)

$B$  - Bottom product flow rate

$D$  - Distillate flow rate

$E_k$  - Residual matrix for  $\Theta$  using  $k$  factors.

EPV - Explained Prediction Variance in %. See Eq. 12

$F$  - feed rate

$k$  - number of factors used in estimator

$K$  - estimator constant

$L$  - reflux flow rate

$L_T$  - logarithmic temperature based on boiling points

$L_\theta$  - logarithmic temperature based on reference temperatures

$t$  - vector of latent variables

$T$  - matrix of latent variables ( $t$ ) for calibration runs

$T^b$  - boiling temperature of pure component

$V$  - boilup from reboiler

$x_B$  - mole fraction of light component in bottom product

$y_D$  - mole fraction of light component in distillate

$y$  - output vector ( $y_D, x_B$ )<sup>T</sup>

$Y$  - matrix of outputs ( $y$ ) for calibration runs

$z_F$  - mole fraction of light component in feed

$W_1, W_2, W_3$  - weight functions for measurement (temperature) scaling

#### Greek symbols

$\alpha_{i,j}$  - relative volatility between components  $i$  and  $j$

$\gamma$  - condition number

$\sigma_i$  -  $i$ 'th singular value

$\theta$  - tray temperature or temperature vector

$\theta_L, \theta_H$  - reference temperature in top and bottom of column

$\Theta$  - matrix of temperatures ( $\theta$ ) for calibration runs

#### Subscripts

$H$  - heavy key component

$i$  - tray number

$L$  - light key component

#### Superscripts

' - pseudobinary basis

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

Whitehead, D. B. and M. Parnis, 1987 "Computer control improves ethylene plant operation." *Hydrocarbon Processing*, (11), 105–108.

## Appendix 1:

### Linearizing effect of logarithmic transform on profile.

1) Consider a binary mixture and assume constant relative volatility. Then

$$\frac{y_i(1-x_i)}{x_i(1-y_i)} = \alpha \quad (26)$$

At infinite (total) reflux we have  $y_i = x_{i+1}$  and we get

$$\ln \frac{x_{i+1}}{1-x_{i+1}} - \ln \frac{x_i}{1-x_i} = \ln \alpha \quad (27)$$

Introduce the logarithmic mole fraction

$$X_i = \ln \frac{x_i}{1-x_i} \quad (28)$$

Then

$$X_{i+1} - X_i = \ln \alpha \quad (29)$$

That is, the profile in terms of  $X_i$  is linear.

2) At finite reflux we derive a similar expression: In the bottom part the material balance is

$$L_B x_{i+1} = V y_i + B x_B \quad (30)$$

where  $L_B = L + F$  is the liquid flow in the bottom part. Eq. 26 then gives:

$$\ln \frac{x_{i+1} - \frac{B}{L_B} x_B}{1 - \frac{L_B}{V} x_{i+1} + \frac{B}{V} x_B} - \ln \frac{x_i}{1-x_i} = \ln \left( \alpha \frac{V}{L_B} \right) \quad (31)$$

In most cases the first term is approximately equal to  $\ln \frac{x_{i+1}}{1-x_{i+1}}$ . (The numerator approximation may not apply to the bottom tray if  $\frac{B}{L_B}$  is close to 1, the denominator approximation does usually not apply close to the feed tray where  $x_{i+1}$  is not small). We then get

$$X_{i+1} - X_i \approx \ln \left( \alpha \frac{V}{L_B} \right) \quad (32)$$

(Here  $\alpha \frac{V}{L_B}$  is the ratio between the slope of the equilibrium line and operating line in the bottom section.) A similar expression but with  $\alpha \frac{V}{L_B}$  replaced by  $\alpha \frac{L}{V}$  is derived for the top. Around the feed tray the expression does not apply, and the profile in terms of  $X_i$  may not be linear.

3) To derive a linearizing transform in terms of temperatures we assume that the boiling temperature is a linear function of  $x_i$ , that is, for the binary case

$$\theta_i = T_L^b x_i + T_H^b (1-x_i) \quad (33)$$