Proceedings of the American Control Conference
Arlington, VA June 25-27, 2001

Inferential Feedback Control of Distillation Composition based on PCR and PLS Models

Jie Zhang
Centre for Process Analytics and Control Technology
Department of Chemical & Process Engineering
University of Newcastle, Newcastle upon Tyne NE1 7RU, U. K.
Tel. +44-191-2227240, Fax: +44-191-2225292
E-mail: jie.zhang@newcastle.ac.uk

Abstract
A principal component regression (PCR) and partial least squares (PLS) model based inferential feedback control strategy for distillation composition control is developed. PCR and PLS model based software sensors are developed from process operational data so that the top and bottom product compositions can be estimated from multiple tray temperature measurements. The PCR and PLS software sensors are used in the feedback control of the top and bottom product compositions. This strategy can overcome the problem of substantial time delay in composition analysers based control and the problem of substantial bias in single tray temperature control. Static estimation bias and the resulting static control offsets are eliminated through mean updating of process measurements. Applications to a simulated methanol-water separation column demonstrate the effectiveness of this control strategy.

Keywords: Software sensors, inferential control, principal component regression, partial least squares, distillation column control.

1. Introduction

In the control of distillation columns, it is usually difficulty to get accurate and reliable product composition measurements without time delay. Many composition analysers such as gas chrotomography usually possess significant time lags. The overall time lags in composition measurements are typically between 10 to 20 minutes (Mejdell and Skogestad, 1991). Such long time lags significantly reduce the achievable performance of composition controllers. A further drawback of composition analysers is that their reliability is usually quite low. Using composition analysers in distillation composition control will therefore incur high maintenance cost. Therefore, in distillation composition control, it is a usual practice to use tray temperatures to represent product compositions. In a binary distillation column, the temperature of a tray at the top of the column is usually used to represent the top product composition while the temperature of a tray at the bottom part of the column is usually used to represent the bottom product composition. Compared with composition measurements, temperature measurements are more reliable and economic and virtually without any measurement time lags. As pointed out by Kister (1990), tray temperatures are usually used in distillation composition control unless the differences between the boiling points are small or tight control of composition can bring in significant economic benefit.

However, using a single tray temperature to represent product composition has the following drawbacks (Mejdell and Skogestad, 1991): 1). even for binary mixtures the relationship between tray temperature and product composition depends on the feed composition and the product composition at the other end of the column; 2). for multicomponent mixtures the presence of off-key components implies that even at the column ends temperature is not an exact indicator of composition; 3). column pressure variations can affect tray temperatures; 4). feed rate jump can also affect tray temperatures. To overcome these problems, multiple tray temperatures should be utilised. Due to the strong correlation among tray temperature measurements, multiple linear regression is usually inappropriate and the PCR or PLS methods should be utilised (Kaspar and Ray, 1992; Kresta *et al.*, 1991). Mejdell and Skogestad (1991) report the estimation of distillation compositions from multiple temperature measurements using the PLS regression technique.

This paper presents a PCR and PLS model based inferential feedback control strategy. A PCR or PLS model is developed from process operational data so that the top and bottom product compositions can be estimated from multiple tray temperature measurements. The estimated product compositions are directly used in a feedback control loop. A technique for eliminating estimation bias and the associated static control offsets through process measurement mean updating is proposed in this paper.

The paper is structured as follows. Section 2 presents PCR and PLS model based software sensors for product compositions. Inferential feedback control based on the PCR and PLS models is detailed in Section 3. Eliminating static estimation bias and the associated static control offsets through mean updating is given in Section 4. The last section contains some concluding remarks.

0-7803-6495-3/01/$10.00

2. PCR and PLS model based software sensors

The distillation column studied in this paper is a comprehensive nonlinear simulation of a methanol-water separation column. A nonlinear tray by tray dynamic model has been developed using mass and energy balances. This simulation has been validated against pilot plant tests and is well known for its use in control system performance studies (Tham *et al.*, 1991a; 1991b). The following assumptions are imposed: negligible vapour holdup, perfect mixing in each stage and constant liquid holdup. The nominal operation data for this column are listed in Table 1.

Table 1. Nominal distillation column operation data

No. of theoretical stages	10
Feed tray	5
Feed composition (z)	50% methanol
Feed flow rate (F)	18.23 g/s
Top composition (y_D)	95% methanol
Bottom composition (y_B)	5% methanol
Top product rate (D)	9.13 g/s
Bottom product rate (B)	9.1 g/s
Reflux rate (L)	10.0 g/s
Steam rate (V)	13.8 g/s

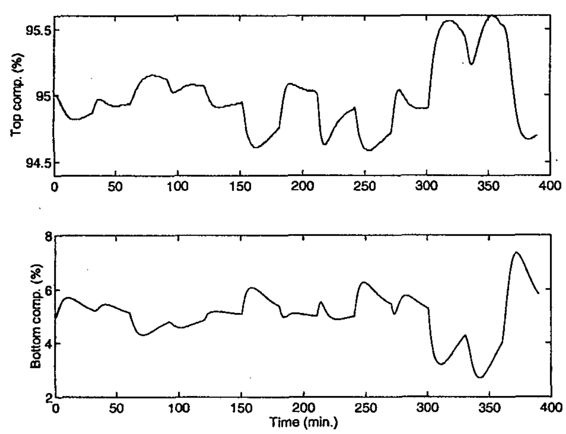

Figure 1. Top and bottom product compositions

In this study the nominal operating point considered is the top composition at 95% and the bottom composition at 5%. To generate data for building PCR and PLS inferential estimation models, random perturbations of ±15% were added to the feed rate and the feed composition. Measurement noises of the distribution N(0°C, 0.1°C) were added to the tray temperature measurements. Figure 1 shows the top and bottom product compositions while Figure 2 shows the tray temperatures. The sampling time used is 1 minute.

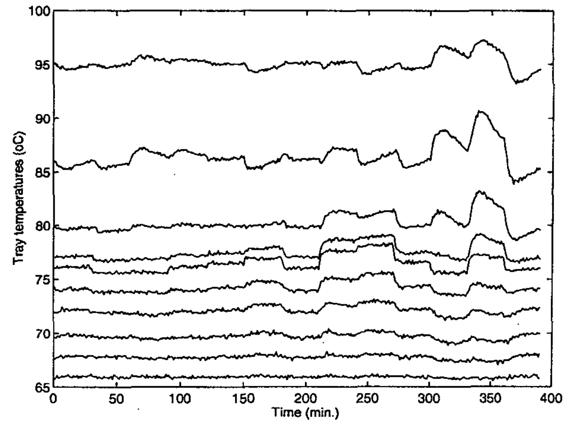

Figure 2. Tray temperatures

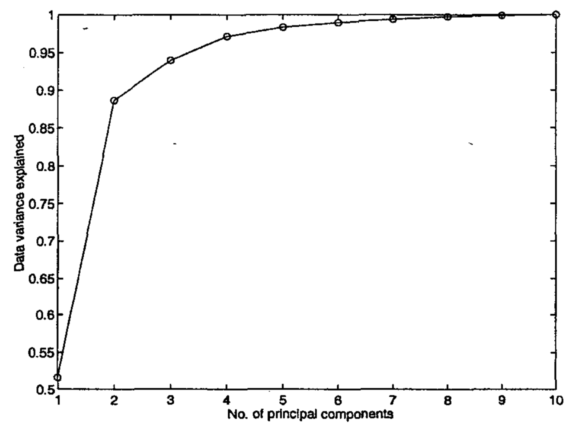

Figure 3. Accumulate data variance explained by principal components

It can be seen from Figure 2 that strong correlation exists among the tray temperatures. Principal component analysis of the tray temperature measurements shows that the first three principal components can explain 93.6% of the data variation. Figure 3 gives the accumulated data variance explanation of the principal components. Due to the strong correlation among tray temperatures, it is not appropriate to build a model between tray temperatures and product compositions using multiple linear regression. Here we use PCR to build the model. The last 240 data points in Figures 1 and 2 were used as training data while the first 150 data points were used as testing data. The appropriate number of principal components retained in the PCR model was determined based on the PCR model errors on the testing data. Figure 4 shows the mean squared errors (MSE) of different PCR models on the testing data. It can be seen that the model with 7 principal components has the smallest MSE on the testing data. Hence, the number of principal components were determined as 7. The identified PCR model is:

$$y_D = 95 + 0.0252\Delta T_1 - 0.0051\Delta T_2 + 0.0036\Delta T_3 + 0.0456\Delta T_4 + 0.1142\Delta T_5 - 0.0790\Delta T_6 - 0.3964\Delta T_7 - 0.3279\Delta T_8 - 0.2375\Delta T_9 - 0.0965\Delta T_{10} \qquad (1)$$

$$y_B = 5 - 0.9916\Delta T_1 - 0.1666\Delta T_2 + 0.1330\Delta T_3 + 0.0968\Delta T_4 - 0.1829\Delta T_5 - 0.0530\Delta T_6 + 0.1271\Delta T_7 + 0.1878\Delta T_8 + 0.0091\Delta T_9 + 0.0483\Delta T_{10} \qquad (2)$$

where y_D and y_B are the top and bottom compositions (%) respectively, ΔT_1 to ΔT_{10} are the deviations of tray temperatures from their nominal mean values. Figure 5 gives the predictions from this PCR model. In Figure 5, the solid lines represent the true compositions from simulation whereas the dashed lines represent PCR model predictions. It can be seen the PCR model predictions are quite accurate.

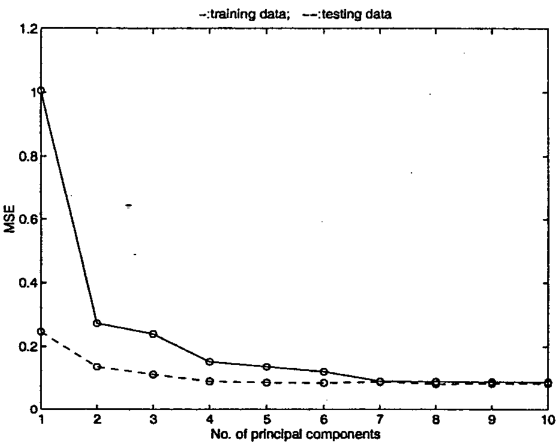

Figure 4. Mean squared errors of different PCR models

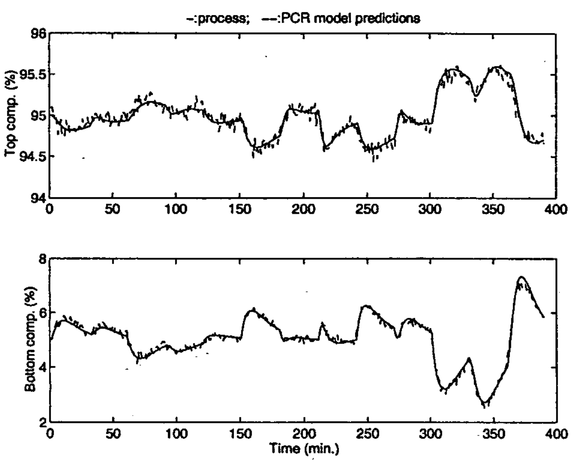

Figure 5. Predictions from the PCR model

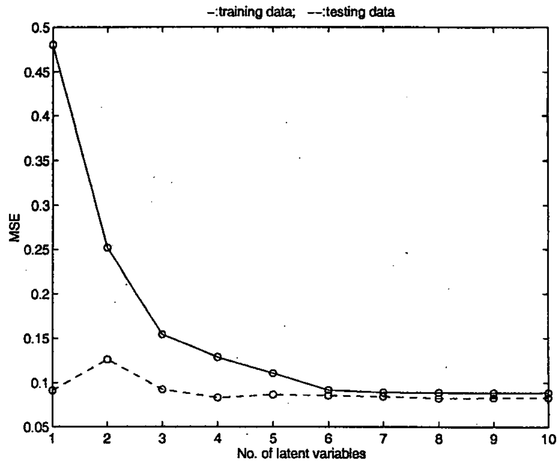

Figure 6. Mean squared errors of different PLS models

A PLS model was also developed for this process. Once again the last 240 data points in Figures 1 and 2 were used as training data while the first 150 data points served as testing data. The appropriate number of latent variables retained in the PLS model was determined based on the model prediction errors on the testing data. Figure 6 gives the MSE of different PLS models on the testing data. It can be seen from Figure 6 that the model with 4 latent variables has the smallest MSE on the testing data. Therefore we selected 4 latent variables in the PLS model. The identified PLS model is:

$$y_D = 95 + 0.0015\Delta T_1 + 0.0045\Delta T_2 + 0.0181\Delta T_3 + 0.0269\Delta T_4 + 0.0387\Delta T_5 - 0.0350\Delta T_6 - 0.2246\Delta T_7 - 0.3419\Delta T_8 - 0.4682\Delta T_9 - 0.0545\Delta T_{10} \qquad (3)$$

$$y_B = 5 - 0.4903\Delta T_1 - 0.2535\Delta T_2 - 0.2208\Delta T_3 - 0.0052\Delta T_4 + 0.1115\Delta T_5 + 0.1467\Delta T_6 + 0.1134\Delta T_7 + 0.1530\Delta T_8 + 0.0864\Delta T_9 - 0.0912\Delta T_{10} \qquad (4)$$

Figure 7 shows the predictions from the PLS model. In Figure 7, the solid lines represent the true simulated product compositions while the dashed lines represent the PLS model predictions. It can be seen that the PLS model predictions are quite accurate.

3. Inferential feedback control of distillation composition

The PCR and PLS software sensors developed in the previous section were used in the feedback control of distillation compositions. The software sensor based feedback control structure is shown in Figure 8. In this control structure, the manipulated variables for composition control are reflux rate (L) and the steam rate to the re-boiler (V). Tray temperature measurements are fed to the PCR (or PLS) based soft-sensor. The predicted product compositions are compared with their setpoints and the errors are fed to a feedback controller. The feedback controller can be of any form such as a multi-loop controller or a multivariable controller. In this study, a multi-loop PI controller was used.

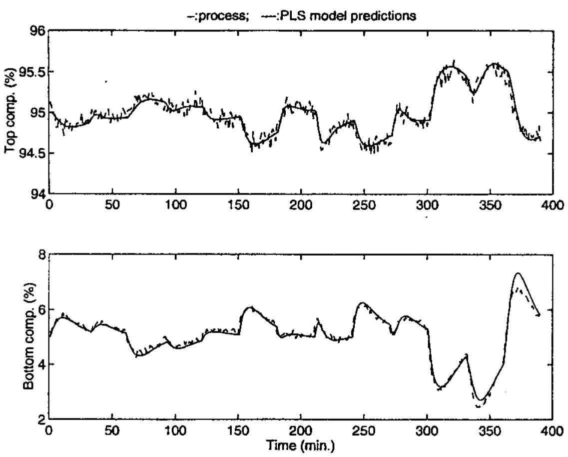

Figure 7. Predictions from the PLS model

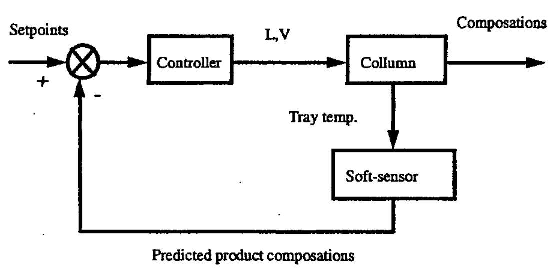

Figure 8. Inferential feedback control structure

For the purpose of comparison, a tray temperature based distillation composition controller and a composition analyser based composition controller were also developed. In the tray temperature based composition control, a single tray temperature was used to represent the product composition. Through analysing the data shown in Figures 1 and 2, it was found that temperature of the 8^{th} tray (from the column bottom) has the largest correlation coefficient (-0.91) with the top product composition while temperature of the 2^{nd} tray has the largest correlation coefficient (-0.93) with the bottom product composition. Therefore, temperatures of the 2^{nd} and the 8^{th} trays were controlled to indirectly control top and bottom product compositions respectively. At the nominal operating point (top composition at 95% and bottom composition at 5%), temperatures at the 2^{nd} and the 8^{th} trays are 86.6°C and 70°C respectively. Therefore, the setpoints for tray 2 and 8 temperatures were set as 86.6°C and 70°C respectively. In the composition analyser based composition control, a 5 min measurement lag was assumed. For all the cases, multi-loop PI controllers were used and tuned using the BLT tuning method (Luyben, 1986).

Figure 9 shows the responses of the composition controllers under feed rate and feed composition disturbances. In Figure 9, the feed rate was increased by 15% at the 51^{st} minutes, the feed composition was increased by 15% at the 251^{st} minutes, the feed rate was decreased by 15% at the 451^{st} minutes, and finally the feed composition was decreased by 15% at the 651^{st} minutes. In Figure 9, the solid, dashed, dash-dotted, and dotted lines represent the responses of the composition analyser based control, the PCR software sensor based control, the PLS software sensor based control, and the tray temperature control respectively.

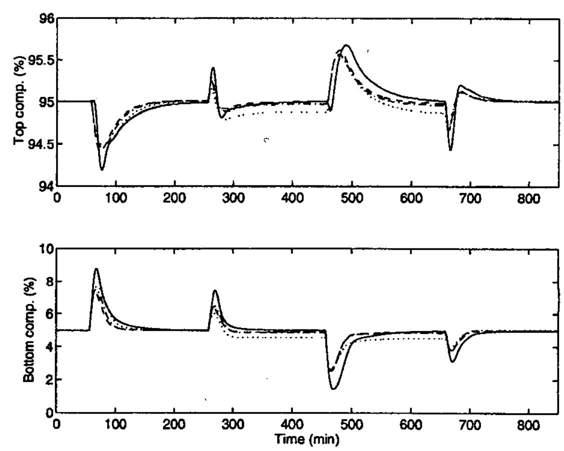

Figure 9. Responses of the composition controllers

Due to the large measurement delay in the composition analyser based control, the controller has to be substantially de-tuned to ensure stability. This resulted in sluggish responses. In the tray temperature control scheme, substantial bias can be observed, especially after the introduction of the first two disturbances. This is due to the fact that the relationship between a single tray temperature and a composition can be significantly affected by process operating condition variations. For the PCR and PLS software sensor based control, much improved control performance are achieved. Some slight static control offsets can be observed especially after the introduction of the first two disturbances. These static control offsets are due to estimation bias caused by the variations in process operating conditions. Table 2 gives the sum of squared errors (SSE) of different control schemes. It can be seen that the PCR and PLS software sensor based inferential feedback control schemes perform much better than the composition analyser based control and the tray temperature control.

Table 2. SSE of different control schemes

	PCR	PLS	Comp. analyser	Tray temp.
Top Comp.	22.1815	20.9823	35.8802	25.1967
Bottom Comp.	230.4493	264.1654	661.7613	373.9638

4. Eliminating estimation/control offsets through mean updating

The PCR and PLS models were developed from process operational data around the nominal operating point: the top and bottom compositions at 95% and 5% respectively. When the column operating condition changes, the PCR and PLS model can give estimation bias. Since the distillation column exhibits some degrees of nonlinearity, a linear PCR or PLS model will inevitably posses some estimation bias, especially when the operating condition changes. Bias in the PCR or PLS model estimation can lead to static control offsets. This can be observed from Figure 10. In Figure 10, the setpoints for the top and bottom compositions were changed to 96% and 4% respectively at the 51st minutes and a 15% increase in feed rate was introduced at the 251st minutes, followed by a 15% increase in feed composition at the 451st minutes. In Figure 10, the solid lines represent the setpoints, the dashed lines represent the product compositions, and the dotted lines represent the PCR model estimations. It can be seen from Figure 10 that, due to operating condition changes, the PCR model has estimation bias leading to static control offsets.

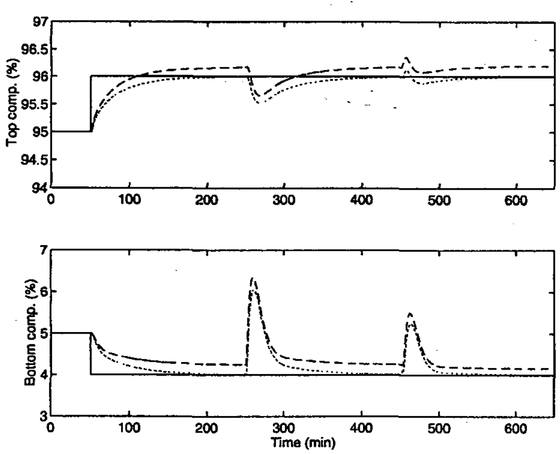

Figure 10. Responses of the PCR model based inferential feedback controllers without mean updating

In this study, we use mean updating to eliminate the static model estimation bias and the associated static control offsets. In the development of a PCR or a PLS model, it is a usual practice to scale the model input and output variables to zero mean and unit variance. These process variable means usually reflect a particular process operating condition. When the process operating condition changes due to, for example, setpoint changes or a disturbance entering the process, the mapping between the scaled model input and output variables (based on the means of the training data) may also change. Here we propose to scale the process variables using their on-line updated means. When the process operating condition changes, a steady state detection method (e.g. Cao and Rhinehart, 1995; Abu-el-zeet *et al.*, 2000) is used to detect if a new steady state has been reached or not. Once it is detected that a new steady state is reached, the process variable means are replaced by their means at that new steady state. The PCR or PLS model input output variables are then scaled using their on-line updated means. To use this mean updating approach, delayed composition measurements from composition analysers are required.

Figure 11 shows the estimation and control performance of the PCR model based inferential feedback controller by using mean updating. In Figure 11, the solid lines represent the setpoints, the dashed lines represent the product compositions, and the dotted lines represent the PCR model estimations. Through mean updating, the PCR model estimation bias and the resulting static control offsets have been eliminated. Similar improvement has also been observed in the PLS model based inferential feedback control. Table 3 summarises the sum of squared errors of the control schemes with and without mean updating. It can be seen that control errors have been reduced through mean updating. The reduction in control errors is mainly due to the elimination of static control offsets as seen from Figure 11.

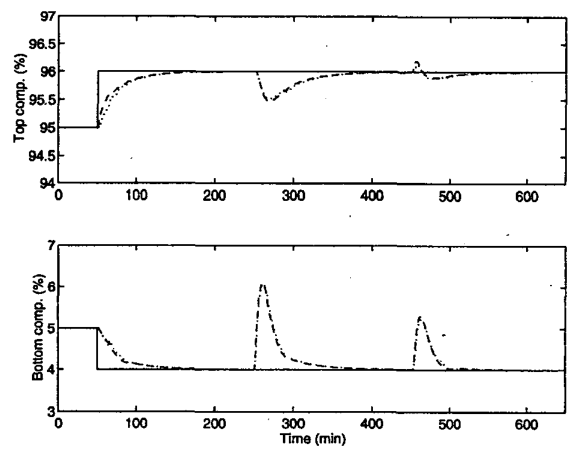

Figure 11. Responses of the PCR model based inferential feedback controllers with mean updating

Table 3. SSE of the inferential feedback control schemes with and without mean updating

Control schemes		Top comp.	Bottom comp.
PCR	without mean updating	27.17	182.64
	with mean updating	22.67	110.65
PLS	without mean updating	26.96	235.74
	with mean updating	23.16	111.72

5. Conclusions

Studies in this paper show that the PCR or PLS software sensor based composition control is superior to single tray temperature control and composition analyser based feedback control where substantial measurement delay exists. By using the PCR or PLS software sensor, substantial measurement delays can be eliminated and, hence, the close loop control performance is improved. By utilising multiple tray temperatures, enhanced correlation between tray temperatures and the top and bottom compositions can be achieved. Colinearity in the multiple tray temperature measurements can be effectively handled by the PCR or PLS method. Through mean updating, static offsets in the PCR or PLS software sensor estimation and the resulting control can be eliminated. Applications to a simulated methanol-water separation column demonstrate the effectiveness of the control strategy.

References

Abu-el-zeet, Z. H., V. M. Becerra, and P. D. Roberts, "Data reconciliation and steady state detection applied to a chemical process", In *Proceedings of Control 2000*, University of Cambridge, England, 4-7 September 2000.

Cao, S. and R. R. Rhinehart, "An efficient method for on-line identification of steady state", *Journal of Process Control*, Vol.6, 1995, pp363-374.

Kaspar, M. H. and W. H. Ray, "Chemometric methods for process monitoring and high-performance controller-design", *AIChE Journal*, Vol.38, No.10, 1992, pp1593-1608.

Kister, H. Z., *Distillation Operation*, McGraw-Hill, New York, 1990;

Kresta, J. V., J. F. MacGregor, and T. E. Marlin, "Multivariate statistical monitoring of process performance", *Canadian Journal of Chemical Engineering*, Vol.69, 1991, pp35-47.

Luyben, W. L., "Simple method for tuning SISO controllers in multivariable systems", *Ind. Eng. Chem. Process Des. Dev.*, Vol.25, 1986, pp654-660.

Mejdell, T. and S. Skogestad, "Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression", *Ind. Eng. Chem. Res.*, Vol.30, 1991, pp2543-2555.

Tham, M. T., F. Vagi, A. J. Morris, R. K. Wood, "Online multivariable adaptive-control of a binary distillation column", *Canadian Journal of Chemical Engineering*, Vol.69, 1991a, pp997-1009.

Tham, M. T., F. Vagi, A. J. Morris, R. K. Wood, "Multivariable and multirate self-tuning control – a distillation column case-study", *IEE Proceedings, Pt.D, Control Theory and Applications*, Vol.138, 1991b, pp9-24.