ELSEVIER

Journal of Process Control 10 (2000) 157–166

JOURNAL OF PROCESS CONTROL

www.elsevier.com/locate/jprocont

Inferential control system of distillation compositions using dynamic partial least squares regression

Manabu Kano*, Koichi Miyazaki, Shinji Hasebe, Iori Hashimoto

Department of Chemical Engineering, Kyoto University, Kyoto 606-8501, Japan

Abstract

In order to control product compositions in a multicomponent distillation column, the distillate and bottom compositions are estimated from on-line measured process variables. In this paper, inferential models for estimating product compositions are constructed using dynamic Partial Least Squares (PLS) regression, on the basis of simulated time series data. It is found that the use of past measurements is effective for improving the accuracy of the estimation. The influence of selection of measurements and sampling intervals on the performance is also investigated. From the detailed dynamic simulation results, it is found that the cascade control system based on the proposed dynamic PLS model works much better than the usual tray temperature control system.

Keywords: Inferential control; Partial Least Squares; Distillation processes

1. Introduction

For product composition control of distillation columns, it is rarely the case that measurements of product compositions are directly used as controlled variables, because on-line accurate measurement of compositions is difficult. Most analyzers, like gas chromatographs and NIR (Near-InfraRed) analyzers, suffer from large measurement delays and high investment and maintenance costs.

1.1. Tray temperature control

In place of composition control using a product analyzer, tray temperature control is widely used. The temperature control is based on the assumption that the product composition can satisfy its specification when an appropriate tray temperature is kept constant at its set-point. For a binary distillation column at constant pressure, the temperature at an end of the column is an exact indicator of the corresponding product composition. However, the temperature variation is very small at the column end and may be difficult to distinguish from measurement noise. Therefore, the use of temperatures removed from the column end is recommended for temperature control. A method used by practitioners to select the appropriate tray is based on the sensitivity of temperatures to changes in feed compositions and in reboiler and reflux flows [1].

In the case where a feed composition or a feed flow rate changes in a multicomponent column, it is quite difficult to keep a product composition at its set-point by using temperature control, because the tray temperature does not correspond exactly to the product composition. In addition, pressure changes also cause temperature variations.

In order to cope with these problems, many approaches have been proposed. The influence of non-key components can be reduced by locating a temperature measurement in the region of the column where their compositions are nearly constant [2]. Yu and Luyben [3] used the other differential temperature for non-key component compensation. Whitehead and Parnis [4] used a weighted average of many differential temperatures for disturbance compensation and also used the temperature difference for pressure compensation. Bozenhardt [5] used multiple temperatures to track the maximum temperature difference between two trays in an alcohol–water–ether column. He found that the position of this maximum difference was strongly correlated to the product composition. However, some of the problems remain unsolved and the performance of the tray temperature control may not be acceptable.

* Corresponding author. Tel.: +81-75-753-5577; fax: +81-75-752-9639.

E-mail address: kano@cheme.kyoto-u.ac.jp (M. Kano).

0959-1524/00/$ - see front matter © 2000 IFAC. Published by Elsevier Science Ltd. All rights reserved.
PII: S0959-1524(99)00027-X

1.2. Inferential control

In order to realize on-line composition control, an inferential control scheme can be used. In inferential control, a product composition is estimated from other measured process variables and the estimates are used for control. Therefore, it is crucial to build a highly accurate inferential model. A first principle model is preferred as far as it is available and provides sufficient accuracy with reasonable computational load. However, if no fundamental model appropriate for real-time use exists, an empirical model determined from process data must be used.

In order to build an empirical model, least squares regression has been widely used. However, when many input variables are used, this method may be unsuccessful due to the highly correlated nature of process data. For example, in distillation, tray temperatures close to each other change in nearly the same way. Applying a linear regression method to such highly correlated data leads to numerical errors and singularity. Even if a numerically accurate solution is reached, the results are not useful for prediction due to the problem called overfitting or overparametrization.

The simplest approach for tackling the problems of correlation and overfitting is to select a few measurements, which are mutually independent, from all measurements. Many articles have been published on this matter [6–9]. However, this simple approach is not optimal, because additional measurements may improve the performance of an estimator.

Brosilow and co-workers [6,7] proposed a composition estimator called the Brosilow estimator, in which temperatures and flow rates were used for estimating unmeasured disturbances and then the derived disturbance values were used to estimate product compositions. This estimator is based on a linearized process model.

In recent years, composition estimators using Partial Least Squares (PLS) regression have been proposed [10–12]. In their work, steady-state inferential models of product compositions were built. That is, their models were constructed by using steady-state data. However, an industrial distillation column is perturbed by disturbances and it is difficult to obtain ideal steady-state data. Even if steady-state data can be obtained, a steady-state estimator may not work well for dynamic operational conditions because it does not take the dynamics into account. Furthermore, Mejdell and Skogestad [11,12] dealt mainly with binary distillation columns. For a multicomponent column, tray temperatures do not correspond exactly to product compositions. Therefore, estimation of product compositions for a multicomponent column is more difficult than that for a binary column. Mejdell and Skogestad [11] have shown that the performance of the steady-state PLS model for a multicomponent column is worse than that for a binary column.

Mejdell and Skogestad [13] compared three different estimators using a linear model of a binary distillation column. They concluded that good control performance could be achieved with the steady-state PCR (Principal Component Regression) estimator, which was almost as good as the dynamic Kalman filter. They also found that the steady-state Brosilow estimator was very sensitive to modeling error for the ill-conditioned plant. Therefore, they recommended using the simple regression estimator. However, the performance of the steady-state PCR estimator deteriorated when feed composition changed. This fact indicates that the achievable performance of the steady-state estimator is limited.

In their paper, only tray temperatures were used as input variables by reason that: (1) the steady-state estimate was not significantly improved by adding measurements, such as flow rates, and (2) the dynamic estimate became even worse by doing that. Result (1) seems quite natural because the steady state of the distillation column can be represented by a few independent variables. In their case, 41 tray temperatures, used as input variables, were enough to represent the steady-state of the column. Result (2) comes from the fact that the manipulated variables cannot affect the product compositions without delay. Therefore, for improving the prediction accuracy by adding manipulated variables to the input variables of the model, past measurements should be used. From the above, the results shown by Mejdell and Skogestad [11,13] seem to indicate the necessity of a dynamic regression estimator, which is the focus of the present paper.

An application of a composition estimator to an industrial packed-bed column was reported [14]. Their inferential model is a static PLS model based on pressure, flow rate, and temperature measurements. This model was built from time-series data, which can be obtained more easily than steady-state data at many different operating conditions.

1.3. Objectives

In the present paper, inferential models are classified into three types, i.e. steady-state models, static models, and dynamic models.

Steady-state models are defined as models determined from steady-state data. When models are built using time-series data, they are called static or dynamic models. Models are classified as dynamic models only when measurements at different sampling times are used as input variables.

In the present paper, a dynamic PLS model, which can estimate the product compositions from multiple temperatures and other on-line measured process variables, is designed for a multicomponent distillation column.

Measurements at several past sampling times are used as input variables in order to incorporate process dynamics into the PLS model. The influence of selection of both input variables and sampling intervals on the performance is also investigated. Furthermore, this paper addresses the performance of on-line composition control systems based on the proposed dynamic PLS model.

2. Problem definition

The problem treated in this paper is to build a dynamic PLS model based on on-line measurements of process variables, such as tray temperatures, reflux flow rate, reboiler heat duty, and pressure. In this section, the example column and the conditions of dynamic simulations are illustrated.

2.1. Example distillation column

The dynamic operation of a multicomponent distillation column was simulated. The schematic diagram of the column is shown in Fig. 1. The column consists of 30 theoretical trays including the reflux drum and the reboiler. The diameter of the column is 1 m. The liquid holdups of the reflux drum and the reboiler are 1.57 and 3.14 m^3, respectively. The feed stream enters the column at the 15th tray and is equimolal flow of methanol, ethanol, 1-propanol, and *n*-butanol. The total flow rate is 128 kmol/h. The set-points of the key components in the distillate and bottom compositions are mole fractions of 0.0010 of propanol and ethanol, respectively. A rigorous model of SPEEDUP® is used for dynamic simulations. The flow dynamics on each tray are expressed by the Francis weir formula. In this study, pressure in the reflux drum is assumed to be kept constant at 1.013×10^5 Pa. That is, the pressure is perfectly controlled by using a total condenser. The pressure drop at each tray changes depending on the vapor flow rate in the column. The base steady-state condition is summarized in Table 1.

Two temperature control loops are used to keep the product compositions at their set-points. Temperatures on the 9th and 22nd trays are used as controlled variables. Reflux flow rate and reboiler heat duty are used as the corresponding manipulated variables. Holdups of the reflux drum and the reboiler are controlled by manipulating the distillate and bottom product flow rates, respectively. PI controllers are used in these control loops. The parameters of the level controllers are determined by using model matching method, in which control parameters are determined so that the closed-loop transfer function of the control system is approximately equivalent to a transfer function with desirable dynamics. The proportional gains of temperature controllers are tuned by trial and error, while the integral times are set to be 0.5 h.

In the simulations, process variables are assumed to be measured every minute. It is also assumed that the propanol mole fraction in the distillate product $x_D^{(\mathrm{PrOH})}$ and the ethanol mole fraction in the bottom product $x_B^{(\mathrm{EtOH})}$ are measured every 10 min.

2.2. Conditions of dynamic simulations

Simulated data for building dynamic inferential models are obtained under the following conditions. The pseudo random binary signals of bounded and varying amplitude (within ±10% of the steady-state value) with

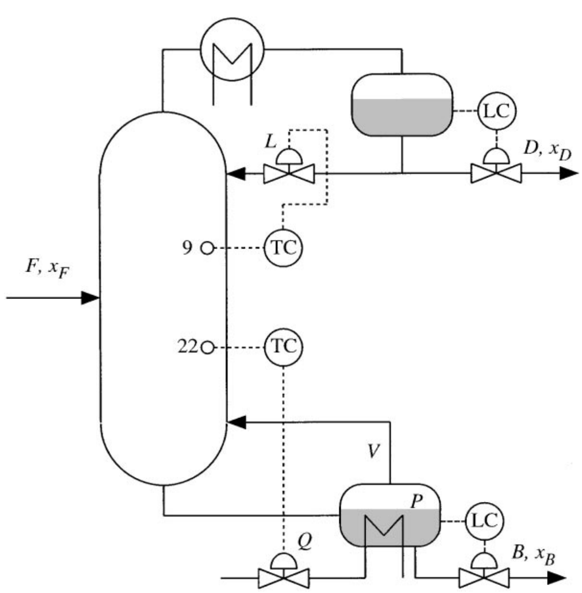

Fig. 1. Tray temperature control scheme of distillation column.

Table 1
Base steady-state condition for the example distillation column

Feed		
F	128.00	kmol/h
T	358.15	K
P	1.150×10^5	Pa
x_F	0.25/0.25/0.25/0.25	–
Reflux drum		
L	110.66 (5.899)	kmol/h (m^3/h)
D	64.00 (3.412)	kmol/h (m^3/h)
T	343.77	K
P	1.013×10^5	Pa
$x_D^{(\mathrm{PrOH})}$	0.0010	–
Reboiler		
V	160.26	kmol/h
B	64.00 (5.965)	kmol/h (m^3/h)
Q	.755	GJ/h
T	386.47	K
P	1.332×10^5	Pa
$x_B^{(\mathrm{EtOH})}$	0.0010	–

different frequency are introduced as component flow rate changes of feed stream during simulations. For realizing slow composition changes, each signal is filtered by a first order lag model. In addition to these random disturbances, the total feed flow rate changes stepwise by $\pm 10\%$ every 2 h, while the fluctuation of the total flow rate is restricted within $\pm 20\%$ of its steady-state value. The total simulation time is 20 h.

Simulated data for validating inferential models are obtained under the almost same conditions as described above. The differences are the seeds of the random signals. The total simulation time is 20 h.

If the time-series data described above are obtained when the temperature controllers are in the automatic mode, the data do not include the operational conditions where the controlled temperatures undergo large changes. However, when inferential composition control is applied instead of temperature control, the tray temperatures fluctuate greatly. Thus, when an inferential model is used for composition control, the accuracy of the estimation may deteriorate due to large changes of the tray temperatures. In order to improve the accuracy, the inferential model must be built using appropriate data, which include large fluctuation of the temperatures. For this purpose, the proportional gains of temperature controllers are changed between 0.5 and 1.5 K every hour in the simulations. Here, K denotes the base controller gain.

The sampling period of the product compositions is 10 min when generating data for modeling. This situation could occur when the compositions are measured by using off-line analyzers. On the other hand, it is not necessary to use this sampling period for model validation. Therefore, when generating data for model validation, the sampling period of the compositions as well as other process variables is set at one minute.

3. Inferential models

In this section, inferential models based on time-series data and steady-state data are built and compared. The influence of selection of both input variables and sampling intervals on the performance is also investigated. The models are evaluated on the basis of mean squared error of prediction (MSEP), which is calculated by applying the models to the validation data,

$$MSEP = \frac{1}{N}\sum_{n=1}^{N}(x(n) - \hat{x}(n))^2 \tag{1}$$

where x is a measurement of the product composition, $\hat{x}$ is its estimate, and N is the number of measurements. Another measure is Explained Prediction Variance (EPV) in percent

$$EPV = \left\{ 1 - \frac{\sum_{n=1}^{N}(x(n) - \hat{x}(n))^2}{\sum_{n=1}^{N}(x(n) - \overline{x})^2} \right\} \times 100 \tag{2}$$

where $\overline{x}$ is a mean value of measurements.

3.1. Partial Least Squares regression

In chemical plants, many variables are measured and those measurements are saved in computers. The effective use of these data is very important for suitable operation.

In the last decade, some multivariate analysis methods, such as principal component analysis (PCA) and PLS [15], have been widely applied for process modeling, monitoring, and control [16–18]. The main advantage of these methods is that they can cope with correlated variables. This characteristic is suitable for analyzing data from chemical processes, because chemical processes are multivariable systems and many variables are mutually correlated.

In these multivariate analysis methods, the transformed independent (orthogonal) variables, which are linear combinations of the original variables and called latent variables, are used.

In the present work, PLS is used for estimating the product compositions from correlated process variables. All input variables are mean centered and their standard deviations are scaled to be unity. No nonlinear transformations are used for dealing with nonlinearity between input variables and product compositions in this paper. When the nonlinearity cannot be ignored, logarithmic transformation of the product compositions is useful [12,19,20]. When using a PLS regression method, the appropriate selection of the number of latent variables is important. The number of latent variables is determined on the basis of the results of cross validation tests and also the results of applying models to validation data.

3.2. Steady-state PLS model

The output variables to be estimated are the propanol and ethanol mole fractions in the products. For clarifying the effect of using pressure measurement, the following two cases are studied:

A1. All 30 tray temperatures are used.
A2. A1 and reboiler pressure are used.

For building a steady-state PLS model, 99 different steady-state data are obtained by considering all possible combinations of the feed flow rates (F) and product compositions (x_D, x_B), as shown in Table 2. The number of the latent variables is selected to be 5, because the mean squared errors (MSE) for a cross validation test are not remarkably decreased by using more than 5 latent variables.

In order to evaluate the steady-state inferential models, they were applied to the time-series validation data. The resulting MSEPs are shown in Table 3 with results for other inferential models. It is found from Table 3 that the estimation accuracy can be much improved by using pressure measurements. However, the MSEPs of the steady-state model (A2) are about five times larger than the results for the dynamic inferential model developed in the following section. It can be safely concluded that the performance of the steady-state PLS model is quite poor. The unacceptable performance of the steady-state PLS model is due to neglecting dynamics.

3.3. Static and dynamic PLS models

In this subsection, static and dynamic inferential models are built from time-series data which can be easily obtained in industrial plants. The conditions for the dynamic simulations are described in the previous section.

In previous studies, temperatures on all trays are used for designing inferential models [10–12]. In the present work, however, a more practical approach is proposed where the product compositions are estimated from fewer temperatures together with reflux flow rate, reboiler heat duty, and pressure. Furthermore, in order to incorporate process dynamics into inferential models, not only measurements at each sampling instant but also past measurements are used as input variables.

Table 2
The steady-state training data

$F^{(EtOH)}$ [kmol/h]	$F^{(PrOH)}$ [kmol/h]	$x_D^{(PrOH)}$ [–]	$x_B^{(EtOH)}$ [–]	F [kmol/h]	$x_D^{(PrOH)}$ [–]	$x_B^{(EtOH)}$ [–]
25.6	25.6	0.0005	0.0005	102.4	0.0005	0.0005
32.0	32.0	0.0010	0.0010	153.6	0.0010	0.0010
38.4	38.4	0.0020	0.0020		0.0020	0.0020

Table 3
Comparison of inferential models

Model	MSEP×10^8 (EPV [%])	
	$x_D^{(PrOH)}$	$x_B^{(EtOH)}$
Steady-state		
A1	3.75 (−5.4)	1.39 (85.6)
A2	1.45 (59.2)	1.26 (86.9)
Static		
B4e,C0	1.12 (68.6)	0.90 (90.6)
Dynamic		
B4e,C21	0.22 (93.8)	0.25 (97.4)
+noise	0.59 (83.3)	0.49 (94.9)

In order to put an inferential model to a practical use, the necessary number of tray temperatures for estimating product compositions should be determined. For this purpose, the following seven cases are studied.

- B1. 2 trays (D:4,11 or B:20,27)
- B2. 3 trays (D:4,9,22 or B:9,22,27)
- B3. 4 trays (4,9,22,27)
- B4. 5 trays (4,9,18,22,27)
- B5. 6 trays (4,9,13,18,22,27)
- B6. 9 trays (3,6,9,12,15,19,22,25,28)
- B7. all 30 trays

where D refers to distillate and B to bottom. As a general rule, the same tray temperatures are used for estimating both the distillate and bottom product compositions. It was found that the estimation accuracy was much worse when only two or three tray temperatures were used. Therefore, the appropriate trays were selected separately for estimating the distillate (D) and bottom (B) product compositions when only two or three tray temperatures were used. The tray selection in these seven cases are almost optimal, because the trays are determined from simulation results including many other cases. Furthermore, for incorporating process dynamics into inferential models, the following 23 cases are investigated. Temperatures at the current sampling instant are used with the following previous value:

- C0. no previous values
- Ck (k = 1,...,20). k min before
- C21. 5, 10, and 15 min before
- C22. 3, 6, 9, 12, 15 and 18 min before

PLS models were built for many cases with different numbers of trays and sampling intervals. These models are evaluated on the basis of MSE given by a cross validation test.

3.4. Selection of tray temperatures

Figs. 2 and 3 show the influence of measurement selection on the performance of the inferential models. The results using the following variables are shown in each graph.

- a. Tray temperatures (T)
- b. T and reflux flow rate (L)
- c. T and reboiler heat duty (Q)
- d. T and pressure at the reboiler (P)
- e. T, L, Q, and P.

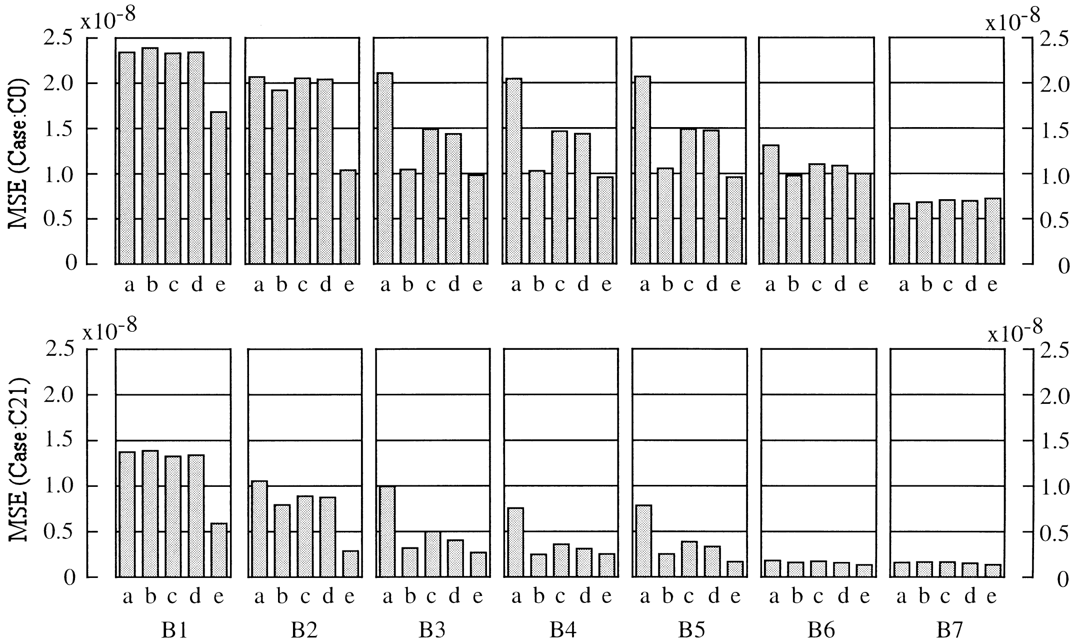

Fig. 2. Influence of measurement selection for estimating the propanol mole fraction in the distillate product $x_D^{(PrOH)}$. Case C0 (top) and C21 (bottom) are compared.

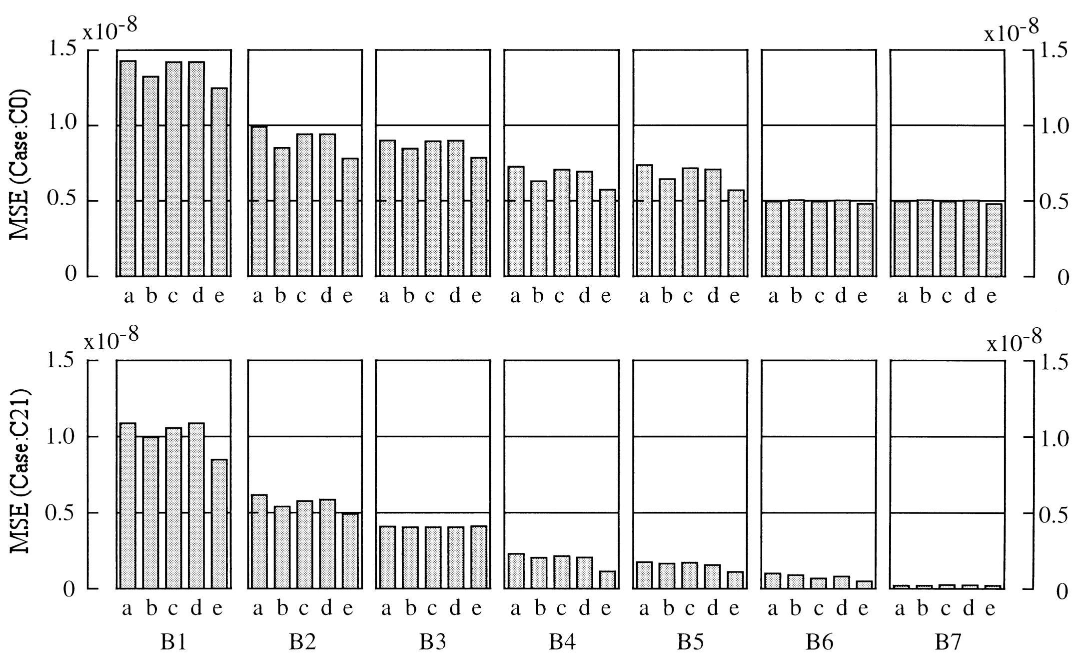

Fig. 3. Influence of measurement selection for estimating the ethanol mole fraction in the bottom product $x_B^{(EtOH)}$. Case C0 (top) and C21 (bottom) are compared.

These figures clarify the effectiveness of using process variables such as reflux flow rate, reboiler heat duty, and pressure, with tray temperatures. The estimation accuracy can be much improved by using not only temperatures but also other process variables. In addition, there seems to be little improvement by using more than five trays, when all process variables (T, L, Q, and P) are used. Similar results have been reported for a binary distillation column [13]. On the basis of these results, five tray temperatures with three other variables (B4e) are adopted in the following sections.

It should be noted that the manipulated variables (L and Q) are determined by using controllers and cannot affect the product compositions without delay. Thus, measurements of the manipulated variables at the current sampling instant may not contribute to the improvement of the estimation accuracy. Therefore, these measurements are not used as input variables in cases C1–C22.

3.5. Selection of sampling intervals

The influence of sampling interval selection for past measurements on the performance of the inferential models is shown in Table 4. Comparing case C0 to C20, the use of temperatures at 16 min earlier (C16) and at 8 min earlier (C8) greatly contributes to the improvement of the accuracy for estimating $x_D^{(PrOH)}$ and $x_B^{(EtOH)}$, respectively. This supports the desirability of using past temperature measurements.

The performance of the inferential models can be improved by increasing the number of measurements. However, case C21 is adopted here, because there seems to be little improvement by choosing C22. From these results, the inferential model (B4e, C21) with 30 input variables was chosen. The number of the latent variables was selected to be 10. The results of applying both the static model (B4e, C0) and the dynamic model (B4e, C21) to the validation data are shown in Fig. 4. The resulting MSEPs are shown in Table 3. The effectiveness of using past measurements is clear from Fig. 4 and Table 3.

Simulated results for other distillation column conditions indicate that the optimal sampling intervals depend on the dynamic characteristics of the column. Table 5 shows the optimal sampling intervals for several column conditions with different holdups of the reflux drum and the reboiler. It should be noted that the levels of the reflux drum and the reboiler are assumed to be kept constant at their set-points and the gains of temperature controllers are constant at their base value when simulated time-series data are generated. It is found that the optimal sampling intervals for estimating $x_D^{(PrOH)}$ and $x_B^{(EtOH)}$ increase with the holdups of the reflux drum and the reboiler.

Table 4
Effect of using past measurements

Case	MSE×10^8	
	$x_D^{(PrOH)}$	$x_B^{(EtOH)}$
C0	0.957	0.576
C2	0.726	0.318
C4	0.587	0.260
C6	0.492	0.195
C8	0.497	0.153
C10	0.405	0.158
C12	0.339	0.184
C14	0.299	0.228
C16	0.274	0.257
C18	0.327	0.305
C20	0.582	0.422
C21	0.252	0.113
C22	0.236	0.113

3.6. Influence of measurement noise

In the previous sections, the ideal case of no measurement noise was studied. In this subsection, for investigating the influence of measurement noise, time series data are generated by adding random noise to the temperature, reflux flow rate, reboiler heat duty, and pressure measurements. The standard deviation of the Gaussian noise is assumed to be 0.1°C for temperature measurements and 10% of standard deviation of time-series data for other measurements. The inferential model (B4e, C21) is rebuilt by using the data with noise.

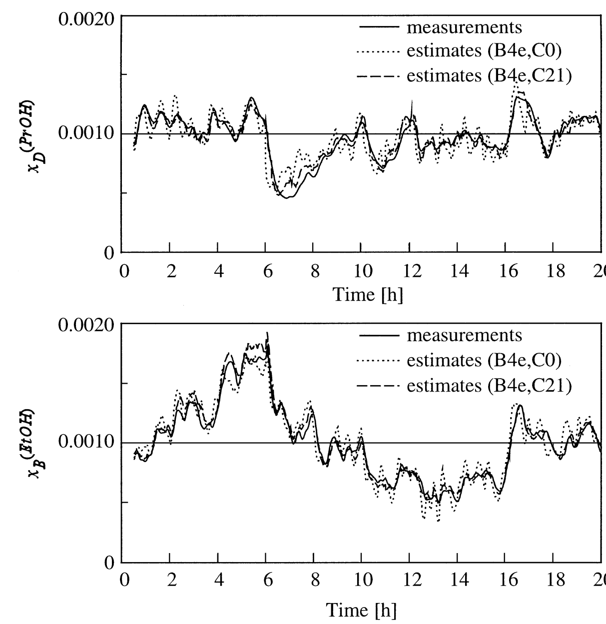

Fig. 4. Estimation results of product compositions by using static and dynamic PLS models. Case (B4e, C0) and Case (B4e, C21) are compared.

Table 5
Optimal sampling intervals for different column conditions

Holdup		Optimal interval	
Reflux drum [m^3]	Reboiler [m^3]	$x_D^{(PrOH)}$ [min]	$x_B^{(EtOH)}$ [min]
0.314	0.314	7	1
0.314	3.142	7	8
0.314	6.284	7	10
1.571	1.571	10	6
1.571	3.142	10	8
1.571	6.284	9	10
3.142	0.314	12	1
3.142	1.571	12	6
3.142	3.142	12	8
3.142	6.284	12	10
6.284	0.314	15	1
6.284	1.571	15	6
6.284	3.142	16	8
6.284	6.284	16	10

The resulting MSEPs are shown in Table 3. The performance is deteriorated by measurement noise. However, it seems to be acceptable, because the dynamic model constructed by using noisy data is better than the steady-state and static models constructed by using data without noise. Furthermore, the influence of measurement noise can be suppressed by filtering or scaling data appropriately. Several scaling methods have been compared [11].

4. Inferential control

The dynamic PLS model is now used for composition control. In this section, conventional tray temperature control, inferential composition control, and cascade control are compared on the basis of MSEs of the controlled variables. For control simulations, the feed disturbances described in the previous section are introduced.

4.1. Tray temperature control

The performance of the tray temperature control system in Fig. 1 is shown in Table 6. The product mole fraction $x_B^{(\mathrm{EtOH})}$ deviates from its set-point when large disturbances are introduced into the process. This result proves that the temperature control is not sufficient for controlling product compositions. The control performance for $x_B^{(\mathrm{EtOH})}$ can be improved by tuning control parameters. However, in such a case, the control performance for $x_D^{(\mathrm{PrOH})}$ will deteriorate due to interaction between the top and bottom control loops.

4.2. Inferential composition control

The inferential control system, in which the estimates of $x_D^{(\mathrm{PrOH})}$ and $x_B^{(\mathrm{EtOH})}$ are controlled by manipulating the reflux flow rate and the reboiler heat duty, is investigated. Multi-loop PI control is used for this purpose. The parameters of the controllers are tuned by trial and error. The simulation results of this control system are shown in Fig. 5.

Comparing the control performance of the temperature control with that of the inferential composition control (cf. Table 6), the inferential control does not seem to function well. In addition, the accuracy of the estimation is much worse when the inferential control is applied. In order to improve the control performance, the controller gain should be increased. However, it is impossible because the influence of estimation error is amplified and as a result the control performance is deteriorated.

The inferential model was built using time-series data, which were obtained by changing the temperature controller gains every hour. In order to clarify the effect of changing controller gains, another inferential model, which was built using time-series data obtained from temperature control with constant parameters, is applied for inferential control. However, this inferential control system was unstable when the same control parameters were used. This unacceptable control performance is due to the deterioration of the estimation accuracy. Since the time-series data are obtained when applying temperature controllers, the data do not include the operational conditions such that the controlled temperatures change significantly. Therefore, when the inferential model is used for composition control, it cannot work well.

These results show that changing the temperature control gains is effective in generating time-series data covering various operational conditions. Such time-series data can also be obtained by changing set-points of controlled variables.

4.3. Cascade control

In general, a controller gain should be increased for improving control performance. However, it is quite

Table 6
Comparison of the performance of three control strategies

Control strategy	MSE of control $\times 10^8$ (MSEP$\times 10^8$)	
	$x_D^{(\mathrm{PrOH})}$	$x_B^{(\mathrm{EtOH})}$
Tray temperature	2.5 (0.14)	8.5 (0.26)
Inferential	4.4 (2.86)	4.3 (1.25)
Cascade	2.6 (1.67)	2.2 (0.56)

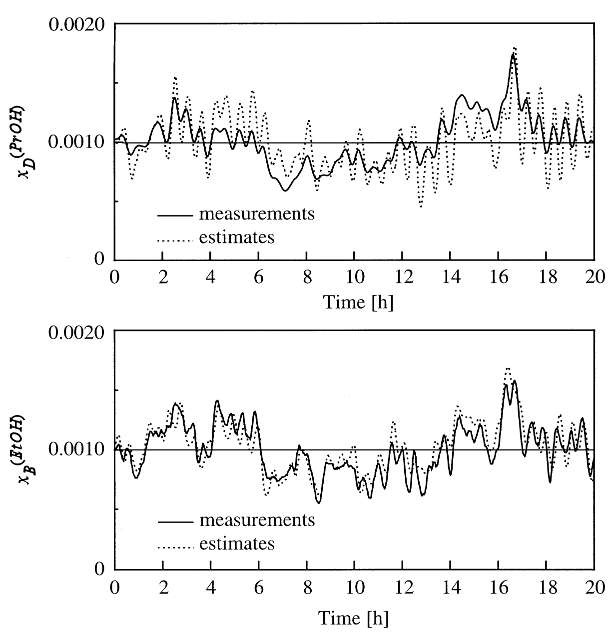

Fig. 5. Closed-loop responses of inferential composition control system.

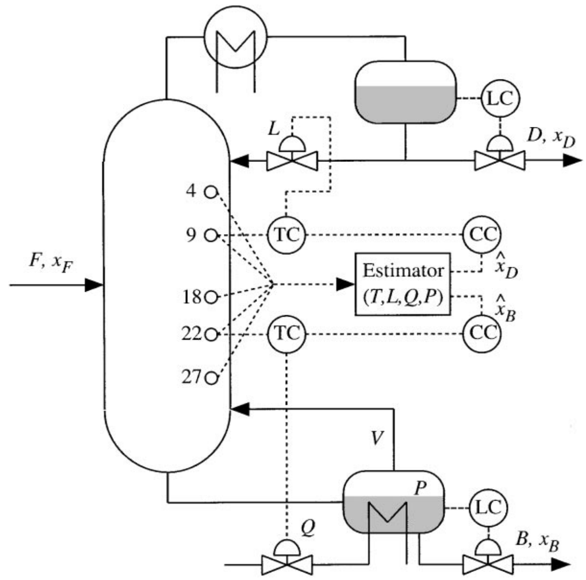

Fig 6. Cascade control scheme with composition estimator.

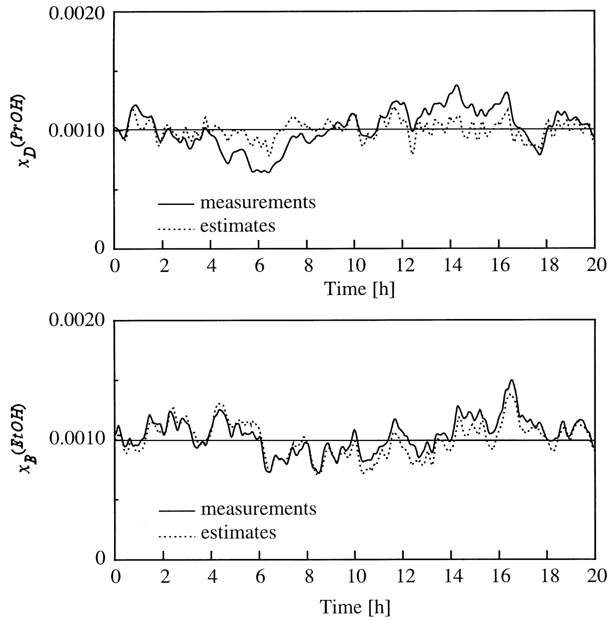

Fig. 7. Closed-loop responses of cascade control system with composition estimator.

difficult when the inferential composition control is adopted, because such a high gain control system tends to suffer from estimation errors. Therefore, another approach must be taken for improving the control performance.

In distillation columns, responses from feed disturbances and a control action of the other control loop $L(Q)$ to the product composition $x_B(x_D)$ have large delays. Since influence of such disturbances and control actions can be detected earlier by monitoring an appropriate tray temperature, the delay can be compensated by using temperature control. However, as the temperature control is not sufficient, the cascade control system shown in Fig. 6 is taken up. In this cascade control system, the same temperature controllers as described before are used for the inner loops, and the set-points of the 9th and 22nd tray temperatures are used as manipulated variables in the outer loops, which are the inferential composition control loops. The composition controllers are PI controllers and their parameters are tuned by trial and error. The simulation results for this control system are shown in Fig. 7. It is found that the performance of this cascade control system is better than the other two control systems (cf. Table 6). In addition, the accuracy of the estimation is much improved in comparison with that of the inferential composition control.

If the control performance as well as the prediction accuracy is not acceptable, the inferential model needs to be rebuilt using time-series data, which are obtained from cascade control system. By executing this iterative modeling, the inferential model will be updated and its performance will be improved. This type of iterative modeling has been proposed by Kresta et al. [10].

5. Conclusion

In the present research, the inferential model, which can estimate the product compositions of the multi-component distillation column from on-line measured process variables such as tray temperatures, reflux flow rate, reboiler heat duty, and pressure, was built by using PLS regression. By using reflux flow rate, reboiler heat duty, and pressure as input variables, a fairly accurate estimation can be attained with fewer temperature measurements. The examples in this paper indicate that five tray temperatures are sufficient for estimating both top and bottom product compositions. It has been found that the performance of an inferential model can be greatly improved by using a dynamic model, i.e. the model based on time-series data, in place of the steady-state or static models.

The application of the inferential model for controlling product compositions was investigated. The cascade control system consisting of inner temperature control loops and outer inferential composition control loops was found to function extremely well. In addition, some simulation results not described in this paper show that the performance of inferential control system based on a dynamic model is much better than that based on a steady-state model demonstrating the advantage of using a dynamic model.

Needless to say, dynamic PLS regression can be applied to various processes in building inferential models. However, we should not try to eliminate the step of developing first principle models, which is important for fully understanding the characteristics of the process. Without such efforts, any results obtained from a statistical method–be it PLS, PCA, or neural network–are unlikely to be truly significant and useful.

Acknowledgements

The authors are grateful to Professor D. E. Seborg (Univ. of California, Santa Barbara) for his helpful comments. The authors also gratefully acknowledge funding from the Japan Society for the Promotion of Science (JSPS-RFTF96R14301).

References

[1] T.L. Tolliver, L.C. McCune, Finding the optimum temperature control trays for distillation columns, InTech 27 (9) (1980) 75–80.
[2] O. Rademaker, J.E. Rijinsdorp, A. Maarleveld, Dynamics and control of continuous distillation units. Elsevier, Amsterdam, 1975.
[3] C.C. Yu, W.L. Luyben, Use of multiple temperatures for the control of multicomponent distillation columns, Ind. Eng. Chem. Res. 23 (1984) 590–597.
[4] D.B. Whitehead and M. Parnis, Computer control improves ethylene plant operation. Hydrocarbon Processing, Nov. (1987) 105–108.
[5] H.F. Bozenhardt, Modern control tricks solve distillation problems, Hydrocarbon Processing, June (1988) 47–50.
[6] R. Weber, C.B. Brosilow, The use of secondary measurements to improve control, AIChE J. 18 (3) (1972) 614–623.
[7] B. Joseph, C.B. Brosilow, Inferential control of processes, AIChE J. 24 (3) (1978) 485–509.
[8] M. Morari, G. Stephanopoulos, Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances, AIChE J. 16 (2) (1980) 247–259.
[9] C. Moore, J. Hackney, D. Canter, Selecting sensor location and type for multivariable processes, in: D.M. Prett, M. Morari (Eds.), Shell process control workshop, Butterworth, Boston, 1987, pp. 291–308.
[10] J.V. Kresta, T.E. Marlin, J.F. MacGregor, Development of inferential process models using PLS, Comp. Chem. Engng. 18 (1994) 597–611.
[11] T. Mejdell, S. Skogestad, Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression, Ind. Eng. Chem. Res. 30 (1991) 2543–2555.
[12] T. Mejdell, S. Skogestad, Composition estimator in a pilot-plant distillation column using multiple temperatures, Ind. Eng. Chem. Res. 30 (1991) 2555–2564.
[13] T. Mejdell, S. Skogestad, Output estimation using multiple secondary measurements: high-purity distillation, AIChE J. 39 (10) (1993) 1641–1653.
[14] H. Fujii, S. Lakshminarayanan, S.L. Shah, Application of the PLS technique to the estimation of distillation tower top composition. Preprints of IFAC symposium on Advanced Control of Chemical Processes (ADCHEM), Banff, Canada, 1997.
[15] W.G. Glen, W.J. Dun III, D.R. Scott, Principal component analysis and partial least squares, Tetrahedron Computer Methodology 2 (1989) 349–376.
[16] P. Nomikos, J.F. MacGregor, Monitoring batch processes using multiway principal component analysis, AIChE J. 40 (8) (1994) 1361–1375.
[17] M. Kano, T. Shiren, M. Ohshima, U. Okamo, S. Aoki, I. Hashimoto, On-line inference system of tube-wall temperature for an industrial olefin pyrolysis plant, AIChE Symposium Series No. 316 93 (1997) 292–296.
[18] S. Lakshminarayanan, S.L. Shah, K. Nandakumar, Modeling and control of multivariable processes: dynamic PLS approach, AIChE J. 43 (9) (1997) 2307–2322.
[19] B. Joseph, C.B. Brosilow, J.C. Howell, W.R.D. Kerr, Multi-temps give better control, Hydrocarbon Processing March (1976) 127–131.
[20] S. Skogestad, M. Morari, Understanding the dynamic behavior of distillation columns, Ind. Eng. Chem. Res. 27 (1988) 1848–1862.