PERGAMON

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Control Engineering Practice 11 (2003) 927–933

CONTROL ENGINEERING PRACTICE

www.elsevier.com/locate/conengprac

Inferential control of distillation compositions: selection of model and control configuration

Manabu Kano*, Natthaphon Showchaiya, Shinji Hasebe, Iori Hashimoto

Department of Chemical Engineering, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Received 6 September 2002; accepted 22 September 2002

Abstract

This article is concerned with a classical problem of how to design an inferential control system for enhancing the performance of distillation composition control. In this article, a new inferential control, termed "predictive inferential control," is proposed. In the predictive inferential control system, future compositions predicted from on-line measured process variables are controlled instead of the estimates of current compositions. The key concept of the predictive inferential control is to realize feedback control with a feedforward effect by the use of the inherent nature of a distillation column. The detailed dynamic simulation results show that the proposed predictive inferential control scheme integrated with cascade control works considerably better than other control schemes. Furthermore, the improvement of the control performance through iterative modeling or control-relevant identification is also demonstrated.

Keywords: Inferential control; Distillation; Partial least squares; Control-relevant identification; Cascade control

1. Introduction

In the last decade or so, inferential control of distillation compositions has been investigated by many researchers. Inferential control became more popular to realize on-line composition control, as chemical processes became more heavily instrumented and process data were more frequently recorded. The design of an inferential control system is a good example to extract useful information from process data and use it for improving process operation. To build an inferential model, which can estimate a product composition from on-line measured process variables, least-squares regression is the simplest approach. However, this method may be unsuccessful due to the highly correlated nature of process data. To solve the collinearity problem, composition estimators using partial least squares (PLS) have been widely used (Kresta, Marlin, & MacGregor, 1994; Mejdell & Skogestad, 1991a, b). In their work, steady-state inferential models of product compositions were built. Mejdell and Skogestad (1993) 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 principal component regression (PCR) estimator, which was almost as good as the dynamic Kalman filter, because the steady-state estimator has a small inherent feedforward effect. An application of a composition estimator to an industrial packed-bed column was reported by Fujii, Lakshminarayanan, and Shah (1997). Their inferential model is a static PLS model based on pressure, flow rate, and temperature measurements.

Later, Kano, Miyazaki, Hasebe, and Hashimoto (2000) further investigated PLS-based inferential models, which can estimate the product compositions of the multicomponent distillation column from on-line measured process variables, such as tray temperatures, reflux flow rate, reboiler heat duty, and pressure. They compared steady-state, static, and dynamic inferential models and found that the estimation accuracy could be greatly improved by the use of dynamic models. Furthermore, a cascade control system, which consisted of an inner temperature control loop and an outer

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

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

0967-0661/02/$ - see front matter
PII: S0967-0661(02)00215-0

inferential composition control loop, was shown to function well.

This article is concerned with a classical problem of how to design an inferential control system for enhancing the performance of distillation composition control. The problem is decomposed into two subproblems: (1) the selection of an inferential model and (2) the selection of a control configuration. These problems are addressed with special emphasis on the inherent feedforward effect of inferential models. Then, a novel control system, based on an inferential model—not for estimating the current product composition but for predicting future product composition—is proposed. This control scheme is termed "predictive inferential control." It is expected that disturbances can be compensated before they affect the product composition by controlling the future composition instead of the current composition. The proposed control system and conventional inferential control systems are compared with applications to a multicomponent distillation column.

2. Problem definition

In this section, the example distillation column (Kano et al., 2000) and the conditions of dynamic simulations are illustrated.

2.1. Example distillation column

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, 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 SPEEDUP^{TM} model 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. The parameters of temperature controllers are tuned by trial and error. Holdups of the reflux drum and the reboiler are controlled by manipulating distillate and bottom product flow rates, respectively. In the simulations, process variables, such as tray temperature, flow rate, pressure, and level, are assumed to be measured every minute. It is also assumed that the propanol mole fraction in the distillate product XD3 and the ethanol mole fraction in the bottom product XB2 are measured every 10 min.

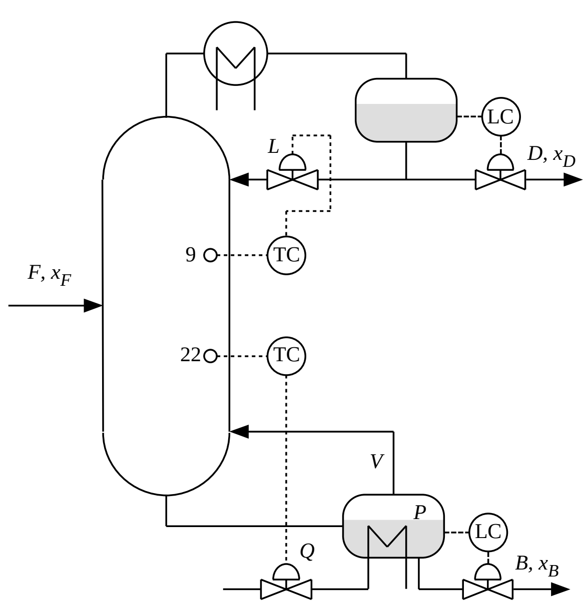

Fig. 1. Tray temperature control scheme of the example distillation column.

Table 1
Base steady-state condition

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
XD3	0.0010	—
Reboiler		
V	160.26	kmol/h
B	64.00 (5.965)	kmol/h (m^3/h)
Q	6.755	GJ/h
T	386.47	K
P	1.332×10^5	Pa
XB2	0.0010	—

2.2. Simulated data for modeling

The pseudo-random binary signals of bounded and varying amplitude (within $\pm 10\%$ of the steady-state value) with different frequencies are introduced as component flow rate changes in the feed stream during simulations. To realize slow composition changes, each signal is filtered by a first-order lag model. In addition to these random disturbances, the feed flow rate changes stepwise by $\pm 10\%$ every 2 h, while the maximum deviation of the flow rate from its steady-state value is restricted within $\pm 20\%$. The total simulation time is 20 h. Simulated data for validating inferential models are obtained under the same conditions as described above.

When inferential composition control is applied instead of temperature control, the tray temperatures will greatly fluctuate. Thus, the estimation accuracy may deteriorate when an inferential model, which is based on the data collected from the process with temperature control, is used for composition control. To improve the accuracy, the inferential model must be built from appropriate data, which include large fluctuations in the temperatures. For this purpose, the set points of temperatures were changed while identification data were generated. The size of set point changes was carefully determined so that the product mole fractions do not exceed the range from zero to 0.003. Changing a set point is an approach used for closed-loop identification.

2.3. Evaluation indexes for estimation and control performance

Inferential models are evaluated on the basis of the explained prediction variance (EPV), which is calculated by the application of the models to the validation data:

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

where x denotes a measurement of the product composition, $\hat{x}$ and $\bar{x}$ are its estimate and mean value, respectively, and N is the number of measurements.

The performance of control systems is evaluated with the mean squared error (MSE) defined as

$$\mathrm{MSE} = \frac{1}{N} \sum_{n=1}^{N} (x(n) - x_{sp}(n))^2, \tag{2}$$

where x_{sp} is the set point of x.

3. PLS-based inferential model

In the last decade, chemometric techniques, such as principal component analysis (PCA) and PLS (Glen, Dun III, & Scott, 1989), have been widely applied for process modeling, monitoring, and control (for example, Wise & Gallagher, 1996; Nomikos & MacGregor, 1994; Kano et al., 1997; Lakshminarayanan, Shah, & Nandakumar, 1997). The main advantage of these methods is that they can cope with correlated input variables. This characteristic is suitable for analyzing data from chemical processes, because chemical processes are multivariable systems and a great number of variables are mutually correlated.

In the present work, PLS is used for estimating the product compositions from correlated process variables. All variables are mean centered and their standard deviations are scaled to be unity. No nonlinear transformation is used for dealing with nonlinearity between input variables and the product compositions because well-known logarithmic transformation of the product compositions did not improve the estimation accuracy in this example. In addition, the number of latent variables is determined on the basis of the results of applying models to validation data. A PLS model in a single output case can be written as

$$\hat{y} = \sum_{i=1}^{r} a_i v_i = \sum_{i=1}^{m} b_i u_i, \tag{3}$$

where $\hat{y}$ is a predicted output variable, u_i is the ith input variable, and v_i is the ith latent variable given as a linear combination of inputs. a_i and b_i are regression coefficients to be estimated. m is the number of input variables and r is the number of latent variables. When input variables are correlated, r should be less than m.

3.1. Steady-state PLS model

The output variables to be estimated are the propanol and ethanol mole fractions in the products (XD3 and XB2). For building a steady-state PLS model, 99 different steady-state data are generated by changing feed flow rate (F) and product compositions. All 30 tray temperatures and a reboiler pressure are used as input variables, and the number of the latent variables is selected to be 5. This model is referred to as SS.

3.2. Static/dynamic PLS model

Both static and dynamic PLS models are built from time-series data, which can be obtained more easily than steady-state data at many different operating conditions. Kano et al. (2000) thoroughly investigated the selection of input variables and sampling intervals. The estimation accuracy can be improved by using not only tray temperatures but also other process variables such as reflux flow rate, reboiler heat duty, and pressure. In addition, there seems to be little improvement by using more than five tray temperatures. Based on these results, five tray temperatures (4th, 9th, 18th, 22nd, and 27th trays) with reflux flow rate, reboiler heat duty, and pressure are used as input variables. In addition, measurements at the current sampling instant are used together with those at 5, 10, and 15 min before when dynamic PLS models are built. It should be noted that the manipulated variables, i.e., reflux flow rate and reboiler heat duty, cannot affect the product compositions without delay. Thus, the manipulated variables at the current sampling instant are not used as input variables in dynamic models. The following four kinds of models are investigated:

STATIC1: static model using all eight variables;

STATIC2: static model using five tray temperatures and a pressure at the bottom;

DYNAMIC1: dynamic model using all eight variables (total 30 input variables);

DYNAMIC2: dynamic model using five tray temperatures and a pressure at the bottom (total 24 input variables).

3.3. Predictive inferential model

Mejdell and Skogestad (1993) pointed out that the temperatures in the middle of the column generally changed slightly faster than at the ends, and therefore the steady-state estimator had a small inherent feedforward effect. In fact, the estimates with STATIC2 precede the measurements as shown in Fig. 2. The cross-correlation r at lag k is defined as $E[x(t)\hat{x}(t+k)]$. The similar cross-correlation plot, in which estimates precede measurements, is obtained when SS model is used. On the other hand, as shown in Fig. 3, there is no lag between measurements and estimates when DYNAMIC1 is used. In addition, the very high cross-correlation at lag zero indicates that DYNAMIC1 can estimate the product compositions with great accuracy.

Kano et al. (2000) proposed the cascade control system, which consisted of an inner temperature control loop and an outer inferential composition control loop. The cascade control system functions better than the conventional inferential control system because the disturbances can be detected and compensated earlier by controlling a tray temperature. In the present work, to detect disturbances before they affect the product compositions and to improve the control performance, an inferential model not for estimating the current product compositions but for predicting future product compositions is proposed. Such an inferential model is termed "predictive inferential model." Conventional and predictive inferential models can be written in general forms as

$$\hat{\boldsymbol{y}}(t) = f_c(\boldsymbol{u}(t), \boldsymbol{u}(t-s_1), \boldsymbol{u}(t-s_2), \ldots), \quad (4)$$

$$\hat{\boldsymbol{y}}(t+\alpha) = f_p(\boldsymbol{u}(t), \boldsymbol{u}(t-s_1), \boldsymbol{u}(t-s_2), \ldots), \quad (5)$$

where $\hat{\boldsymbol{y}}$ is a predicted output vector, and $\boldsymbol{u}$ is an input vector. In addition, s_i and α denote intervals for sampling and prediction, respectively.

By using the proposed predictive inferential model instead of a conventional inferential model in an inferential control system, the control system will be able to compensate the effect of disturbances before they affect product compositions.

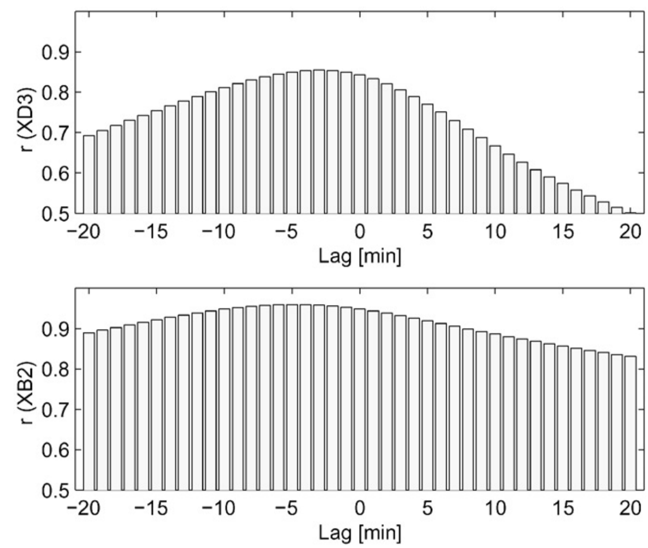

Fig. 2. Cross-correlation between measurements and estimates of the top and bottom product compositions (Model: STATIC2).

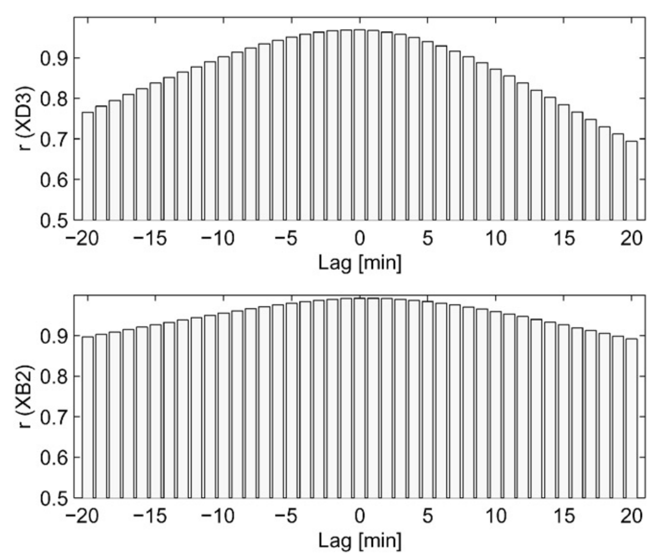

Fig. 3. Cross-correlation between measurements and estimates of the top and bottom product compositions (Model: DYNAMIC1).

4. Inferential control

The inferential models designed in the last section are now used for composition control.

At first, to investigate the selection of inferential models, inferential control systems are introduced for controlling the bottom composition XB2 while the top temperature control system remains. The control results are shown in Fig. 4. The MSEs of both product

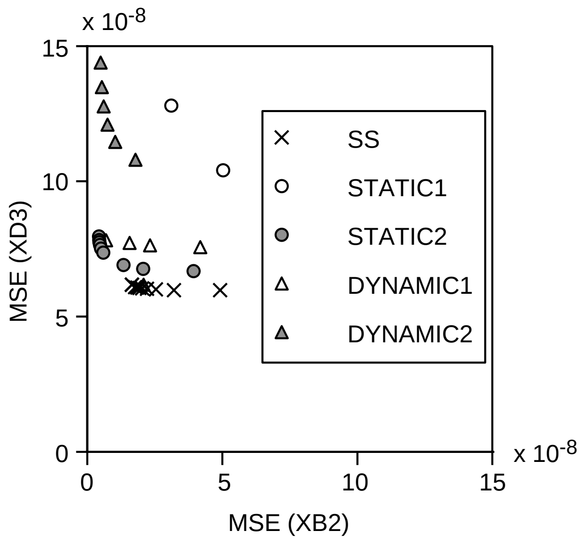

Fig. 4. Control results of the top and bottom product compositions. XD3 is indirectly controlled by TC and XB2 by inferential control.

Table 2
Comparison of temperature control system and inferential control systems using static and dynamic models: MSE $\times 10^7$ (EPV [%]); control performance improvement through iterative modeling

Iteration	0	1	2
Temperature control (TC)			
XD3	6.35		
XB2	9.41		
Total	15.76		
Inferential control with STATIC2			
XD3	9.14 (70.4)	9.02 (72.5)	8.86 (76.6)
XB2	9.15 (75.3)	8.97 (77.4)	8.75 (79.2)
Total	18.29	17.99	17.61
Inferential control with DYNAMIC1			
XD3	8.67 (96.2)	7.94 (97.7)	7.43 (98.0)
XB2	8.75 (98.5)	8.05 (99.1)	7.84 (99.2)
Total	17.42	15.99	15.27

compositions are calculated by changing a proportional gain and an integral time of the inferential controller at the bottom. Several MSEs resulted from each control system are plotted to show the interaction between the top and bottom control loops. The use of inferential control can improve the performance of bottom composition control while the control performance of top product composition deteriorates due to the interaction. The improvement of control performance with SS is limited in comparison with other models. It should be noted that STATIC2 is considerably better than STATIC1. This result indicates that manipulated variables should not be used as inputs when a static model is used for inferential control. On the other hand, DYNAMIC1 is better than DYNAMIC2. Therefore, manipulated variables in the past should be used in a dynamic model to capture the process dynamics and improve the control performance. In addition, STATIC2 outperforms DYNAMIC1. This result contravenes the supposition that better estimation accuracy results in better control performance. The advantage of using static models results from their inherent feedforward effect. Based on these results, STATIC2 and DYNAMIC1 are further investigated in the following sections.

4.1. Conventional inferential control

The inferential control systems, in which the estimates of XD3 and XB2 are controlled by manipulating the reflux flow rate and the reboiler heat duty, are investigated. Multiloop proportional-integral (PI) control is used for this purpose. The simulation results are summarized in Table 2. The inferential control systems do not function well in comparison with TC. To improve the control performance, the inferential model need to be rebuilt by using time-series data, which are obtained from the corresponding inferential control system. Therefore, the inferential control system with the updated inferential model is designed and applied. Executing this iterative modeling updates the inferential model and improves estimation accuracy and control performance. In this example, two iterations are sufficient. However, even after iterative modeling, which is time and effort consuming, inferential control is not at all attractive.

4.2. Predictive inferential control

Predictive inferential models with a different prediction time ($\alpha = 5$, 10, and 15 min) as well as conventional inferential models ($\alpha = 0$ min) are built and compared. The simulation results are summarized in Table 3. The same control parameters are used in all cases to neutralize the effect of tuning.

Table 3 shows that the estimation accuracy becomes worse as prediction time α becomes larger. DYNAMIC1 has considerably better estimation accuracy than STATIC2. However, the control performance of the predictive inferential control systems with DYNAMIC1 and STATIC2 are not significantly different. It should be noted that the best control performance could not be achieved by the use of the most accurate inferential model. These results contradict the presumption that better estimation accuracy results in better control performance. The control performance of the predictive inferential control, however, is worse than that of the temperature control shown in Table 2.

To improve both the estimation accuracy and the control performance, an iterative modeling technique is

Table 3
Comparison of conventional and predictive inferential control systems using static and dynamic models: MSE $\times 10^7$ (EPV [%]); effects of prediction horizon α

α (min)	0	5	10	15
Model: STATIC2				
XD3	9.14	8.97	8.94	10.31
	(70.5)	(66.4)	(64.3)	(63.2)
XB2	9.15	8.02	9.14	11.22
	(75.3)	(75.1)	(74.8)	(73.7)
Total	18.29	16.99	18.08	21.53
Model: DYNAMIC1				
XD3	8.67	9.03	7.97	8.48
	(96.2)	(94.3)	(93.2)	(92.1)
XB2	8.75	8.98	8.02	9.16
	(98.5)	(83.3)	(80.2)	(78.1)
Total	17.42	18.01	15.99	17.54

Table 4
Comparison of predictive inferential control systems using static and dynamic models: MSE $\times 10^7$ (EPV [%]); control performance improvement through iterative modeling

Iteration	0	1	2
Predictive inferential control with STATIC2 ($\alpha = 5$ min)			
XD3	8.97 (66.4)	8.74 (68.1)	8.59 (70.0)
XB2	8.02 (75.1)	7.82 (77.2)	7.69 (79.5)
Total	16.99	16.56	16.28
Predictive inferential control with DYNAMIC1 ($\alpha = 10$ min)			
XD3	7.97 (93.2)	7.54 (94.6)	6.99 (94.8)
XB2	8.02 (80.2)	7.88 (81.5)	7.41 (82.1)
Total	15.99	15.42	14.40

used. The simulation results are summarized in Table 4. Table 4 shows the advantage of using predictive inferential control with the DYNAMIC1 model. After two iterations, the control performance of predictive inferential control with DYNAMIC1 becomes certainly better than that of TC.

4.3. Cascade control

For further improvement of the control performance, the cascade control system suggested by Kano et al. (2000) is investigated. The cascade inferential control scheme is shown in Fig. 5. The results summarized in Table 5 clearly show the advantage of using cascade control together with predictive inferential control. The cascade control system can achieve the excellent control performance, which is almost the same as that of the ideal composition control, in which the product compositions are assumed to be measured online without delay or noise. Furthermore, it is very important from the practical viewpoint that good performance can be given without iteration.

The basic idea of the iterative modeling is that the inferential model must be built from appropriate data. In the present work, the set points of tray temperatures are changed while identification data are generated in order to make the data include large fluctuation of the temperatures. This situation is quite similar to the situation realized by cascade control, because an inferential controller changes a set point of a tray temperature. Therefore, the cascade control system can give good performance without iteration.

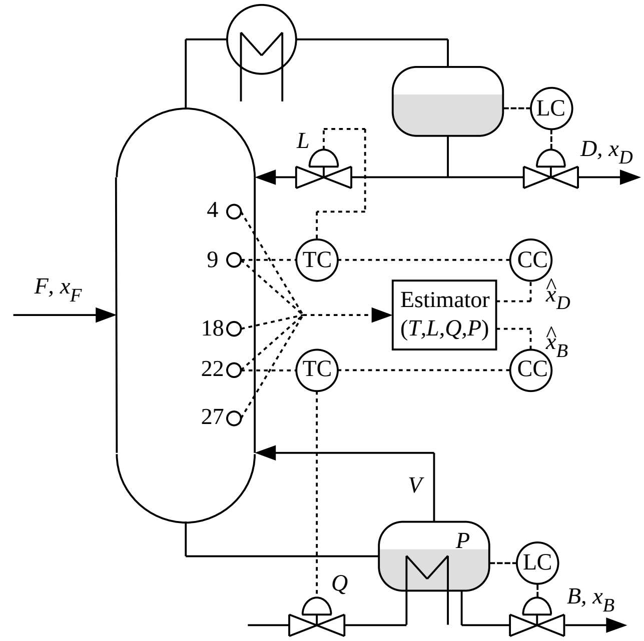

Fig. 5. Cascade inferential control scheme.

Table 5
Evaluation of cascade control systems integrated with conventional and predictive inferential control: MSE $\times 10^7$ (EPV [%]); control performance improvement through iterative modeling

Iteration	0	1	2
Inferential control with DYNAMIC1			
XD3	6.94 (96.2)	6.88 (96.6)	6.82 (96.8)
XB2	7.15 (98.5)	7.11 (98.7)	7.09 (98.9)
Total	14.09	13.99	13.91
Predictive inferential control with DYNAMIC1 ($\alpha = 10$ min)			
XD3	6.62 (93.2)	6.60 (94.0)	6.56 (94.2)
XB2	6.54 (80.2)	6.43 (80.9)	6.42 (81.0)
Total	13.16	13.03	12.98
Ideal composition control (ideal CC)			
XD3	5.15		
XB2	7.77		
Total	12.92		

At the end of this section, it should be emphasized that the proposed predictive inferential control is essentially different from model predictive control (MPC). MPC needs a dynamic process model, which can describe the influence of manipulated variables on controlled variables. Predictive inferential control, however, does not require such a dynamic model. In fact, any measured variable can be used as an input variable in an inferential model. Furthermore, an inferential model can be static. A dynamic model is recommended simply because of its high estimation accuracy.

5. Conclusion

This article is concerned with a classical problem of how to design an inferential control system for enhancing the performance of distillation composition control. By comparing several inferential control systems, it has been demonstrated that the better control performance cannot always be achieved with dynamic models although dynamic models can outperform static models from the viewpoint of estimation accuracy. The advantage of using static models results from their inherent feedforward effect, which is confirmed through cross-correlation plots.

To detect disturbances before they affect the product compositions, an inferential model for predicting the future product compositions, not for estimating the current product compositions, has been proposed. The proposed control scheme, in which the predicted compositions are used as controlled variables, is termed "predictive inferential control."

The inferential control does not function well in comparison with conventional temperature control. To improve control performance, the iterative modeling approach has been tested. As expected, control performance has been improved through the iterations. However, even after iterative modeling, which is time and effort consuming, inferential control is not attractive except for the predictive inferential control with the dynamic model.

For further improvement of the control performance, the cascade control system has been investigated. The advantage of using cascade control together with predictive inferential control has been clearly shown. Furthermore, it is very important from the practical viewpoint that the cascade control system can give good performance without iteration of modeling and control system design.

The results of the present research suggest to use predictive inferential control with a dynamic inferential model within the cascade control configuration to achieve good performance without demanding iterative modeling approach. The proposed control system is a feedback control system with a feedforward control effect.

Acknowledgements

The authors gratefully acknowledge the financial support from the Japan Society for the Promotion of Science (JSPS) Research for the Future Program (JSPS-RFTF96R14301).

References

Fujii, H., Lakshminarayanan, S., & Shah, S. L. (1997). 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.

Glen, W. G., Dun III, W. J., & Scott, D. R. (1989). Principal component analysis and partial least squares. *Tetrahedron Computer Methodology*, *2*, 349–376.

Kano, M., Miyazaki, K., Hasebe, S., & Hashimoto, I. (2000). Inferential control system of distillation compositions using dynamic partial least squares regression. *Journal of Process Control*, *10*, 157–166.

Kano, M., Shiren, T., Ohshima, M., Okamo, U., Aoki, S., & Hashimoto, I. (1997). On-line inference system of tube-wall temperature for an industrial olefin pyrolysis plant. *A.I.Ch.E. Symposium Series No. 316*, *93*, 292–296.

Kresta, J. V., Marlin, T. E., & MacGregor, J. F. (1994). Development of inferential process models using PLS. *Computers in Chemical Engineering*, *18*, 597–611.

Lakshminarayanan, S., Shah, S. L., & Nandakumar, K. (1997). Modeling and control of multivariable processes: Dynamic PLS approach. *A.I.Ch.E. Journal*, *43*(9), 2307–2322.

Mejdell, T., & Skogestad, S. (1991a). Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression. *Industrial and Engineering Chemistry Research*, *30*, 2543–2555.

Mejdell, T., & Skogestad, S. (1991b). Composition estimator in a pilot-plant distillation column using multiple temperatures. *Industrial and Engineering Chemistry Research*, *30*, 2555–2564.

Mejdell, T., & Skogestad, S. (1993). Output estimation using multiple secondary measurements: High-purity distillation. *A.I.Ch.E. Journal*, *39*(10), 1641–1653.

Nomikos, P., & MacGregor, J. F. (1994). Monitoring batch processes using multiway principal component analysis. *A.I.Ch.E. Journal*, *40*(8), 1361–1375.

Wise, B. M., & Gallagher, N. B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, *6*, 329–348.