INTEGRATION OF PRODUCT QUALITY ESTIMATION AND OPERATING CONDITION MONITORING FOR EFFICIENT OPERATION OF INDUSTRIAL ETHYLENE FRACTIONATOR

Hirokazu Kamohara * **Akitoshi Takinami** *
Makoto Takeda * **Manabu Kano** ** **Shinji Hasebe** **
Iori Hashimoto **

* *Showa Denko K.K., Oita 870-0189, Japan*
** *Kyoto University, Kyoto 606-8501, Japan*

Abstract: In this industry-university collaboration, a soft sensor for measuring a key product quality and a monitoring system for testing the validity of the soft sensor were developed to realize highly efficient operation of the ethylene production plant. To estimate impurity concentrations in ethylene products from online measured process variables, dynamic partial least squares (PLS) models were developed. The developed soft sensor can estimate the product quality very well, but it does not function well when the process is operated under unexperienced conditions. Therefore, a monitoring system was developed to judge whether the soft sensor is reliable based on the dynamic PLS model. In addition, simple rules were established for checking the performance of a process gas chromatograph by combining the soft sensor and the monitoring system. The soft sensor and the monitoring system have functioned successfully.

Keywords: soft sensor, partial least squares, statistical process control, monitoring, inferential control, ethylene fractionator

1. INTRODUCTION

Soft sensors are key technologies for producing high quality products when hard sensors of product quality are either not available or too expensive to install. Soft sensors are based on a first principle model, a black-box model, or their combination. Currently, a huge amount of process data is stored in computers, and the effective use of such data is anxiously expected. This situation motivates us to develop a black-box model rather than a first principle model.

Much research has been conducted to develop data-based soft sensors for various processes. A data-based soft sensor, however, does not always function well, because a black-box model is not valid when a process is operated outside certain condition where operation data used for modeling were obtained. The product quality and process performance will deteriorate if the estimates of the soft sensor are blindly believed by operators and used in a control system. To avoid such a situation, the validity of the soft sensor should be monitored online. When a soft sensor is judged to be invalid, the control system and operators should not use the estimates for any purpose.

In this industry-university collaboration, soft sensors for measuring key product qualities and monitoring systems for testing the validity of the soft sensors were developed to realize highly efficient operation of two ethylene fractionators at the SDK (Showa Denko K.K.) Oita plant in Japan. To estimate ethane concentration in ethylene products from online measured

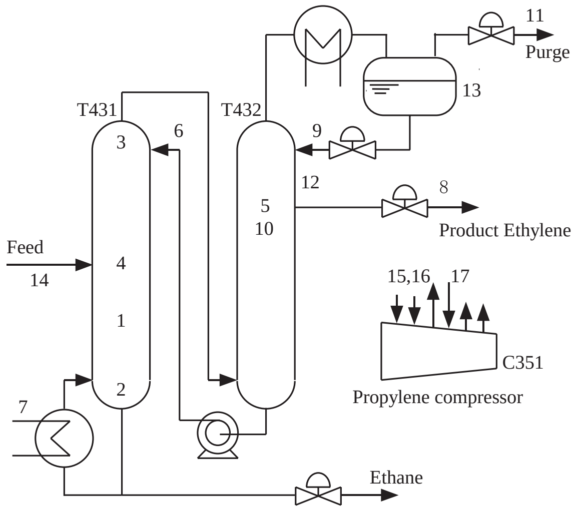

Fig. 1. Schematic diagram of the ethylene fractionator T431/2 at SDK Oita plant.

process variables, dynamic partial least squares (PLS) models were developed. Chemometric techniques, such as principal component analysis (PCA) and PLS, have been widely applied for process modeling, monitoring, and control (for example, Wise and Gallagher, 1996; Nomikos and MacGregor, 1994; Lakshminarayanan, 1997). The main advantage of those 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. The soft sensors developed in the present work are based on PLS. Therefore, in order to test the validity of the soft sensor and judge whether the current operating condition is normal or not, a multivariate statistical process control (MSPC) technique based on PLS is used. In addition, a method is proposed for checking the performance of process gas chromatographs by using the developed soft sensors and monitoring systems. A process gas chromatograph sometimes gives a disturbed and incorrect measurement, which may disturb the operation and deteriorate the process performance. Therefore, it is very important to detect the incorrect measurement and report it to operators. In this article, application results of the PLS-based soft sensor and monitoring system are presented with real industrial data.

2. ETHYLENE FRACTIONATOR

The schematic diagram of the ethylene fractionator, referred to as T431/2, at the SDK Oita plant is shown in Fig. 1. This ethylene fractionator consists of two columns: the bottom column T431 and the top column T432. The feed stream enters the bottom column, and the product ethylene is drawn from the top column. The main specification is ethane concentration in the ethylene product. The ethane concentration must not exceed its upper bound. In order to keep the operation cost as low as possible, the ethane concentration should be kept as high as possible. This fractionator is controlled by using multivariable model predictive control. The number of controlled variables, manipulated variables, and disturbance variables is seven, four, and three, respectively. The controlled variables are ethane concentration and methane concentration in the ethylene product, T431 tray #29 temperature (1), T431 differential pressure, T432 differential pressure, condenser pot level, and reboiler pot level. Manipulated variables are T431 reboiler flow rate (7), T432 internal reflux flow rate (10), T432 purge flow rate (11), and T432 top pressure (13). The disturbance variables are T431 feed flow rate, T431 feed ethane concentration (14), and C351 #4 suction pressure (17). Here, C351 is a propylene compressor. Its #4 suction pressure affects propylene refrigerant temperature and then reboiler heat duty. The numbers in parentheses correspond to those shown in Fig. 1.

3. SOFT SENSOR

Soft sensors are key technologies for producing high quality products when quality hard-sensors are either not available or too expensive to install. Distillation compositions are such quality variables. The compositions can be measured by using, for example, gas chromatographs and near-infrared analyzers, but gas chromatographs suffer from large measurement delays, and most analyzers suffer from high investment and maintenance costs. Therefore, many researchers have investigated soft sensors and inferential control of distillation compositions.

3.1 *Overview*

In order to build a soft sensor by using past operation data stored automatically in a computer, linear or nonlinear black-box models have been widely used. When many process variables are used as input variables, the highly correlated nature of process data must be taken into account. In distillation processes, for example, tray temperatures close to each other change in nearly the same way. Applying a statistical modeling method to such highly correlated data causes a collinearity problem.

The simplest approach for tackling the collinearity problem is to select a few variables, which are mutually independent, from all process variables. Many articles have been published on this matter, for example, Weber and Brosilow (1972), Joseph and Brosilow (1978), Morari and Stephanopoulos (1980), and Moore et al. (1987). However, this simple approach would not be optimal, because additional measurements may improve the performance of an estimator.

To solve the collinearity problem, composition estimators using PLS have been widely used (Kresta et al., 1994; Mejdell and 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 PCR (Principal Component Regression) estimator, which was almost as good as the dynamic Kalman filter, because the steady-state estimator has an inherent feedforward effect. The inherent feedforward effect was investigated in more detail by Kano et al. (2002). They suggested using predictive inferential control with a dynamic inferential model within the cascade control configuration to achieve good performance without demanding the iterative modeling approach. The proposed control system is a feedback control system with a feedforward control effect. An application of a composition estimator to an industrial packed-bed column was reported by Fujii et al. (1997). Their inferential model is a static PLS model based on pressure, flow rate, and temperature measurements. Kano et al. (2000) further investigated PLS-based inferential models, which can estimate the product compositions of the multicomponent distillation column from on-line measured process variables. They compared steady-state, static, and dynamic inferential models and found that the estimation accuracy could be greatly improved by using dynamic models.

In the present work, dynamic PLS is used for estimating the output variable, i.e., the ethane concentration in the ethylene product, from correlated process variables.

3.2 *Dynamic PLS Model*

Kano et al. (2000) thoroughly investigated the selection of input variables and sampling intervals. They concluded that the estimation accuracy was improved by using not only tray temperatures but also other process variables such as reflux flow rate, reboiler heat duty, and pressure. In addition, they strongly recommended using a dynamic inferential model to improve estimation accuracy and control performance. Based on these results, almost all measured process variables are considered as candidates for input variables. Then, optimal selection of input variables and sampling intervals was carried out by trial and error. The selection was mainly based on the correlation analysis, engineers' knowledge, and validation tests.

Table 1. Input variables of PLS model.

No.	Variable
1.	T431 tray #29 temperature
2.	T431 bottom temperature
3.	T431 top temperature
4.	T431 tray #37 temperature
5.	T432 tray #129 temperature
6.	Flow rate from T432 to T431
7.	T431 reboiler flow rate
8.	Product ethylene flow rate
9.	T432 reflux flow rate
10.	T432 internal reflux flow rate (=9-8)
11.	T432 purge flow rate
12.	T432 reflux ratio
13.	T432 top pressure
14.	T431 feed ethane concentration
15.	C351 #2 discharge pressure
16.	C351 #2 discharge temperature
17.	C351 #4 suction pressure

C351 is a propylene compressor. Propylene is used for heating and cooling at the reboiler and the condenser.

For building a dynamic PLS model, the operation data obtained in the period from December 2001 to February 2002 were used. A part of the operation data, which represents abnormal operation or sensor malfunction, was excluded. Although measurements of process variables are stored in the computer every minute, moving averages of five points are used for modeling to reduce measurement noise. Therefore, the minimum sampling interval, which is available for modeling, is five minutes. All variables were mean-centered and scaled to have unit variances. The selected input variables are listed in Table 1. A total of 17 process variables, including temperatures, pressures, flows, and reflux ratio, were selected. In addition, measurements at the current sampling instant were used together with those at 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275 minutes before when a dynamic PLS model was built. Therefore, the total number of input variables is 340 ($= 17 \times 20$). The developed estimator has the form of

$$\hat{y}(t) = \sum_{i=1}^{17} \sum_{j=1}^{20} \alpha_{ij} x_i(t - s_j) \tag{1}$$

where $\hat{y}(t)$ and $x_i(t)$ denote the estimated product concentration and the input variables at the time t, respectively. s_j is sampling instants and α_{ij} is regression coefficients.

The number of latent variables was determined on the basis of cross-validation tests. Twenty latent variables are used in the final model. Furthermore, no nonlinear transformation is used for dealing with nonlinearity between input variables and the output, i.e., the ethane concentration, because a well-known logarithmic transformation of the product quality did not improve the estimation accuracy in this application.

3.3 *Estimation Results*

The developed soft sensor has been applied to the ethylene fractionator T431/2. The estimation results in three periods are shown in Fig. 2: (A) Nov. 10 2001 through Nov. 30 2001, (B) May 1 2002 through June 13 2002, and (C) Aug. 1 2002 through Aug. 31 2002. Here, measurements and estimates of ethane concentration (Ethane conc.) and estimation errors (Error) are scaled. The soft sensor functions very well. In Fig. 2(top), it is difficult to distinguish between measurements and estimates because estimation errors are so small. The relative estimation error is less than 10 % at almost all times except in the period from 1100 to 1200 hours. In this period, a trouble occurred and thus operators changed the operating condition considerably.

The product quality will deteriorate if estimates of the soft sensor are blindly believed by the control system and operators, while the soft sensor does not function well. To avoid such a situation, an operating condition should be monitored online and the validity of the soft sensor should be tested. The control system and operators should not use the estimates for any purpose when the soft sensor is judged to be invalid. In the next section, a monitoring system for testing the validity of the soft sensor is proposed.

4. MONITORING SYSTEM

A black-box model does not always function well. Since it is a data-based model, a black-box model is valid only when a process is operated within a certain condition where operation data used for modeling were obtained. Therefore, to successfully apply a data-based soft sensor to an industrial process, the validity of the soft sensor should be tested. In other words, the operating condition should be monitored online to judge whether the estimated value is reliable or not.

The soft sensor developed in the present work is based on PLS. Therefore, the operating condition, where operation data used for modeling were obtained, can be easily defined in the subspace spanned by major latent variables retained in the PLS model and its orthogonal complement space. In order to test the validity of the soft sensor, it is necessary to judge whether the current operating condition is inside the defined operating condition or outside. Such a monitoring system can be realized by using a multivariate statistical process control (MSPC) technique based on PLS.

4.1 *Overview*

Chemical processes are multivariable systems consisting of a large number of mutually correlated variables. MSPC was developed to monitor such multivariable processes. The original Shewhart-type control chart for correlated variables is the Hotelling T^2 control chart. Jackson (1959) used principal component analysis (PCA) and proposed a T^2 control chart for principal components. Later, Jackson and Mudholkar (1979) investigated PCA as a tool of MSPC and introduced a residual analysis. The control chart was introduced for the sum of squared residuals Q as well as T^2.

$$T^2 = \sum_{r=1}^{R} \frac{t_r^2}{\sigma_{t_r}^2} \qquad (2)$$

$$Q = \sum_{p=1}^{P} (x_p - \hat{x}_p)^2 \qquad (3)$$

where t_r is the r-th principal component score and $\sigma_{t_r}^2$ is the variance of t_r. x_p and $\hat{x}_p$ are a measurement of the p-th variable and its predicted (reconstructed) value, respectively. R and P denote the number of principal components retained in the PCA model and the number of process variables, respectively. The T^2 statistic is a measure of the variation within the PCA model, and the Q statistic is a measure of the amount of variation not captured by the PCA model. These two statistics, T^2 and Q, can be used for PLS-based MSPC by substituting latent variables for principal components. Many successful applications of PLS-based MSPC to industrial data have shown its practicability (Kourti et al., 1995; Macgregor and Kourti, 1995; Kourti and MacGregor, 1995).

4.2 *Monitoring System Design and Results*

The monitoring system is based on the dynamic PLS model designed for the soft sensor. The number of input variables is 340, and the number of latent variables is 20. That is, $R = 20$ and $P = 340$ in Eqs. (2) and (3). The developed monitoring system integrated with the soft sensor has been applied to the ethylene fractionator T431/2. The control limits of T^2 and Q are 200 and 100, respectively. These control limits were determined so that they represent 99% confidence limits. The monitoring results in three periods

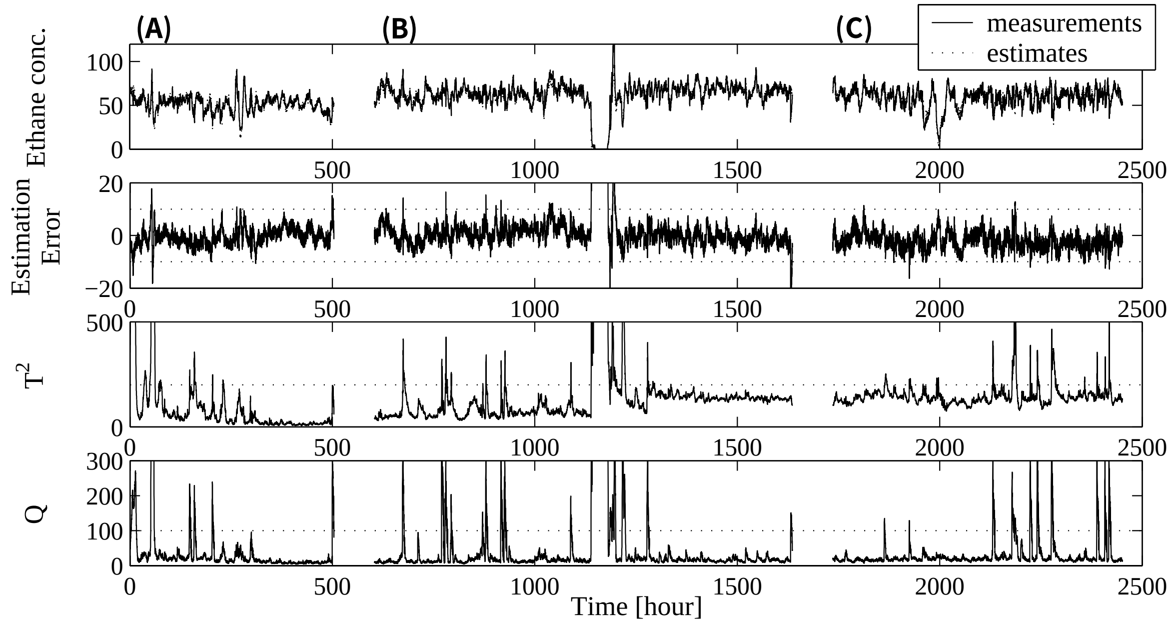

Fig. 2. Estimation and monitoring results of ethane concentration in the ethylene product.

are shown in Fig. 2. In the period from 1100 to 1200 hours, when the process was operated under the abnormal operating condition and the estimation error was crucial, both T^2 and Q statistics considerably exceeded their control limits. That is, the monitoring system indicates that the operating condition is abnormal and the soft sensor is not reliable in this period. It should be noted here that "abnormality" in this context does not necessarily mean the occurrence of faults. It means that the current operating condition is different from the past operating condition in which data used for modeling were obtained. In Fig. 2, estimation errors tend to become large and exceed ± 10 when T^2 or Q exceeds its control limit. Spikes of T^2 or Q are observed when load on the tower increases. These results demonstrate that the developed monitoring system can judge whether the soft sensor is reliable in the current operating condition or not. The soft sensor and the monitoring system have functioned successfully.

Furthermore, the T^2 statistic indicates that the operating condition in the period (C) differs from that in (A). The operating condition in (A) is quite similar to the normal operating condition, where reference data were obtained. On the other hand, an improved multivariable control system was installed and the process has been operated near the optimal condition in (C). Therefore, T^2 in (C) is larger than that in (A). Estimation errors are quite small even when T^2 indicates the significant operating condition change. That is, the developed soft sensor is valid under a sufficiently wide range of operating conditions.

Table 2. Rules for checking the performance of a process gas chromatograph.

Error of GC and soft sensor	Operating condition (Reliability of soft sensor)	Performance of GC
small	good	good
small	bad	good (unclear)
large	good	bad
large	bad	good

4.3 *Validation of GC*

The developed monitoring system can be used for another purpose. A process gas chromatograph (GC) sometimes gives a disturbed and incorrect measurement, and such an incorrect measurement disturbs the operation and deteriorates the process performance. In addition, a GC needs to be repaired if its measurements are not reliable. Therefore, it is important to detect the incorrect measurement and report it to operators. The performance of a GC can be checked by comparing its measured value with an estimated value of the soft sensor. The rules summarized in Table 2 were established in the present work. Here, an error of GC and soft sensor is a difference between a measured value of GC and an estimated value of the soft sensor. The operating condition is judged by the developed PLS-based monitoring system. Estimates of the soft sensor are reliable when operating conditions are good. Therefore, the soft sensor and the GC function well, when an error is small and the operating condition is good. On the other hand, a GC measurement is judged to be incorrect if an error is large while the operating condition is good, because the soft sensor must be reliable when the operating condition is good. In such a situation, a large estimation error is caused by malfunction of GC.

5. CONCLUSION

In this industry-university collaboration, a soft sensor for measuring ethane concentrations in ethylene products and a monitoring system for testing the validity of the soft sensor were developed to realize highly efficient operation of two ethylene fractionators at the SDK Oita plant.

To estimate the ethane concentration from online measured process variables, dynamic PLS models were developed. The developed soft sensor can estimate the product quality very well. The relative estimation error is less than 10 % at T431/2; however, the soft sensor does not function well when the process is operated under abnormal conditions. To test the validity of the soft sensor, a monitoring system was developed based on the dynamic PLS model designed for the soft sensor. The monitoring system can judge whether the soft sensor is reliable or not. The usefulness of the developed monitoring system was demonstrated with real operation data. In addition, the performance of a process gas chromatograph can be checked by using the soft sensor and the monitoring system. Simple rules were established for this purpose. The soft sensor and the monitoring system have functioned successfully in the SDK Oita plant.

REFERENCES

Fujii, H., S. Lakshminarayanan, and S. L. Shah (1997). Application of the PLS technique to the estimation of distillation tower top composition. *Prep. of ADCHEM*, Banff, Canada.

Jackson, J. E. (1959). Quality Control Methods for Several Related Variables. *Technometrics*, **1**, 359-377.

Jackson, J. E. and G. S. Mudholkar (1979). Control Procedures for Residuals Associated with Principal Component Analysis. *Technometrics*, **21**, 341-349.

Joseph, B. and C. B. Brosilow (1978). Inferential Control of Processes. *AIChE J.*, **24**, 485-509.

Kano, M., K. Miyazaki, S. Hasebe, and I. Hashimoto (2000). Inferential control system of distillation compositions using dynamic partial least squares regression. *J. Proc. Cont.*, **10**, 157-166.

Kano, M., N. Showchaiya, S. Hasebe, and I. Hashimoto (2002). Inferential control of distillation compositions: selection of model and control configuration. *Control Engineering Practice* (in press)

Kourti, T. and J. F. MacGregor (1995). Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and intelligent laboratory systems*, **28**, 3-21.

Kourti, T., P. Nomikos, and J. F. MacGregor (1995). Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *J. Proc. Cont*, **5**, 277-284.

Kresta, J. V., T. E. Marlin, and J. F. MacGregor (1994). Development of inferential process models using PLS. *Comp. Chem. Engng*, **18**, 597-611.

Lakshminarayanan, S., S. L. Shah, and K. Nandakumar (1997). Modeling and control of multivariable processes: dynamic PLS approach. *AIChE J.*, **43**, 2307-2322.

MacGregor, J. F. and T. Kourti (1995). Statistical process control of multivariate processes. *Control Eng. Practice*, **3**, 403-414.

Mejdell, T. and S. Skogestad (1991a). Estimation of distillation compositions from multiple temperature measurements using Partial-Least-Squares regression. *Ind. Eng. Chem. Res.*, **30**, 2543-2555.

Mejdell, T. and S. Skogestad (1991b). Composition estimator in a pilot-plant distillation column using multiple temperatures. *Ind. Eng. Chem. Res.*, **30**, 2555-2564.

Mejdell, T. and S. Skogestad (1993). Output estimation using multiple secondary measurements: high-purity distillation. *AIChE J.*, **39**, 1641-1653.

Moore, C., J. Hackney, and D. Canter (1987). Selecting sensor location and type for multivariable processes In: *Shell Process Control Workshop* (D. M. Prett and M. Morari, Ed.). pp.291-308, Butterworth, Boston.

Morari, M. and G. Stephanopoulos (1980). Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances. *AIChE J.*, **26**, 247-259.

Nomikos, P. and J. F. MacGregor (1994). Monitoring batch processes using multiway principal component analysis. *AIChE J.*, **40**, 1361-1375.

Weber, R. and C. B. Brosilow (1972). The use of secondary measurements to improve control. *AIChE J.*, **18**, 614-623.

Wise, B. M. and N. B. Gallagher (1996). The process chemometrics approach to process monitoring and fault detection. *J. Proc. Cont.*, **6**, 329-348.