Development of a New Soft Sensor Method Using Independent Component Analysis and Partial Least Squares

Hiromasa Kaneko, Masamoto Arakawa, and Kimito Funatsu
Dept. of Chemical System Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

DOI 10.1002/aic.11648
Published online October 29, 2008 in Wiley InterScience (www.interscience.wiley.com).

Soft sensors are used widely to estimate a process variable which is difficult to measure online. One of the crucial difficulties of soft sensors is that predictive accuracy drops due to changes of state of chemical plants. To cope with this problem, a regression model can be updated. However, if the model is updated with an abnormal sample, the predictive ability can deteriorate. We have applied the independent component analysis (ICA) method to the soft sensor to increase fault detection ability. Then, we have tried to increase the predictive accuracy. By using the ICA-based fault detection and classification model, the objective variable can be predicted, updating the PLS model appropriately. We analyzed real industrial data as the application of the proposed method. The proposed method achieved higher predictive accuracy than the traditional one. Furthermore, the nonsteady state could be detected as abnormal correctly by the ICA model. © 2008 American Institute of Chemical Engineers *AIChE J,* 55: 87–98, 2009
Keywords: soft sensor, process control, independent component analysis, partial least squares, fault detection

Introduction

In operating chemical plants, an operator has to monitor operating condition of the plants and control process variables, for example, temperature, pressure, liquid level, concentration of products, and so on. Therefore, these variables need to be measured online, but all of them are not easy to measure online because of technical difficulty, large measurement delays, high investment cost, and so on.

In chemical plants, soft sensors are used widely to estimate a process variable which is difficult to measure online. An inferential model is constructed between variables which are easy to measure online and one which is difficult to measure online, and a value of an objective variable is estimated by the model. Mainly, partial least squares (PLS) method[1–16] is used as a modeling method for the soft sensors. Also, principle component regression (PCR) method,[14,17,18] nonlinear PLS method,[7,10,19–23] artificial neural network,[10,13,18,24–51] support vector machine based regression method,[52–58] and so on, are researched as the soft sensor method. By using soft sensors, a value of objective variables can be estimated with high accuracy.

However, soft sensors have some practical difficulties. One of the crucial difficulties is that predictive accuracy drops due to changes of state of chemical plants, catalyst performance loss, sensor and process drifting, and so on. If a problem of the degradation of soft sensors is not solved, it is difficult to identify reasons of abnormal situations. On the site of plants, when the prediction error of an objective variable **y** is above a threshold, abnormal situation is detected. There is no effective method to judge whether the reason of it is the trouble of process, the trouble of analyzer, or the degradation of the soft sensor, under the circumstances.

The regression model can be updated with a new sample to solve the degradation of the soft sensor model.[59–63] If the model is updated with an abnormal sample, the predictive

Correspondence concerning this article should be addressed to H. Kaneko at hkaneko@chemsys.t.u-tokyo.ac.jp.

ability can deteriorate. The abnormal sample needs to be detected, but by checking only the abnormal data by using the prediction error of **y**, it is difficult to detect the abnormal samples and to judge whether the reason of them is the trouble of process or the trouble of analyzer. Thus, it is desirable that a fault detection model with high accuracy be constructed to increase prediction ability of updating soft sensor model.

Then, we have applied the independent component analysis (ICA)[64] method to increase fault detection ability and to classify the abnormal situation. ICA is a method that is used in many fields such as signal processing. ICA can effectively resolve the independent components from the measured mixed signals without any additional information about the source signals. In other words, ICA is a method for transforming observed multivariate data into statistically independent components expressed as the linear combinations of observed variables. It has been widely applied in fields such as spectral analysis,[65,66] quantitative structural–property relationships,[67] and so on. If reasons of changes of each process variable measured in chemical plants are mutually independent, it is conceivable that each of them can be extracted by using the ICA method.

ICA is often compared with principal component analysis (PCA). PCA can extract mutually uncorrelated components from explanatory variables, through making full use of the second-order central moment. On the other hand, ICA can extract mutually independent components from explanatory variables, through making full use of the fourth-order central moment. ICA attracts much attention because ICA uses the higher-order statistical characteristics of the source than PCA, and the fourth-order central moment represents high values for components including outliers. In fact, the ICA method is applied to statistical process control[68–77] and soft sensors.[48,55] In addition, if the fault detection ability improves, it could contribute to the abnormal diagnosis.[78,79]

We have developed the new soft sensor method that combines ICA and PLS. Monitoring methods based on ICA have been developed recently. However, even if traditional methods can detect the abnormal situation, it is difficult to classify the reason of it appropriately.[80] Independent components extracted from explanatory variables by using ICA method are mutually independent and sensitive to outliers. Therefore, we propose to select independent components, which are sensitive to the same outliers, and classify the reason of abnormal situation by using these components individually. By using the ICA-based fault detection and classification model, it is conceivable that the objective variable can be predicted, updating the PLS model with only normal samples.

To verify prediction ability and fault detection and classification ability of the proposed method, we applied this method to real industrial data. We analyzed the data obtained from the operation of the distillation column at Mizushima works, Mitsubishi Chemical. We prepared 10,000 samples, then first 4000 samples are training data, and we tried to predict next 6000 samples. To verify the superiority of the proposed method, below four methods are applied to real industrial data.

A. Do not update the PLS model, and detect the abnormal data by using the prediction error of **y**.

B. Update the PLS model, and detect the abnormal data by using the prediction error of **y**.

C. Update the PLS model, and detect and classify the abnormal data by using the PCA-based fault detection model[81] and the prediction error of **y**.

D. Update the PLS model, and detect and classify the abnormal data by using the ICA-based fault detection and classification model and the prediction error of **y**.

Method A is a traditional soft sensor method, and method D is the proposed one. By comparing the results of them, we verified the superiority of the proposed method. Furthermore, by using the ICA-based fault detection and classification model, the nonsteady state, which could not be detected as abnormal by the prediction error of **y**, could be detected.

PLS

Modeling a relationship between explanatory variables $\mathbf{X} \in R^{n\times d}$ and objective variable $\mathbf{y} \in R^{n\times 1}$ (where n is the number of samples and d is the number of variables) is done by using MLR, which works well as long as **X** variables are few and uncorrelated. However, it is impossible to construct a regression model when the number of **X**-variables is more than the number of samples. Thus, attempts to pretreat **X** have been proposed. One of these methods is PCR. It is a method that constructs a regression model of **y** by means of the score matrix $\mathbf{T} \in R^{n\times a}$ (where a is the number of components) calculated by PCA. Since **T** is mutually orthogonal, a stable model would be constructed by the PCR method.

PLS is a method for relating **X** and **y**, by a linear multivariate model, but goes beyond traditional regression methods in that it models also the structures of **X** and **y**. In PLS modeling, the covariance between score vector $\mathbf{t}_i \in R^{n\times 1}$ and **y** is maximized. Generally, PLS models have higher predictive power than those of MLR and PCR.

A PLS model consists of two equations as follows:

$$\begin{aligned} \mathbf{X} &= \mathbf{TP}' + \mathbf{E} \\ \mathbf{Y} &= \mathbf{Tq} + \mathbf{f} \end{aligned} \tag{1}$$

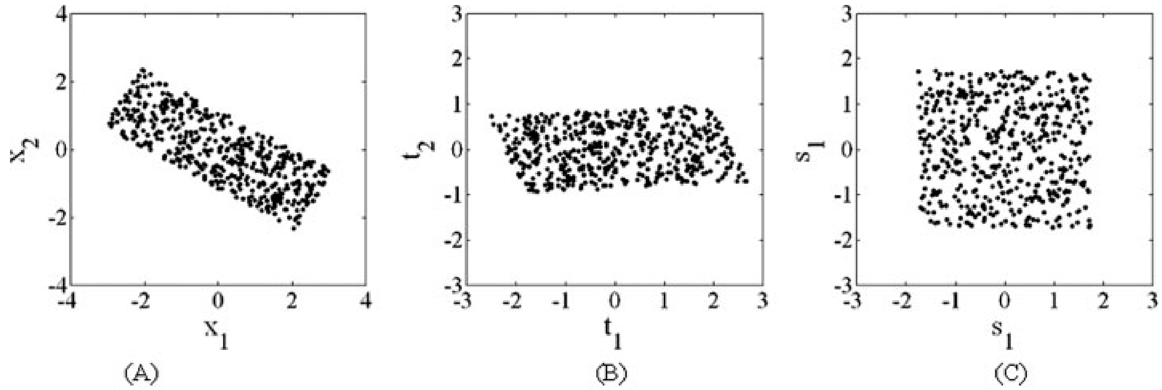
where $\mathbf{P} \in R^{d\times a}$ is an **X**-loading matrix, $\mathbf{q} \in R^{1\times a}$ is a **y**-loading vector, $\mathbf{E} \in R^{n\times d}$ is a matrix of **X** residuals, and $\mathbf{f} \in R^{n\times 1}$ is a vector of **y** residuals. The PLS-regression model is as follows:

$$\begin{aligned} \mathbf{y} &= \mathbf{Xb} + \mathbf{const} \\ \mathbf{b} &= \mathbf{W}(\mathbf{P}'\mathbf{W})^{-1}\mathbf{q} \end{aligned} \tag{2}$$

where $\mathbf{W} \in R^{d\times a}$ is an **X**-weight matrix and $\mathbf{b} \in R^{d\times 1}$ is a vector of regression coefficients. The number of components must be appropriately decided to construct a highly predictive model. R^2 and Q^2 values are used as the measure and defined as follows:

$$\begin{aligned} R^2 &= 1 - \frac{\sum (y_{\mathrm{obs}} - y_{\mathrm{calc}})^2}{\sum (y_{\mathrm{obs}} - \bar{y})^2} \\ Q^2 &= 1 - \frac{\sum (y_{\mathrm{obs}} - y_{\mathrm{pred}})^2}{\sum (y_{\mathrm{obs}} - \bar{y})^2} \end{aligned} \tag{3}$$

where y_{obs} is the actual **y** value, y_{calc} is the calculated **y** value, and y_{pred} is the predicted **y** value in the procedure of cross-validation such as leave-one-out. In this paper, the



ED1

Figure 1. Comparison between PCA and ICA.

(A) Explanatory variables; (B) PCA; (C) ICA.

number of components is determined by the first local maximum of Q^2.

ICA

ICA is a method for transforming observed multivariate data into statistically independent components expressed as the linear combinations of observed variables. Statistical independence is a different concept from decorrelation. Denote $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$ as random variables with a joint probability density function (pdf) of $p(\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n)$, and assume that these variables have zero-mean; then, they are said to be mutually statistically independent if the following condition holds:

$$p(\mathbf{s}_1, \mathbf{s}_2, ., \mathbf{s}_n) = \prod_{i=1}^{n} p_i(\mathbf{s}_i) \tag{4}$$

where $p_i(\mathbf{s}_i)$ (for $i = 1, 2, \ldots, n$) denotes a marginal pdf of $\mathbf{s}_i$, that of, the pdf of $\mathbf{s}_i$ when it is considered alone. The basic problem of ICA is then to estimate the original components $\mathbf{s}_i$ from measured variables $\mathbf{X}$. For performing ICA, $\mathbf{X}$ is first transformed into mutually uncorrelated variables. By defining the sphering matrix as $\mathbf{M} \in R^{d \times a}$, transformed matrix $\mathbf{Z} \in R^{n \times a}$ is given as

$$\mathbf{Z} = \mathbf{XM} \tag{5}$$

Generally, this pretreatment can be accomplished by singular value decomposition. Next, $\mathbf{Z}$ is transformed into mutually independent variables as follows:

$$\mathbf{S} = \mathbf{ZB} \tag{6}$$

where $\mathbf{S} \in R^{n \times a}$ is an independent component matrix and $\mathbf{B} \in R^{a \times a}$ is a transformation matrix. Several techniques that calculate $\mathbf{B}$ have been proposed. In this paper, FastICA[79] is used. This algorithm finds, one at a time, all non-Gaussian components, that is, independent components, regardless of their probability distributions. The relationship between $\mathbf{X}$ and $\mathbf{S}$ is given as

$$\begin{gathered} \mathbf{S} = \mathbf{XMB} = \mathbf{XW} \\ \mathbf{W} = \mathbf{MB} \end{gathered} \tag{7}$$

where $\mathbf{W} \in R^{d \times a}$ is a separating matrix. Mutually independent components underlying in $\mathbf{X}$ can be extracted by $\mathbf{W}$.

Figure 1 shows a comparison between PCA and ICA. Figure 1A is a plot of explanatory variables $\mathbf{x}_1$ and $\mathbf{x}_2$, and then PCA and ICA method are applied to this data set. A plot of principal components $\mathbf{t}_1$ and $\mathbf{t}_2$ extracted by PCA method is shown in Figure 1B and a plot of independent components $\mathbf{s}_1$ and $\mathbf{s}_2$ extracted by the ICA method is shown in Figure 1C.

Mutually uncorrelated components whose dispersion is large are extracted in order by the PCA method. On the one hand, the distribution of $\mathbf{s}_1$ and $\mathbf{s}_2$ is like that in Figure 1C. In other words, even if a value of one of these components is obtained, the information of another one is not obtained, because $\mathbf{s}_1$ and $\mathbf{s}_2$ are mutually independent.

Fault Detection and Classification Based on ICA

Monitoring methods based on PCA have been developed. PCA can extract mutually uncorrelated components from explanatory variables, through making full use of the second-order central moment. On the other hand, ICA can extract mutually independent components from explanatory variables, through making full use of the fourth-order central moment. The fourth-order central moment represents high values for components including outliers. ICA shows superior performance for detecting outliers over PCA, because independent components are sensitive to outliers.

However, even if traditional monitoring methods based on ICA can detect the abnormal situation, it is difficult to classify the reason of it appropriately. Then, we propose to select independent components, which are sensitive to the same outliers, and classify the reason of abnormal situation by using these components individually. By using the proposed method, fault detection model with high accuracy and classification of abnormal situation are achieved.

We construct ICA-based fault detection and classification model. The algorithm is described as follows:

1. Extract independent components $\mathbf{S}$ from measured variables $\mathbf{X}$ of training samples by using ICA method like as Eq. (7).
2. Check each independent component and separate $\mathbf{S}$ into $\mathbf{S}_1$, $\mathbf{S}_2$, ..., each of which is sensitive to the same outliers.

Also prepare $\mathbf{W}_i$ corresponding to $\mathbf{S}_i$ as follows:

$$\begin{aligned} \mathbf{S}_1 &= \mathbf{X}\mathbf{W}_1 \\ \mathbf{S}_2 &= \mathbf{X}\mathbf{W}_2 \\ &\cdots \\ \mathbf{S}_i &= \mathbf{X}\mathbf{W}_i \\ &\cdots \end{aligned} \tag{8}$$

3. Calculate I^2 statistic[73] which is defined for each $\mathbf{S}_i$ as follows:

$$\begin{aligned} I_{1,j}^2 &= \mathbf{s}_{1,j}\mathbf{s}_{1,j}^T \\ I_{2,j}^2 &= \mathbf{s}_{2,j}\mathbf{s}_{2,j}^T \\ &\cdots \\ I_{i,j}^2 &= \mathbf{s}_{i,j}\mathbf{s}_{i,j}^T \\ &\cdots \end{aligned} \tag{9}$$

where j is the number of sample.

4. Check each $I_{i,j}^2$ and set the threshold to distinguish the particular outliers of each I_i^2 from other samples.

After constructing the model, we detect the abnormal data and classify the reason of it online. The algorithm is described as follows:

1. Calculate $\mathbf{s}_{i,\text{test}}$ from the new test sample $\mathbf{x}_{\text{test}}$ and $\mathbf{W}_i$ as follows:

$$\begin{aligned} \mathbf{s}_{1,\text{test}} &= \mathbf{X}_{\text{test}}\mathbf{W}_1 \\ \mathbf{s}_{2,\text{test}} &= \mathbf{X}_{\text{test}}\mathbf{W}_2 \\ &\cdots \\ \mathbf{s}_{i,\text{test}} &= \mathbf{X}_{\text{test}}\mathbf{W}_i \\ &\cdots \end{aligned} \tag{10}$$

2. Calculate I_i^2 from $\mathbf{s}_{i,\text{test}}$ as follows:

$$\begin{aligned} I_{1,\text{test}}^2 &= \mathbf{s}_{1,\text{test}}\mathbf{s}_{1,\text{test}}^T \\ I_{2,\text{test}}^2 &= \mathbf{s}_{2,\text{test}}\mathbf{s}_{2,\text{test}}^T \\ &\cdots \\ I_{i,\text{test}}^2 &= \mathbf{s}_{i,\text{test}}\mathbf{s}_{i,\text{test}}^T \\ &\cdots \end{aligned} \tag{11}$$

3. Compare $I_{i,\text{test}}^2$ with the threshold and detect the abnormal data. If $I_{i,\text{test}}^2$ is above the threshold, $\mathbf{x}_{\text{test}}$ is the similar sample as $\mathbf{S}_i$ is sensitiveto.

4. Return to step 1.

Thus, we can judge whether $\mathbf{x}_{\text{test}}$ is a normal sample or the similar sample as $\mathbf{S}_i$ is sensitive to, one after another.

Soft Sensor Method Based on ICA and PLS

To solve the degradation of the soft sensor model, the regression model can be updated with a new sample. If the model is updated with an abnormal sample, the predictive ability can deteriorate. Therefore, the abnormal sample needs to be detected. However, by checking the abnormal data by only using the prediction error of $\mathbf{y}$, it is difficult to detect the abnormal samples and to classify the reason of them.

Then, we propose to judge whether a new sample is normal or abnormal by using the ICA-based fault detection and classification model and the prediction error of $\mathbf{y}$. If the sample is detected as abnormal by the ICA model, the reason of it can be trouble of process. If the sample is detected as abnormal by the prediction error of $\mathbf{y}$, the reason of it can be trouble of analyzer. If the sample is detected as normal, we can update the soft sensor model based on PLS. The algorithm is described as follows:

1. Construct an ICA-based fault detection and classification model with training samples $\mathbf{X}_{\text{train-ICA}}$, and a PLS model with first training samples $\mathbf{y}_{\text{train}}$ and $\mathbf{X}_{\text{train-PLS}}$ as follows:

$$\mathbf{Y}_{\text{train}} = \mathbf{X}_{\text{train-PLS}}\mathbf{b} + \mathbf{const} \tag{12}$$

The number of samples of $\mathbf{X}_{\text{train-ICA}}$ and that of $\mathbf{X}_{\text{train-PLS}}$ are not necessarily the same, and the type of measured variables of $\mathbf{X}_{\text{train-ICA}}$ and that of $\mathbf{X}_{\text{train-PLS}}$ are not necessarily the same either.

2. Judge whether the new test data $\mathbf{x}_{\text{test-ICA}}$ is normal or abnormal by using the ICA-based fault detection and classification model explained in the previous section. If the sample is detected as abnormal here, the reason of it can be trouble of process.

3. Predict an objective variable from the new test data $\mathbf{x}_{\text{test-PLS}}$ by using the PLS model as follows:

$$y_{\text{pred}} = \mathbf{x}_{\text{test-PLS}}\mathbf{b} + \text{const} \tag{13}$$

4. Judge whether the objective variable is normal or abnormal by comparing measured value y_{test} with predicted value y_{pred} as follows:

$$\text{prediction error} = y_{\text{test}} - y_{\text{pred}} \tag{14}$$

If the sample is detected as abnormal here, the reason of it can be trouble of analyzer.

5. Determine whether the PLS model is updated or not by using Table 1.

6. If the PLS model is updated, prepare the new training data $\mathbf{y}_{\text{new}}$ and $\mathbf{X}_{\text{new}}$, that is, add the new data y_{test} and $\mathbf{x}_{\text{test-PLS}}$ to the training data, omit the oldest data from the training data.

7. Construct the new PLS model with $\mathbf{y}_{\text{new}}$ and $\mathbf{X}_{\text{new}}$ as follows:

$$\mathbf{y}_{\text{new}} = \mathbf{X}_{\text{new}}\mathbf{b} + \mathbf{const} \tag{15}$$

8. Return to step 2.

Results and Discussion

We applied this method to real industrial data to verify prediction ability and fault detection and classification ability of the proposed method. We analyzed the data obtained from the operation of the distillation column at Mizushima works, Mitsubishi Chemical. Figure 2 shows a schematic representation of the distillation column and Table 2 shows the process variables. An objective variable y is concentration of bottom product, which is lower boiling point, and explanatory variables $\mathbf{X}$ are 19 variables, which are temperature, pressure, liquid level, and so on. Inputs variables are F3 and F4, and operational variables are F1 and F2. Sampling interval is 1 h and we prepared 10,000 samples (10,000 h). Figure 3 shows a plot of $\mathbf{y}$. Variations appeared in around 3600, 6800–7000, 8100, and 9000–9500 h are caused by disturbance, check of

Table 1. Judgment About Whether the PLS Model Is Updated or Not

		ICA Model (Explanatory Variables)	
		Normal	Abnormal
PLS model (predictive error of objective variable)	Normal	Normal Update the PLS model	Process is abnormal Do not update the PLS model
	Abnormal	Analyzer is abnormal Do not update the PLS model	Process is abnormal Do not update the PLS model

plant, trouble of a concentration analyzer, and plant test, respectively. The variations caused by the disturbance and the plant test do not need to be detected, because they are not abnormal variations. The variation caused by the trouble of the concentration analyzer needs to be detected as faults of **y** and one which is caused by the check of plant needs to be detected as faults of **X**. First 4000 samples are training data and we tried to predict next 6000 samples.

To verify the superiority of the proposed method, below four methods are applied to real industrial data.

A. Do not update the PLS model, and detect the abnormal data by using the prediction error of **y**.

B. Update the PLS model, and detect the abnormal data by using the prediction error of **y**.

C. Update the PLS model, and detect and classify the abnormal data by using the PCA-based fault detection model and the prediction error of **y**.

D. Update the PLS model, and detect and classify the abnormal data by using the ICA-based fault detection and classification model and the prediction error of **y**.

Method A is a traditional soft sensor method, and method D is the proposed one. The three-sigma method was used to detect and classify the abnormal data of **y**. The explanatory variables used to construct the ICA-based fault detection and classification model are 19 variables in Table 2 and those which are used to construct the PLS model are 25 variables which are 12 variables (No. 2, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19) in Table 2, and the 12 variables and **y** whose time lag is 1 h. The method that constructs PLS model with dynamics of process variables is called Dynamic PLS.[2,6]

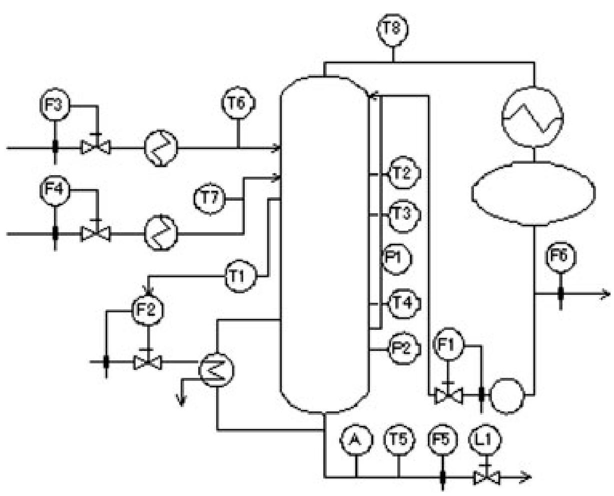


Figure 2. A schematic representation of the distillation column.

PLS-based fault detection method, which distinguishes soft sensor errors from analyzer troubles, has been proposed.[6] By using PLS, not PCA, as a dimension reducing method, the model can contain the information of **y**. However, it is conceivable that **X**-variables used to predict the value of **y** with high predictive accuracy are different from those used to detect the fault with high accuracy. We did not use this method, because we regarded the fault detection differently from the prediction of **y**.

Method A

We constructed the PLS model with first 4000 samples. The number of component was 12, R^2 value was 0.874, and Q^2 value was 0.827. By using the PLS method, a rather accurate predictive model could be constructed.

Table 3 shows the result of prediction and fault detection of 6000 samples. The ${R_{\mathrm{pred}}}^2$ value is an R^2 value that is calculated with the test samples, and RMSE (Root Mean Square Error) is defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum (y_{\mathrm{obs}} - y_{\mathrm{pred}})^2}{n}} \tag{16}$$

where y_{obs} is the actual **y** value, and y_{pred} is the predicted **y** value, and n is the number of test samples. ${R_{\mathrm{pred}}}^2$ and RMSE are calculated with the samples from which samples detected as abnormal are omitted. The number of samples detected as

Table 2. Process Variables

	Symbol	Objective Variable
	A	Bottom product concentration
No.	Symbol	Explanatory variables
1	F1	Reflux flow
2	F2	Reboiler flow
3	F3	Feed 1 flow
4	F4	Feed 2 flow
5	F5	Bottom flow
6	F6	Top flow
7	L1	Liquid level
8	P1	Pressure 1
9	P2	Pressure 2
10	T1	Temperature 1
11	T2	Temperature 2
12	T3	Temperature 3
13	T4	Temperature 4
14	T5	Bottom temperature
15	T6	Feed 2 temperature
16	T7	Feed 1 temperature
17	T8	Top temperature
18	F1/F6	Feed flow ratio
19	F4/F3	Reflux ratio

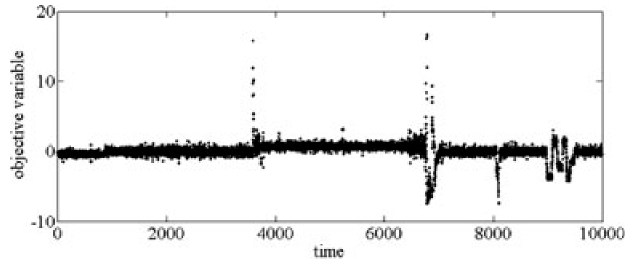


Figure 3. Objective variable.

abnormal, R_{pred}^2 value, and RMSE were 258, 0.829, and 1.18, respectively. It is conceivable that the degradation of soft sensor model happened because RMSE was large. Then, R_{pred}^2 value was rather large. It could be because correlation between y_{obs} and y_{pred} was large because of high standard deviation of **y**. Thus, we used RMSE to compare the prediction ability of each method rather than R_{pred}^2 value.

Figure 4 shows a plot of measured and predicted **y** in method A. Figure 4A is the plot before omitting the samples detected as abnormal. Red circles represent them. Figure 4B is the plot after omitting them. The plot shows that the samples, which deviate from the diagonal, are only detected as abnormal. Figure 5 shows the result of fault detection in method A. Red circles represents the samples detected as abnormal. Some samples around 6800–7000 and 8100 h could be detected as abnormal, and others could not. For example, samples around the end of the check of plant were not detected as abnormal. Furthermore, it is difficult to distinguish between the trouble of the concentration analyzer and the check of plant, because all of the abnormal samples were detected as faults of **y**.

Method B

We predicted 6000 samples by using method B. The abnormal samples are detected by using the prediction error of **y**. If the sample is normal, the PLS model is updated. Table 3 shows the result of this calculation. The number of samples detected as abnormal, R_{pred}^2 value, and RMSE were 53, 0.908, and 0.349, respectively. RMSE of method B was lower than that of method A, because the predictive ability increased by updating the PLS model. It is conceivable that the degradation of soft sensor model was reduced. However, it could occur overfitting because the number of samples detected as abnormal was too small.

Figure 6 shows a plot of measured and predicted **y** in method B. Figure 6A is the plot before omitting the samples detected as abnormal. Red circles represent them. Figure 6B is the plot after omitting them. The plot shows a much tighter cluster of predicted values along the diagonal in method B than in method A. However, the number of samples detected as abnormal in method B is smaller than in method A. Figure 7 shows the result of fault detection in method B. Red circles represents the samples detected as

Table 3. Number of Samples, R_{pred}^2, RMSE in Each Method

Method	A	B	C	D
No. of abnormal samples	258	53	61	307
R_{pred}^2	0.829	0.908	0.907	0.910
RMSE	1.18	0.349	0.387	0.281

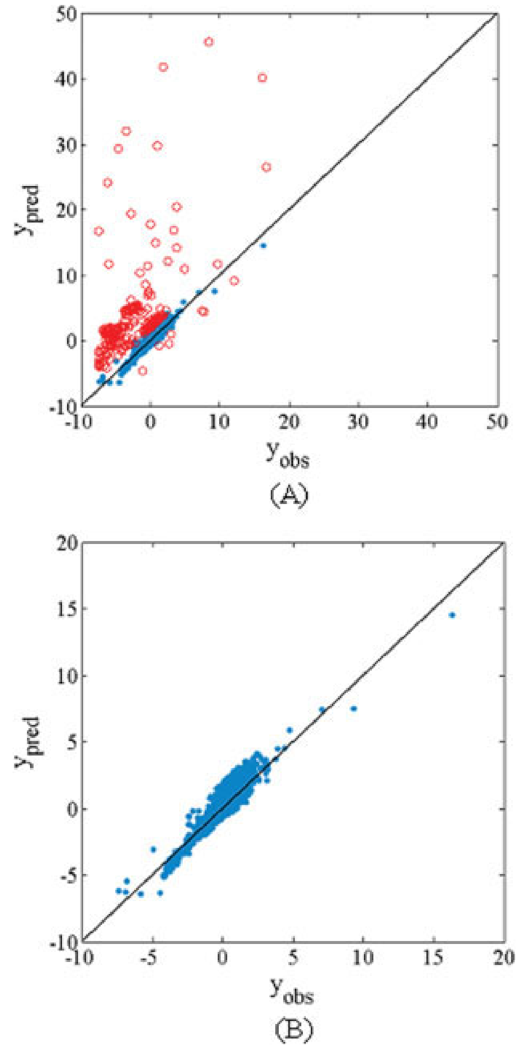


Figure 4. The relationship between measured and predicted y in method A.

(A) Before omitting the samples diagnosed as abnormal (red circles). (B) After omitting the samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

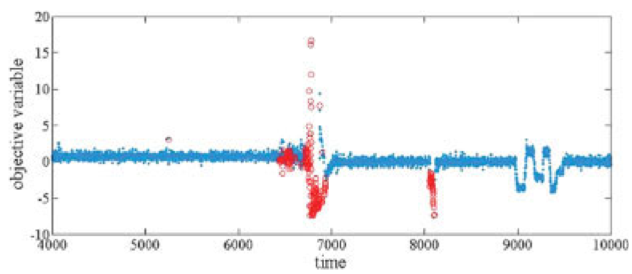


Figure 5. Fault detection in method A.

Red circles indicate samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

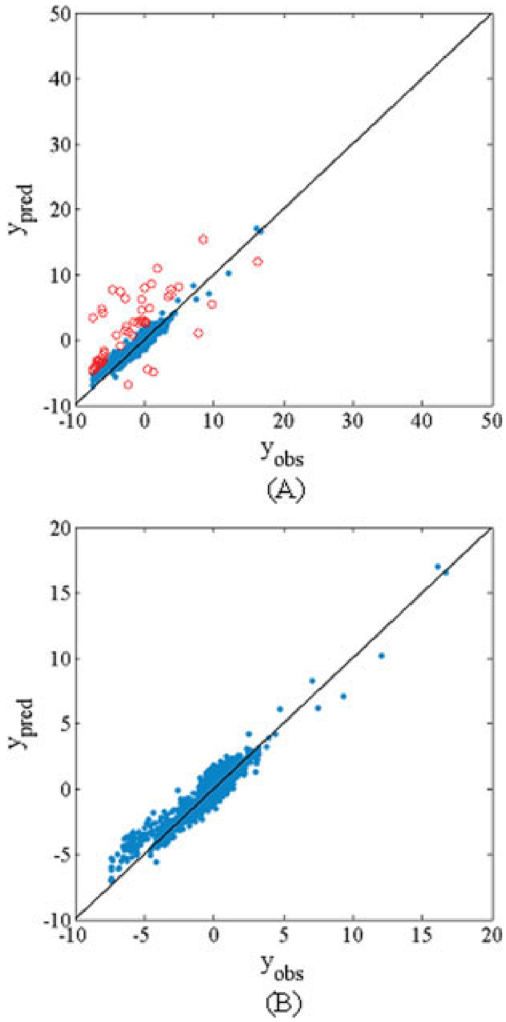


Figure 6. The relationship between measured and predicted *y* in method B.

(A) Before omitting the samples diagnosed as abnormal (red circles). (B) After omitting the samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

abnormal. Method B could not detect many samples around 6800–7000 h as abnormal, and could not almost samples around 8100 h as abnormal.

Method C

To compare PCA-based fault detection method with the proposed method, method D, the PCA-based fault detection model was constructed. Figure 8 shows the Hotelling T^2 statistic and the sum of squared residuals Q statistic.[82] The number of components is 5 where the cumulative contribution ratio is higher than 0.9. In addition, the number of components did not affect the results significantly. The threshold of T^2 statistic was set in 310 and that of Q statistic was set in 150, because a variation around 3600 is not abnormal. If T^2 or Q statistic were above the thresholds, test samples were detected as abnormal.

Then, we predicted 6000 samples by using method C. The abnormal samples were detected by using the prediction error of **y** and the PCA-based fault detection model. If the sample was normal, the PLS model was updated. Table 3 shows the result of this calculation. The number of samples detected as abnormal, ${R_{pred}}^2$ value, and RMSE were 61, 0.907, and 0.387, respectively. Although ${R_{pred}}^2$ value of method B and that of method C are almost the same, RMSE of method C was higher than that of method B. The PCA-based fault detection model did not work well. In addition, it could occur overfitting as in method B, because the number of samples detected as abnormal was too small.

Figure 9 shows a plot of measured and predicted **y** in method C. Figure 9A is the plot before omitting the samples detected as abnormal. Red circles represent them. Figure 9B is the plot after omitting them. The plot shows a much tighter cluster of predicted values along the diagonal in method C than in method A, but many differences between method B and C are not seen. By using the PCA-based fault detection model, not only samples which deviate from the diagonal but also samples which are near the diagonal were detected as abnormal. However, the number of them was too small. Figure 10 shows the result of fault detection and classification in method C. Red circles represent the samples detected as abnormal by the prediction error of **y**, and black asterisks represent the samples detected as abnormal by the PCA-based fault detection model. As in method B, there were many samples which did not detected as abnormal when check of plant or trouble of a concentration analyzer occurred. It was conceivable that Q statistic did not detect an unknown variation because it included a variation around 3600.

Method D

First, we constructed ICA-based fault detection and classification model. Figure 11 shows independent components extracted from the explanatory variables by using ICA. We selected the components $\mathbf{S}_1(\mathbf{s}_1, \mathbf{s}_3, \mathbf{s}_6, \mathbf{s}_{11}, \mathbf{s}_{15})$ which are sensitive to variation of around 3600 h and the components $\mathbf{S}_2(\mathbf{s}_2, \mathbf{s}_{12}, \mathbf{s}_{14}, \mathbf{s}_{17})$ which are not sensitive to variation of around 3600 h. $\mathbf{s}_7$, $\mathbf{s}_{10}$, and $\mathbf{s}_{16}$ are not included in $\mathbf{S}_1$ because samples which deviate from the distribution in $\mathbf{s}_7$, $\mathbf{s}_{10}$, and $\mathbf{s}_{16}$ are the same as those which deviate from the distribution in $\mathbf{s}_1$. $\mathbf{s}_4$ and $\mathbf{s}_5$ are not included in $\mathbf{S}_1$ and $\mathbf{S}_2$ because they have a special variation around 3600 h, but samples around 3600 h do not deviate from the distribution in $\mathbf{s}_4$ and $\mathbf{s}_5$. $\mathbf{s}_9$ and $\mathbf{s}_{18}$ are not included in $\mathbf{S}_1$ and $\mathbf{S}_2$ because they have a

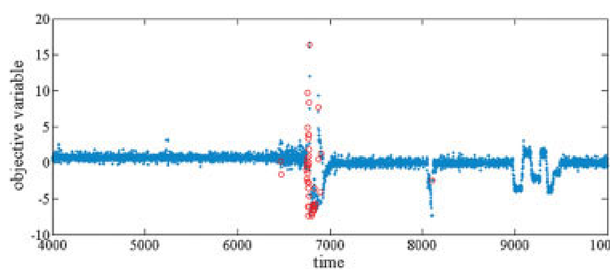


Figure 7. Fault detection in method B.

Red circles indicate samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

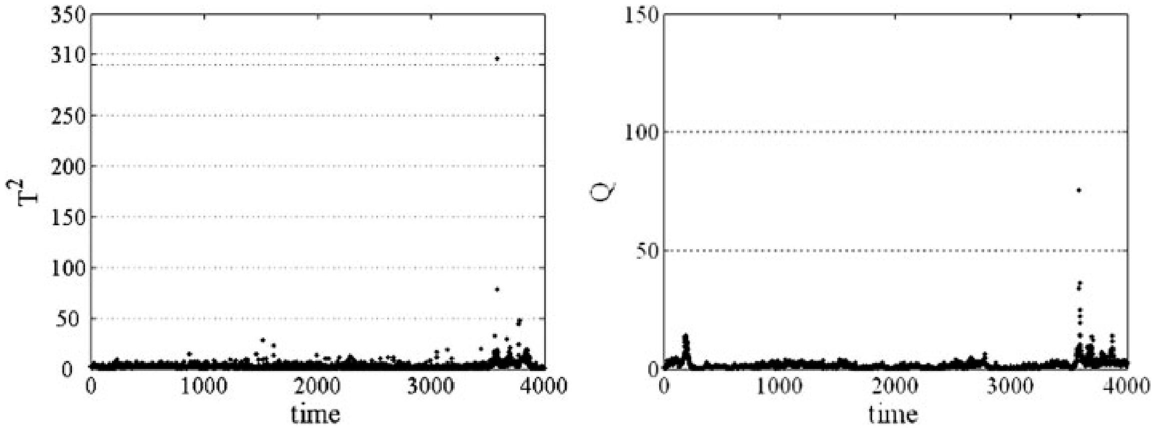


Figure 8. T^2 and Q.

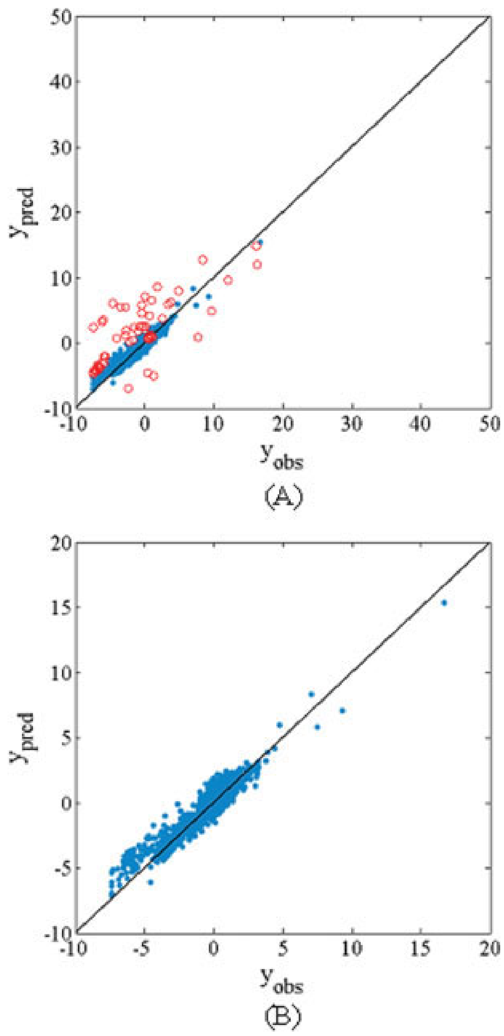


Figure 9. The relationship between measured and predicted *y* in method C.

(A) Before omitting the samples diagnosed as abnormal (red circles). (B) After omitting the samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

special variation in early hours. $\mathbf{s}_{13}$ is not included in $\mathbf{S}_2$ because it is almost the same as the liquid level.

Figure 12 shows I^2 statistic calculated from $\mathbf{S}_1$ and $\mathbf{S}_2$. The threshold of I_1^2 was set in 25 to detect the same variation as that of around 3600 h, and the threshold of I_2^2 was set in 70 to detect an unknown variation. By using I_1^2 and I_2^2, a variation around 3600 can be separated. Table 4 shows a judgment about whether the situation of plant is normal or abnormal. If I_1^2 and I_2^2 are under threshold, the situation of plant can be normal. If I_1^2 is above threshold and I_2^2 is under threshold, the situation of plant can be normal, because it is conceivable that it is the similar variation as that of around 3600 h. If I_1^2 is under threshold and I_2^2 is above threshold, the situation of plant can be abnormal, because it is conceivable that it is an unknown variation. If I_1^2 and I_2^2 are above threshold, the situation of plant can be abnormal because it is conceivable that it is the variation of whole plant.

Then, we predicted 6000 samples by using method D. The abnormal samples were detected and classified by using the prediction error of $\mathbf{y}$ and the ICA-based fault detection and classification model. If the sample was normal, the PLS model was updated. Table 3 shows the result of this calculation. The number of samples detected as abnormal, R_{pred}^2 value, and RMSE were 307, 0.910, and 0.281, respectively. It is conceivable that the degradation of soft sensor model is

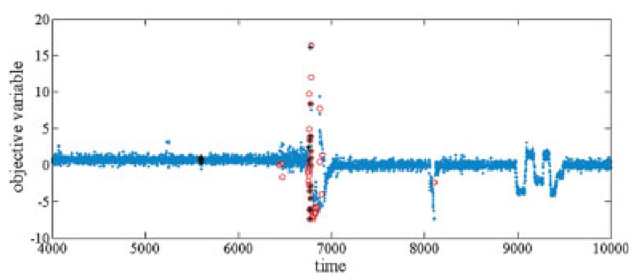


Figure 10. Fault detection in method C.

Samples diagnosed as abnormal by the prediction error of y (red circles). Samples diagnosed as abnormal by the PCA-based fault detection model (black asterisks). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

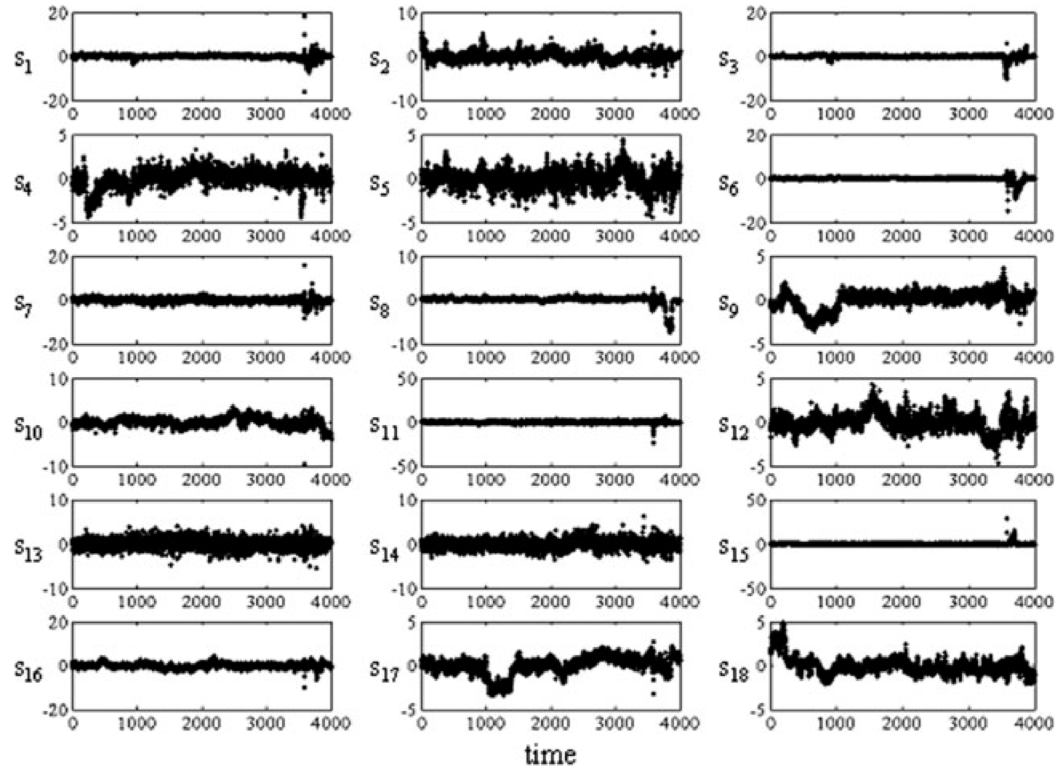


Figure 11. Independent components.

solved because of the low RMSE. Furthermore, RMSE of method D was lower than that of method B and C, because the predictive ability increased by detecting and classifying the abnormal samples by using the ICA-based fault detection and classification model. We verified that the model with high predictive accuracy could be constructed by using the proposed method.

Figure 13 shows a plot of measured and predicted **y** in method D. Figure 13A is the plot before omitting the samples detected as abnormal. Red circles represent them. Figure 13B is the plot after omitting them. The plot shows a much tighter cluster of predicted values along the diagonal in method D, reflecting the higher prediction of **y**. By using the ICA-based fault detection and classification model, not only samples which deviate from the diagonal but also those which are near the diagonal were detected as abnormal. In fact, there were abnormal samples near the diagonal. For example, in Figure 6B and 9B, the samples whose **y** value is under -5 are not detected as abnormal. However, these samples need to be detected so, because the trouble of the concentration analyzer and the check of plant happen there from Figure 3. Most part of samples whose **y** value is under -5 is detected as abnormal by method D as shown in Figure 13.

Figure 14 shows the result of fault detection and classification in method D. Red circles represent the samples detected as abnormal by the prediction error of **y**, and black asterisks represent the samples detected as abnormal by the ICA-based fault detection and classification model. The samples around

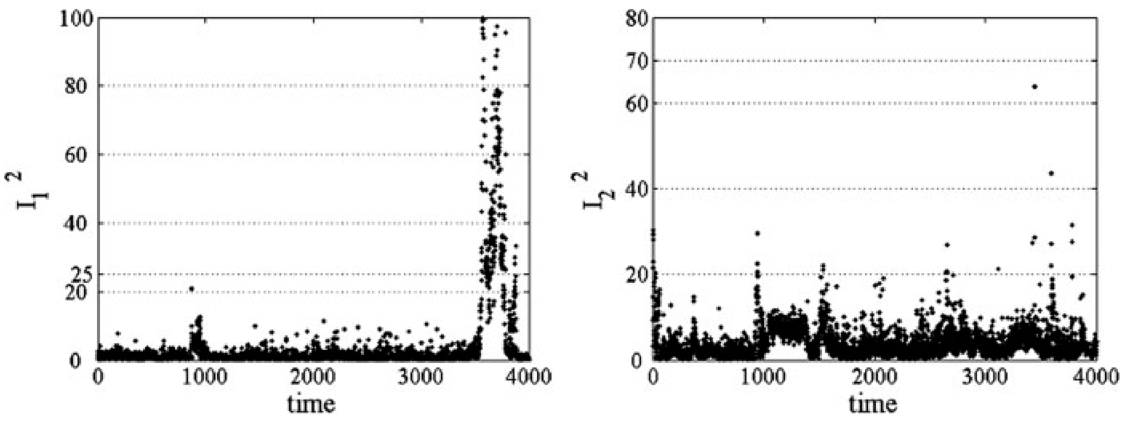


Figure 12. I_1^2 and I_2^2.

Table 4. Judgment About Whether the Situation of Plant Is Normal or Abnormal

		I_1^2	
		Under Threshold	Above Threshold
I_2^2	Under threshold	Normal	Same as around 3600 h Normal
	Above threshold	Unknown variation	Variation of whole plant
		Abnormal	Abnormal

6800–7000 could be detected as abnormal by the ICA model. Then, the result of them could be classified as fault of **X**. The samples around 8100 could be detected as abnormal by the prediction error of **y**, although method B and C could not detect the samples. Then, the result of them could be classi-

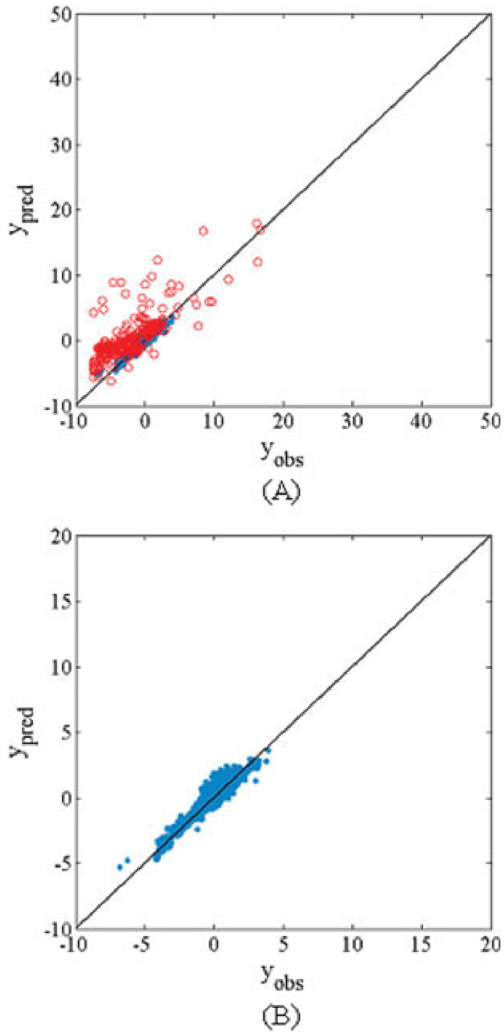


Figure 13. The relationship between measured and predicted *y* in method D.

(A) Before omitting the samples diagnosed as abnormal (red circles). (B) After omitting the samples diagnosed as abnormal. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

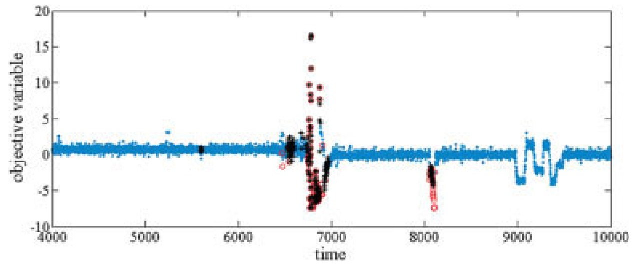


Figure 14. Fault detection in method D.

Samples diagnosed as abnormal by the prediction error of *y* (red circles). Samples diagnosed as abnormal by ICA-based fault detection and classification model (black asterisks). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

fied as fault of **y**, that is, the concentration analyzer. The samples around 5600 could be also detected as abnormal by the ICA model. It is nonsteady state caused by the variation of liquid level. The nonsteady state samples unconcerned with **y** can decrease the predictive ability of the PLS model. Then, by using the ICA model, the nonsteady state could be detected as abnormal and the degradation of soft sensor model could be prevented. It was shown that we could comprehend the state of plant by the ICA model and estimate **y** by the PLS model, updating it appropriately.

Conclusion

To increase prediction ability and fault detection ability, we have developed the new soft sensor method that combines ICA and PLS. By using the ICA method, independent components were extracted from explanatory variables and we selected independent components which were sensitive to the same outliers because independent components were mutually independent and sensitive to outliers. Then, we classified the reason of abnormal situation by using these components. By using the ICA-based fault detection and classification model, the objective variable could be predicted, updating the PLS model with only normal samples.

We analyzed real industrial data as the application of the proposed method and confirmed superiority of the proposed method over traditional methods. It was shown that we could comprehend the state of plant by the ICA model and estimate **y** by the PLS model, updating it appropriately. Furthermore, by using the ICA-based fault detection and classification model, the nonsteady state, which could not be detected as abnormal by the prediction error of **y**, could be detected.

The selection of independent components has arbitrary property. Someone who constructs the ICA-based fault detection and classification model must select independent components, manually. Therefore, it is desired that independent components be selected automatically by using some sort of measure. The fourth-order central moment and validation data might be used as the measure. In this paper, only the PLS model is updated online. However, by updating also the ICA model appropriately, fault detection ability will rise and the model will be in no need of maintenance.

To prevent occurring overfitting, we used a linear regression method, PLS. Many nonlinear regression methods are reported to construct soft sensor models. A linear model could not represent all the relationships among process variables. Therefore, prediction ability will rise by taking well into account nonlinearity, for example, using nonlinear regression methods, constructing multiple models, and so on.

The proposed method could be applied to not only the analyses of the distillation column but also the other time series analyses. By applying this method to process control, plants could be operated stably.

Acknowledgments

The authors acknowledge the support of Mizushima works, Mitsubishi Chemical.

Literature Cited

1. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst*. 2001;58:109–130.
2. Kano M, Miyazaki K, Hasebe S, Hashimoto I. Inferential control system of distillation compositions using dynamic partial least squares regression. *J Process Control*. 2000;10:157–166.
3. Kano M, Showchaiya N, Hasebe S, Hashimoto I. Inferential control of distillation compositions: selection of model and control configuration. *Control Eng Practice*. 2003;11:927–933.
4. Gonzalez GD, Orchard M, Cerda JL, Casali A, Vallebuona G. Local models for soft-sensors in a rougher flotation bank. *Miner Eng*. 2003;16:441–453.
5. Hagedorn A, Legge RL, Budman H. Evaluation of spectrofluorometry as a tool for estimation in fed-batch fermentations. *Biotechnol Bioeng*. 2003;83:104–111.
6. Kamohara H, Takinami A, Takeda M, Kano M, Hasebe S, Hashimoto I. Product quality estimation and operating condition monitoring for industrial ethylene fractionator. *J Chem Eng Jpn*. 2004;37:422–428.
7. Zamprogna E, Barolo M, Seborg DE. Estimating product composition profiles in batch distillation via partial least squares regression. *Control Eng Practice*. 2004;12:917–929.
8. Chu YH, Lee YH, Han CH. Improved quality estimation and knowledge extraction in a batch process by bootstrapping-based generalized variable selection. *Ind Eng Chem Res*. 2004;43:2680–2690.
9. Lennox B. Integrating fault detection and isolation with model predictive control. *Int J Adapt Control Signal Process*. 2005;19:199–212.
10. Zamprogna E, Barolo M, Seborg DE. Optimal selection of soft sensor inputs for batch distillation columns using principal component analysis. *J Process Control*. 2005;15:39–52.
11. Kim MJ, Lee YH, Han IS, Han CH. Clustering-based hybrid soft sensor for an industrial polypropylene process with grade changeover operation. *Ind Eng Chem Res*. 2005;44:334–342.
12. Ookita K. Operation and quality control for chemical plants by soft-sensors. *CICSJ Bull*. 2006;24:31–33 (in Japanese).
13. Sharmin R, Sundararaj U, Shah S, Griend LV, Sun YJ. Inferential sensors for estimation of polymer quality parameters: Industrial application of a PLS-based soft sensor for a LDPE plant. *Chem Eng Sci*. 2006;61:6372–6384.
14. Aguado D, Ferrer A, Seco A, Ferrer J. Comparison of different predictive models for nutrient estimation in a sequencing batch reactor for wastewater treatment. *Chemom Intell Lab Syst*. 2006;84:75–81.
15. Bruwer MJ, MacGregor JF, Bourg WM. Soft sensor for snack food textural properties using on-line vibrational measurements. *Ind Eng Chem Res*. 2007;46:864–870.
16. Lin B, Recke B, Knudsen JKH, Jorgensen SB. A systematic approach for soft sensor development. *Comput Chem Eng*. 2007;31:419–425.
17. Dodds SA, Heath WP. Construction of an online reduced-spectrum NIR calibration model from full-spectrum data. *Chemom Intell Lab Syst*. 2005;76:37–43.
18. Qin SJ, Yue HY, Dunia R. Self-validating inferential sensors with application to air emission monitoring. *Ind Eng Chem Res*. 1997;36:1675–1685.
19. Baffi G, Martin EB, Morris AJ. Non-linear projection to latent structures revisited: the quadratic PLS algorithm. *Comput Chem Eng*. 1999;23:395–411.
20. Baffi G, Martin EB, Morris AJ. Non-linear projection to latent structures revisited (the neural network PLS algorithm). *Comput Chem Eng*. 1999;23:1293–1307.
21. Park S, Han CH. A nonlinear soft sensor based on multivariate smoothing procedure for quality estimation in distillation columns. *Comput Chem Eng*. 2000;24:871–877.
22. Wilson DJH, Irwin GW. PLS modelling and fault detection on the Tennessee Eastman benchmark. *Int J Syst Sci*. 2000;31:1449–1457.
23. Zhao SJ, Zhang J, Xu YM, Xiong ZH. Nonlinear projection to latent structures method and its applications. *Ind Eng Chem Res*. 2006;453:843–3852.
24. Dacosta P, Kordich C, Williams D, Gomm JB. Estimation of inaccessible fermentation states with variable inoculum sizes. *Artif Intell Eng*. 1997;11:383–392.
25. Su HB, Fan LT, Schlup JR. Monitoring the process of curing of epoxy/graphite fiber composites with a recurrent neural network as a soft sensor. *Eng Appl Artif Intell*. 1998;11:293–306.
26. Wang Y, Rong G. A self-organizing neural-network-based fuzzy system. *Fuzzy Sets Syst*. 1999;103:1–11.
27. Yang SH, Chen BH, Wang XZ. Neural network based fault diagnosis using unmeasurable inputs. *Eng Appl Artif Intell*. 2000;13:345–356.
28. de Assis AJ, Maciel R. Soft sensors development for on-line bioreactor state estimation. *Comput Chem Eng*. 2000;24:1099–1103.
29. Radhakrishnan VR, Mohamed AR. Neural networks for the identification and control of blast furnace hot metal quality. *J Process Control*. 2000;10:509–524.
30. Meleiro LAC, Maciel R. A self-tuning adaptive control applied to an industrial large scale ethanol production. *Comput Chem Eng*. 2000;24:925–930.
31. Casali A, Gonzalez G, Vallebuona G, Perez C, Vargas R. Grindability soft-sensors based on lithological composition and on-line measurements. *Miner Eng*. 2001;14:689–700.
32. Fortuna L, Rizzo A, Sinatra M, Xibilia MG. Soft analyzers for a sulfur recovery unit. *Control Eng Practice*. 2003;11:1491–1500.
33. Chen LZ, Nguang SK, Li XM, Chen XD. Soft sensors for on-line biomass measurements. *Bioprocess Biosyst Eng*. 2004;26:191–195.
34. Feil B, Abonyi J, Pach P, Nemeth S, Arva P, Nemeth M, Nagy G. Semi-mechanistic models for state-estimation - Soft sensor for polymer melt index prediction. *Lect Note Artif Intell*. 2004;3070:1111–1117.
35. Dufour P, Bhartiya S, Dhurjati PS, Doyle FJ. Neural network-based software sensor: Training set design and application to a continuous pulp digester. *Control Eng Practice*. 2005;13:135–143.
36. Fortuna L, Graziani S, Xibilia MG. Soft sensors for product quality monitoring in debutanizer distillation columns. *Control Eng Practice*. 2005;13:499–508.
37. Chen GC, Yu JS. Particle swarm optimization neural network and its application in soft-sensing modeling. *Lect Notes Comput Sci*. 2005;3611:610–617.
38. Yang H, Yang CY, Song CH, Chai TY. Intelligent optimal control in rare-earth countercurrent extraction process via soft-sensor. *Lect Notes Comput Sci*. 2005;3611:214–223.
39. Luo JX, Shao HH. Developing soft sensors using hybrid soft computing methodology: a neurofuzzy system based on rough set theory and genetic algorithms. *Soft Comput*. 2006;10:54–60.
40. Negro C, Alonso L, Blanco K, Tijero J. Optimization of the fiber cement composite process. *Ind Eng Chem Res*. 2006;45:197–205.
41. Dam M, Saraf DN. Design of neural networks using genetic algorithm for on-line property estimation of crude fractionator products. *Comput Chem Eng*. 2006;30:722–729.
42. Karakuzu C, Turker M, Ozturk S. Modelling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation. *Control Eng Practice*. 2006;14:959–974.
43. Dai XZ, Wang WC, Ding YH, Sun ZY. "Assumed inherent sensor" inversion based ANN dynamic soft-sensing method and its application in erythromycin fermentation process. *Comput Chem Eng*. 2006;30:1203–1225.
44. Geng H, Xiong ZH, Mao SA, Xu YM. Dynamic soft-sensing model by combining diagonal recurrent neural network with Levinson predictor. *Lect Notes Comput Sci*. 2006;3973:1059–1064.
45. Liu XG. Product quality prediction with support vector machines. *Lect Notes Comput Sci*. 2006;3973:1126–1131.

46. Zhou Q, Xiong ZH, Zhang J, Xu YM. Hierarchical neural network based product quality prediction of industrial ethylene pyrolysis process. *Lect Notes Comput Sci*. 2006;3973:1132–1137.
47. Xu YF, Chen GC, Yu JS. The Kalman particle swarm optimization algorithm and its application in soft-sensor of acrylonitrile yield. *Lect Notes Comput Sci*. 2006;4222:176–179.
48. Shi J, Liu XG, Sun YX. Melt index prediction by neural networks based on independent component analysis and multi-scale analysis. *Neurocomputing*. 2006;70:280–287.
49. Fortuna L, Giannone P, Graziani S, Xibilia MG. Virtual instruments based on stacked neural networks to improve product quality monitoring in a refinery. *IEEE Trans Instr Meas*. 2007;56:95–101.
50. Abdel-Aal RE. Predictive modeling of mercury speciation in combustion flue gases using GMDH-based abductive networks. *Fuel Process Technol*. 2007;88:483–491.
51. Minari RJ, Stegmayer GS, Gugliotta LM, Chiotti OA, Vega JR. Industrial SBR process: Computer simulation study for online estimation of steady-state variables using neural networks. *Macromol React Eng*. 2007;1:405–412.
52. Vapnik VN. *The Nature of Statistical Learning Theory*. New York: Springer; 1999.
53. Yan WW, Shao HH, Wang XF. Soft sensing modeling based on support vector machine and Bayesian model selection. *Comput Chem Eng*. 2004;28:1489–1498.
54. Lee DE, Song JH, Song SO, Yoon ES. Weighted support vector machine for quality estimation in the polymerization process. *Ind Eng Chem Res*. 2005;44:2101–2105.
55. Chen AJ, Song ZH, Li P. Soft Sensor Modeling Based on DICA-SVR. *Lect Notes Comput Sci*. 2005;3644:868–877.
56. Desai K, Badhe Y, Tambe SS, Kulkarni BD. Soft-sensor development for fed-batch bioreactors using support vector regression. *Biochem Eng J*. 2006;27:225–239.
57. Shi J, Liu XG. Melt index prediction by weighted least squares support vector machines. *J Appl Polym Sci*. 2006;101:285–289.
58. Yuan P, Mao ZZ, Wang FL. Endpoint prediction of EAF based on multiple support vector machines. *J Iron Steel Res*. 2007;14:20–24.
59. Dayal BS, MacGregor JF. Recursive exponentially weighted PLS and its applications to adaptive control and prediction. *J Process Control*. 1997;7:169–179.
60. Qin SJ. Recursive PLS algorithms for adaptive data modeling. *Comput Chem Eng*. 1998;22:503–514.
61. Mu SJ, Zeng YZ, Liu RL, Wu P, Su HY, Chu J. Online dual updating with recursive PLS model and its application in predicting crystal size of purified terephthalic acid (PTA) process. *J Process Control*. 2006;16:557–566.
62. Fu YF, Su HY, Chu JA. MIMO soft-sensor model of nutrient content for compound fertilizer based on hybrid modeling technique. *Chin J Chem Eng*. 2007;15:554–559.
63. Liu JL. On-line soft sensor for polyethylene process with multiple production grades. *Control Eng Practice*. 2007;15:769–778.
64. Comon P. Independent component analysis, A new concept?. *Signal Processing*. 1994;36:287–314.
65. Chen J, Wang XZ. A New Approach to Near-Infrared Spectral Data Analysis Using Independent Component Analysis. *J Chem Inf Comput Sci*. 2001;41:992–1001.
66. Shao X, Wang W, Hou Z, Cai W. A new regression method based on independent component analysis. *Talanta*. 2006;69:676–680.
67. Kaneko H, Arakawa M, Funatsu K. Development of a New Regression Analysis Method Using Independent Component Analysis. *J Chem Inf Model*. 2008;48:534–541.
68. Kermit M, Tomic O. Independent Component Analysis Applied on Gas Sensor Array Measurement Data. *IEEE Sensors J*. 2003;3:218–228.
69. Kano M, Tanaka S, Hasebe S, Hashimoto I. Monitoring Independent Components for Fault Detection. *AIChE J*. 2003;49:294–298.
70. Carpentier AS, Riva A, Tisseur P, Didier G, Henaut A. The operons, a criterion to compare the reliability of transcriptome analysis tools: ICA is more reliable than ANOVA, PLS and PCA. *Comput Biol Chem*. 2004;28:3–10.
71. Kano M, Hasebe S, Hashimoto I, Ohno H. Evolution of multivariate statistical process control: application of independent component analysis and external analysis. *Comput Chem Eng*. 2004;28:1157–1166.
72. Lee JM, Yoo C, Lee IB. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chem Eng Sci*. 2004;59:2995–3006.
73. Lee JM, Qin SJ, Lee IB. Fault detection and diagnosis based on modified independent component analysis. *AIChE J*. 2006;52:3501–3514.
74. Albazzaz H, Wang XZ. Historical data analysis based on plots of independent and parallel coordinates and statistical control limits. *J Process Control*. 2006;16:103–114.
75. Liu F, Wu CY. Improvement on multivariate statistical process monitoring using multi-scale ICA. *Lect Notes Comput Sci*. 2006;3889:376–383.
76. Zhang YW, Qin SJ. Fault detection of nonlinear processes using multiway kernel independent component analysis. *Ind Eng Chem Res*. 2007;46:7780–7787.
77. Albazzaz H, Wang XZ. Introduction of dynamics to an approach for batch process monitoring using independent component analysis. *Chem Eng Comm*. 2007;194:218–233.
78. Nomikos P. Detection and diagnosis of abnormal batch operations based on multi-way principal component analysis. *ISA Trans*. 1996;35:259–266.
79. Westerhuis JA, Gurden SP, Smilde AK. Generalized contribution plots in multivariate statistical process monitoring. *Chemom Intell Lab Syst*. 2000;51:95–114.
80. Raich A, Cinar A. Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *AIChE J*. 1996;42:995–1009.
81. Kourti T, MacGregor JF. Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemom Intell Lab Syst*. 1995;28:3–21.
82. Hyvärinen A, Karhunen J, Oja E. A fast fixed-point algorithm for independent component analysis. *Neural Computing*. 1997;9:1483–1492.

Manuscript received Jan. 23, 2008, revision received May 26, 2008, and final revision received July 25, 2008.