APPLICATION OF STATISTICAL PROCESS MONITORING WITH EXTERNAL ANALYSIS TO AN INDUSTRIAL MONOMER PLANT

Tetsuhiko Yamamoto * **Akirou Shimameguri** *
Morimasa Ogawa ** **Iori Hashimoto** *** **Manabu Kano** ***

* *Mitsubishi Chemical Corporation, Ibaraki 314-0102, Japan*
** *Mitsubishi Chemical Corporation, Okayama 712-8054, Japan*
*** *Kyoto University, Kyoto 606-8501, Japan*

Abstract: The main objective of this industry-university collaboration is to develop an on-line process monitoring system that can detect a particular malfunction in an industrial monomer plant. The most serious malfunction is a blockage caused by an accumulation of polymers inside a cooling unit. Since the blockage requires shutdown maintenance, it is crucial to detect its symptom as early as possible and properly adjust the operating condition to avoid further polymer accumulation. The developed on-line monitoring system can detect the symptom of the blockage by using multivariate statistical process control, distinguish it from normal changes in operating conditions by using external analysis, and persuade operators to take appropriate action.

Keywords: process monitoring, fault detection, fault diagnosis, statistical process control, principal component analysis, external analysis, industrial application

1. INTRODUCTION

Long-term stable operation is becoming increasingly important in the chemical industry, because 1) a trouble shutdown of one plant inflicts a heavier loss on the company as production sites become more consolidated, and 2) plant managers have to get the best out of existing equipments and maximize the production efficiency. To achieve long-term efficient operation, one needs to recognize that:

- It seems impossible to entirely avoid troubles due to process upsets or equipment malfunction.
- Unexpected trouble may happen at an unexpected location during a high load operation that has never been experienced.
- The integration of operating rooms and the deployment of advanced process control systems reduce the number of operators; each operator's responsibilities have increased.

In modern chemical plants, operators must monitor a large number of process variables one after another for safe operation. Since measured process variables are highly correlated, it is difficult for operators to detect every fault without monitoring the correlation between process variables. A more difficult task than fault detection is to identify a real cause of the fault and to take prompt and appropriate action. To support operators, automation of process monitoring is greatly desired in the industry.

Multivariate statistical process control (MSPC) has been investigated as a data-based technique for multivariable process monitoring (Kresta et al., 1991; Kourti and MacGregor, 1995; Ku et al., 1995; Kano et al., 2002). MSPC is based on chemometric techniques such as principal component analysis (PCA) and partial least squares (PLS). PCA is a tool for data compression and information extraction; it finds linear combinations of variables that describe major trends in a data set. On the other hand, PLS relates output variables to latent variables, which are given as linear combinations of input variables. A typical application of PLS in the chemical industry is to estimate product quality from measurable variables (Kano et al., 2000). These chemometric techniques are very useful for modeling and monitoring chemical processes where a great number of measured variables are highly correlated. Many researchers and practitioners have investigated MSPC to extract useful information from process data and use it for process monitoring.

In conventional SPC, a process is assumed to be operated in a particular steady state, and deviations of measurements from their steady-state values are used for monitoring. However, operating conditions cannot be constant in many processes due to production rate adjustments, product grade transitions, and so on. Therefore, it is crucial to develop a new SPC method that can cope with changes in operating conditions. In order to develop a new monitoring system for distinguishing between faults and normal changes in operating conditions, external analysis was proposed and integrated with MSPC (Kano et al., 2003).

In the present work, an on-line monitoring system is developed to detect a blockage, caused by an accumulation of polymers, in an industrial monomer plant. Since the blockage requires shutdown maintenance, it is crucial to detect its symptom as early as possible and properly adjust the operating condition for avoiding further polymer accumulation.

2. MSPC WITH EXTERNAL ANALYSIS

In the present work, changes in operating conditions, which should be distinguished from faults, are assumed to be given from the outside of a process as changes in a feed flow rate, set-points of controllers, and so on. Thus, variables that are used for monitoring can be classified into two groups. The first group consists of variables representing operating conditions such as a feed flow rate and a set-point, hereafter referred to as external variables. The second group consists of variables affected by external variables and other unmeasured disturbances. Those variables are referred to as main variables. Changes in external variables are not faults. Therefore, both the changes in external variables and their influence on main variables should be distinguished from faults. To achieve this goal, operation data of main variables are decomposed into two parts: one is a part explained by external variables, and the other is a part not explained by them. As a result, the influence of changes in external variables can be removed from operation data. This technique is called external analysis and it can be integrated with any SPC method (Kano et al., 2003).

In this section, it is briefly shown that the external analysis can be used for removing the influence of external variables from operation data and it can be integrated with PCA-based SPC.

2.1 *External Analysis*

Consider a data matrix $\boldsymbol{X} \in \Re^{k\times m}$, where k and m are the number of samples and that of variables, respectively. For simplicity, each variable is assumed to be normalized. When m_g of m variables are classified as external variables and $m_h(= m - m_g)$ are main variables, the data matrix is described as

$$\boldsymbol{X} = \left[\, \boldsymbol{H} \;\; \boldsymbol{G} \,\right] \tag{1}$$

where $\boldsymbol{G} \in \Re^{k\times m_g}$ consists of external variables and $\boldsymbol{H} \in \Re^{k\times m_h}$ consists of main variables. The data matrix $\boldsymbol{H}$ of main variables should be decomposed into two parts: a part explained by the data matrix $\boldsymbol{G}$ of external variables and the other part not explained. For this purpose, multiple linear regression analysis can be used by regarding external variables and main variables as inputs and outputs, respectively. That is, a regression coefficient matrix $\boldsymbol{C} \in \Re^{m_g\times m_h}$ is determined so that the sum of squared errors or the squared Frobenius norm of an error matrix is minimized.

$$\boldsymbol{C} = (\boldsymbol{G}^T\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{H} \tag{2}$$

where the error matrix $\boldsymbol{E} \in \Re^{k\times m_h}$ is defined as

$$\boldsymbol{E} = \boldsymbol{H} - \boldsymbol{G}\boldsymbol{C}. \tag{3}$$

As a result, the main data matrix $\boldsymbol{H}$ can be decomposed into two parts, $\boldsymbol{GC}$ and $\boldsymbol{E}$. $\boldsymbol{GC}$ is a part explained by the external variables, and $\boldsymbol{E}$ is the other part that cannot be explained by the external variables. Any SPC method can be used for monitoring error variables. Equation (2) can be used only if external variables are linearly independent of each other. When external variables are highly correlated to each other, a multivariate data analysis technique such as PLS, which can cope with a collinearity problem, should be used instead of ordinary least squares.

When process dynamics cannot be ignored, the influence of changes in external variables cannot be removed from operation data by using the static external analysis. In such a case, a dynamic model must be built. Kano et al. (2003) have proposed dynamic external analysis, and they have shown that the dynamic external analysis can be successfully applied to a chemical process.

2.2 *MSPC Integrated with External Analysis*

The basic statistic to monitor $\boldsymbol{E}$ in Eq. (3) is the Hotelling T^2 statistic. The Hotelling T^2 control chart is an original Shewhart-type control chart for correlated variables, and it is related to PCA-based SPC. PCA-based SPC was further investigated and a residual analysis was developed (Jackson and Mudholkar, 1979). In recent years, the T^2 statistic of several important principal components and the Q statistic, which is the sum of squared residuals or the sum of prediction errors (SPE), are usually used for statistical process monitoring. The T^2 statistic of principal components is defined as

$$T^2 = \sum_{r=1}^{R} \frac{t_r^2}{\sigma_{t_r}^2} \tag{4}$$

where t_r is a score of the r-th principal component and $\sigma_{t_r}^2$ is its variance. R denotes the number of principal components retained in the PCA model. The score t_r is defined as

$$\left[\, t_1 \ t_2 \ \cdots \ t_R \,\right] = \boldsymbol{e}\boldsymbol{P} \tag{5}$$

where $\boldsymbol{e} \in \Re^{1\times m_h}$ is an error vector, which is a row of $\boldsymbol{E}$, and $\boldsymbol{P} \in \Re^{m_h \times R}$ is a loading matrix. On the other hand, the Q statistic is defined as

$$Q = \sum_{i=1}^{m_h} (e_i - \hat{e}_i)^2 \tag{6}$$

where e_i and $\hat{e}_i$ are a calculated value of the i-th error variable and its predicted (reconstructed) value, respectively. $\hat{e}_i$ is derived from

$$\left[\, \hat{e}_1 \ \hat{e}_2 \ \cdots \ \hat{e}_{m_h} \,\right] = \boldsymbol{e}\boldsymbol{P}\boldsymbol{P}^T \tag{7}$$

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.

3. MONITORING A MONOMER PLANT

This section introduces an application of MSPC integrated with external analysis to a monomer plant of Mitsubishi Chemical Corporation. The main objective of this collaborative research project is to develop a monitoring system that can detect a particular malfunction as early as possible. The malfunction to detect is a blockage in an equipment of the monomer plant. Since the blockage, caused by an accumulation of polymers, requires shutdown maintenance, it is crucial to detect its symptom as early as possible and properly adjust the operating condition to prevent polymers from further accumulating. Conventional MSPC does not function well because it cannot distinguish the blockage from normal changes in operating conditions such as load changes. In the present work, therefore, external analysis is used to remove the influence of operating condition changes from process variables, and the error is monitored by using PCA-based SPC.

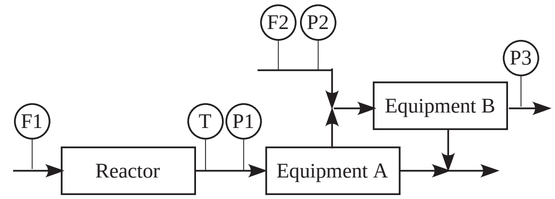

Fig. 1. Simplified PFD of the monomer plant.

Table 1. Process variables.

Symbol in Fig.1	Variable
F1	Raw material feed flow rate
F2	Recovery feed flow rate
P1	Equipment A inlet pressure
P2	Equipment A outlet pressure
P3	Equipment B outlet pressure
T	Reactor outlet temperature

3.1 *Malfunction in the Monomer Plant*

The process flow of the monomer plant is shown in Fig. 1. The product monomers are produced in the reactor, and then the reactant is condensed in the equipment A. Undesirable polymerization reactions take place under specific conditions in the equipment A although the operating condition is controlled to prevent monomers from polymerizing. The accumulation of polymers inside the equipment A blocks the flow and makes stable operation impossible.

Several important process variables are listed in Table 1. The symptom of the blockage could be detected by monitoring changes in differential pressure, P1-P2, because the blockage affects the pressure drop in the equipment A. The differential pressure will increase as more polymers accumulate. This monitoring strategy based on the differential pressure is very simple and easy to understand, but it is useful only when the differential pressure is not affected by other factors. In practice, not only polymer accumulation but also flow rates affect the differential pressure. For efficient monitoring, it is necessary to take into account the influence of operating conditions on the differential pressure.

3.2 *Analysis of Abnormal Conditions*

Trend graphs of the measured process variables listed in Table 1 are shown in Fig. 2. The sampling period of each variable is one hour, and each graph includes 7500 samples (about 10 months). All six variables, except the differential pressure, P1-P2, are mean-centered.

The trend of P1-P2 shows that the differential pressure began to increase around 2000 hours. The uptrend of the differential pressure indicates the possibility of the polymer accumulation in the equipment A. Finally, at 2300 hours, operators gave up carrying on the operation and shut down the monomer plant. The blockage caused by polymer accumulation was found inside the equipment A. The monomer plant was restarted at 2800 hours after a considerable part of accumulated polymers were removed. However, further polymer accumulation proceeded after 3500 hours, and then the plant was shut down again. The plant was restarted at 4500 hours after the whole accumulated polymers were removed. The differential pressure increased again after the second start-up. In particular, the differential pressure after 6500 hours is higher than that in the period when polymers blocked the flow (2000-2300 hours). However, polymers did not accumulate and a blockage did not occur in that period. This fact indicates that a rise in the differential pressure does not necessarily mean a blockage and the differential pressure is affected by other factors. After 6500 hours, the recovery feed flow rate F2 decreased and consequently the pressures P1, P2, and P3 decreased. This change caused the differential pressure to increase. In addition, a load change also affects the differential pressure. The pressure measurements P1, P2, and P3 increased from 4500 to 6000 hours as the feed flow rate F1 increased. In this period, the differential pressure P1-P2 increased because of high throughput.

3.3 *Design of Monitoring System*

A rise in the differential pressure is a useful indicator for detecting the blockage, but it is also affected by operating conditions such as a feed flow rate. Therefore, the influence of operating conditions has to be removed from the differential pressure. For this purpose, the static external analysis was used. In this application, only static properties of the process should be taken into account because the sampling period is one hour.

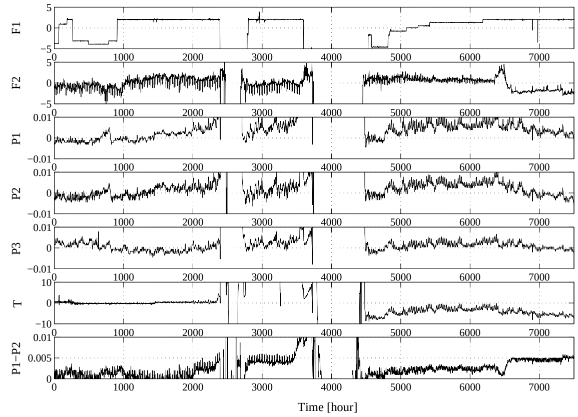

Fig. 2. Time-series plot of process variables.

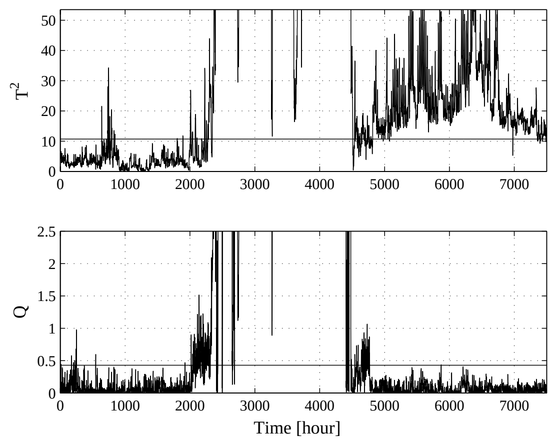

Fig. 3. Time-series plot of the T^2 and Q statistics.

The variables listed in Table 1 were used for monitoring. The external variables are the raw material feed flow rate F1 and the recovery feed flow rate F2. Those variables represent operating conditions. The other four variables were used as main variables.

External analysis and PCA were applied to data in the period when the process is operated under the normal condition (0-2000 hours), and the developed model was validated by the other data (2001-7500 hours except two shutdown periods). The number of principal components was selected so that the Q statistic could increase markedly when polymers accumulated. The selected number of principal components was three, and 98% of the variance in the reference data can be explained by three principal components. The control limits of two statistics are determined so that the number of samples outside the control limit is 1% of the entire samples while the process is operated under a normal condition. The control limits of T^2 and Q are 10.7 and 0.43, respectively.

3.4 *Monitoring Results*

The monitoring results, the trend graphs of T^2 and Q, are shown in Fig. 3. Although the

T^2 statistic exceeds its control limit from 2300 through 4500 hours, it also exceeds its control limit after 4500 hours. Therefore, the T^2 statistic is not a suitable index for detecting polymer accumulation. This result is not surprising because the T^2 statistic is a measure of the variation within the PCA model. The T^2 statistic is mainly affected by operating condition changes, which do not affect the correlation structure. As shown in Fig. 2, the operating condition of the monomer plant before 2000 hours is considerably different from that after 4500 hours. For example, the reactor outlet temperature is almost constant before 2000 hours, but it becomes lower and fluctuates wildly after 4500 hours. Such changes make the T^2 statistic exceed its control limit even though any fault does not occur. Since changes in the reactor outlet temperature cannot be explained by the external variables F1 and F2, the changes affect the T^2 statistic even when external analysis is conducted.

On the other hand, the Q statistic remarkably increases after 2000 hours and exceeds its control limit. The Q statistic is about 50 from 2800 to 3800 hours when polymers are blocking the flow. In addition, the Q statistic is under its control limit after accumulated polymers are removed. The Q statistic after 4800 hours is similar to that of the reference data even though the operating conditions are quite different from each other and the differential pressure increases considerably in this period. This result demonstrates the usefulness of the Q statistic for detecting polymer accumulation. It should be noted here, however, that conventional PCA-based SPC does not function well in this application. It cannot distinguish between polymer accumulation and operating condition changes. The key to success is to remove the influence of operating condition changes from monitored variables by conducting the external analysis.

Figure 3 shows that the Q statistic is a suitable index for detecting polymer accumulation. However, it is necessary to confirm that polymer accumulation is the real cause of the abnormal condition because another factor may affect the Q statistic and make it exceed the control limit. To identify the variables that contribute significantly to an out-of-control value of the Q statistic, contributions from process variables to the Q statistic can be used (Nomikos, 1996). This information helps operators to further diagnose an actual cause of the fault. A contribution of the i-th variable to the Q statistic is defined as

$$C_i = e_i - \hat{e}_i. \tag{8}$$

Contributions from four main variables to the Q statistic at the 2350th step are shown in Fig. 4.

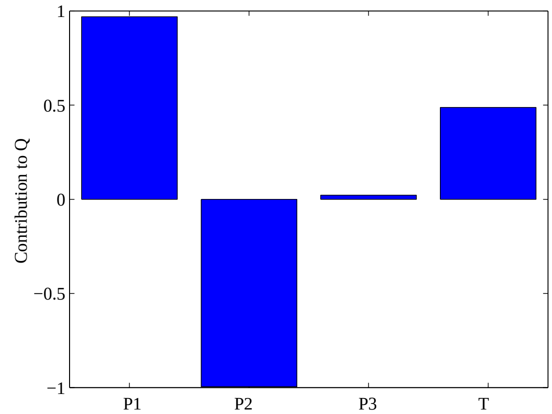

Fig. 4. Contribution plot of the Q statistic.

It is clear from this contribution plot that the contributions of the pressure of the equipment A, P1 and P2, are significant. P1 is positive large, and P2 is negative large. This result indicates that the differential pressure P1-P2 contributes significantly to the out-of-control value of the Q statistic. Therefore, polymer accumulation is the most possible cause. On the basis of this diagnosis, the operating condition should be adjusted to avoid a blockage caused by polymer accumulation.

3.5 *On-line Monitoring*

To monitor this monomer plant, in particular, to detect polymer accumulation, an on-line monitoring system was developed. The developed monitoring system performs the following procedures:

(1) Calculates the mean and the standard deviation of each monitored (external and main) variable, determines the regression coefficient matrix used for the external analysis, and builds the PCA model. This step is conducted off-line by using the reference data.
(2) Collects data every hour.
(3) Normalizes the data.
(4) Applies the external analysis to the normalized data.
(5) Calculates the Q statistic.
(6) Compares the calculated Q statistic and its control limit, and gives an alarm if Q exceeds its control limit.

The calculated Q statistic is stored in the database every hour, and its trend graph can be checked with trend graphs of monitored variables if necessary.

After the installation of the on-line monitoring system, polymer accumulation proceeded again in the equipment A. The external variables, F1 and F2, and the Q statistic are shown in Fig. 5. The recovery feed flow rate F2 was kept

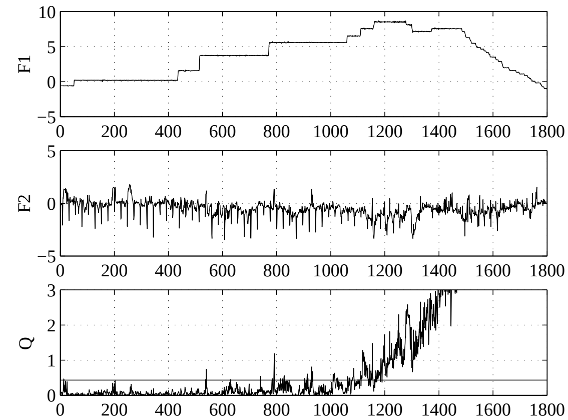

Fig. 5. On-line monitoring result.

almost constant during this period, but the raw material feed flow rate F1 was increased stepwise several times. The changes in the raw material feed flow rate did not affect the Q statistic, because the influence of the external variables was successfully removed from the monitored variables by using the external analysis. The Q statistic, however, increased steadily and exceeded its control limit after 1100 hours. This chart helped operators to suspect polymer accumulation and persuaded them to examine the equipment A. It was confirmed that polymer accumulation was proceeding. To avoid further accumulation of polymers and also to cope with the decrease in heat transfer efficiency, the raw material feed flow rate was decreased after 1280 hours. This result demonstrates that the developed on-line monitoring system, which integrates MSPC with external analysis, is useful for detecting polymer accumulation and avoiding a serious blockage in the monomer plant.

Another approach to avoid further polymer accumulation is to increase an inhibitor feed flow rate or a diluent water feed flow rate, which is used for cooling the reactant. Those approaches are useful, but they cannot remove the accumulated polymers. In the monomer plant, the equipment A consists of a number of parallel units. Therefore, the equipment A can be partially shut down and the accumulated polymers can be removed. As a result, by detecting polymer accumulation and avoiding a serious blockage, the developed monitoring system enables long-term, safe, and efficient operation of the monomer plant.

4. CONCLUSIONS

In this research project, PCA-based SPC was integrated with external analysis and applied to an industrial monomer plant. The developed monitoring system can distinguish polymer accumulation, which causes a serious blockage of the flow, from normal changes in operating conditions by using the static external analysis, and thus the system can successfully detect the polymer accumulation at its early stage. Although the developed monitoring system focuses only on the polymer accumulation in a particular equipment, it can detect symptoms of the most serious malfunction and persuade operators to take prompt and appropriate action. In practice, a reliable *specialist* is preferable to a moderate *generalist*. Various MSPC methods have been developed for general purposes in the last decade or so, but more important problems to investigate are how to diagnose the real cause of a serious fault and how to help operators to take appropriate action. Those problems seem to remain unsolved.

REFERENCES

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

Kano, M., S. Hasebe, I. Hashimoto, and H. Ohno (2003). Evolution of Multivariate Statistical Process Control: Application of Independent Component Analysis and External Analysis. *Proc. of The Foundations of Computer Aided Process Operations Conference (FOCAPO)*, 358-388, Coral Springs, US, Jan. 12-15.

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., K. Nagao, H. Ohno, S. Hasebe, I. Hashimoto, R. Strauss, and B.R. Bakshi (2002). Comparison of Multivariate Statistical Process Monitoring Methods with Applications to the Eastman Challenge Problem. *Comput. Chem. Engng*, **26**, 161-174.

Kourti, T. and J.F. MacGregor (1995). Process Analysis, Monitoring and Diagnosis, Using Multivariate Projection Methods. *Chemometrics and intelligent laboratory systems*, **28**, 3-21.

Kresta, J.V., J.F. MacGregor, and T.E. Marlin (1991). Multivariate Statistical Monitoring of Process Operating Performance. *Can. J. Chem. Eng.*, **69**, 35-47.

Ku, W., R.H. Storer, and C. Georgakis (1995). Disturbance Detection and Isolation by Dynamic Principal Component Analysis. *Chemometrics and intelligent laboratory systems*, **30**, 179-196.

Nomikos, P. (1996). Detection and diagnosis of abnormal batch operations based on multi-way principal component analysis. *ISA Trans.*, **35**, 259-266.