Korean J. Chem. Eng., **20**(6), 977-984 (2003)

Application of Partial Least Squares Methods to a Terephthalic Acid Manufacturing Process for Product Quality Control

In-Su Han, Minjin Kim, Chang-Hyun Lee*, Woonou Cha*, Byoung-Kyoung Ham, Jy-Hyo Jeong**, Haksoo Lee**, Chang-Bock Chung*** and Chonghun Han†**

Department of Chemical Engineering, Pohang University of Science and Technology, Pohang, Gyeongbuk 790-784, Korea
*e-Management Team, SK Chemicals, Suwon, Gyeonggi 440-745, Korea
**Petrochemical Production Team, SK Chemicals, Ulsan 680-160, Korea
***Faculty of Chemical Engineering, Chonnam National University, Gwangju 500-757, Korea

(*Received 31 September 2002 • accepted 8 August 2003*)

Abstract−This paper deals with an application of partial least squares (PLS) methods to an industrial terephthalic acid (TPA) manufacturing process to identify and remove the major causes of variability in the product quality. Multivariate statistical analyses were performed to find the major causes of variability in the product quality, using the PLS models built from historical data measured on the process and quality variables. It was found from the PLS analyses that the variations in the catalyst concentrations and the process throughput significantly affect the product quality, and that the quality variations are propagated from the oxidation unit to the digestion units of the TPA process. A simulation-based approach was addressed to roughly estimate the effects of eliminating the major causes on the product quality using the PLS models. Based on the results that considerable amounts of the variations in the product quality could be reduced, we have proposed practical approaches for removing the major causes of product quality variations in the TPA manufacturing process.

Key words: Terephthalic Acid, Partial Least Squares (PLS), Quality Control, Empirical Modeling, Multivariate Analysis

INTRODUCTION

Terephthalic acid (TPA) is a monomer used to manufacture polyethylene terephthalate (PET), which then is formed into films, textiles, bottles, and plastic molds. The manufacturing processes are receiving increased attention due to steady growth in the demand for PET in the expanding market. Several types of commercial processes each licensed by Amoco, Eastman-Kodak, or Mitsubishi are being operated to manufacture TPA in the world [Kroschwitz, 1991]. In all these commercial processes, *p*-xylene is partially oxidized by air to TPA, where 4-carboxybenzaldehyde (4-CBA) is inevitably formed as an undesirable by-product. The amount of the 4-CBA contained in the TPA product mainly determines the product quality since the 4-CBA hinders stable polymerization of TPA in the subsequent PET manufacturing process [Kim et al., 2001]. Hence, it is desirable to minimize not only the formation of 4-CBA but also its fluctuation in the TPA manufacturing process to keep up the uniformity in the PET product quality during a certain period of operation. To achieve the goal, TPA manufacturers have adopted various process systems approaches such as modeling, optimization, control, and statistical analysis of the process [Jaisinghani et al., 1997; Cincotti et al., 1999].

Multivariate statistical projection methods, such as partial least squares (PLS) methods, principal component analysis (PCA) or principal component regression (PCR) are widely used as powerful tools to identify the sources of quality deviations and process faults and to model processes and product qualities in various industries. Eriksson et al. [1995] carried out multivariate analyses of aquatic toxicity data using a PLS method and compared its performance with an MLR (multiple linear regression) method. MacGregor and Kourti [1995] gave an overview of these methods and applied PCA and PLS methods to an industrial polymerization process for online monitoring and fault diagnosis. Fujii et al. [1997] used a PLS method to select the important variables for empirical modeling as well as to build the models for predicting the top composition in a distillation column. Recently, Hong et al. [1999] employed a PLS method to design a soft sensor for the prediction of toluene composition in an industrial splitter column. They divided a data set into several sub-groups using a PCA clustering method to improve the prediction capability of the soft sensor by constructing several local PLS models. Liu et al. [2000] proposed a nonlinear PLS method by extending the conventional linear PLS method to a nonlinear framework and applied their PLS method to a distillation column. More recently, Han and Han [2003] developed a hybrid model by combining a thermodynamic compression/expansion model into a PLS model to predict the power consumption/generation rates of an industrial compression/expansion system.

In this article, we present an application of a PLS method to an industrial TPA manufacturing process to investigate the major causes of variability in the quality (4-CBA concentration) of the TPA product. First, we present a brief overview of PLS methods and the TPA manufacturing process. Then, the PLS methods are applied to the process to find the major causes of variability in the product quality and to simulate the effects of eliminating these causes. Finally, practical approaches are proposed to get rid of the revealed causes in

---

†To whom correspondence should be addressed.
E-mail: chan@postech.ac.kr

the TPA manufacturing process.

BACKGROUND ON PARTIAL LEAST SQUARES METHODS

Multivariate statistical projection methods are recognized as state-of-the-art techniques for the analysis and modeling of chemical processes, and comprise several specific methods such as PCA, PCR, and PLS [Fujii et al., 1997; Neogi and Schlags, 1998]. Since this study mainly relies on a PLS method, we give a brief overview of it, referring the readers to the literature for the other methods [Wold et al., 1987; Wise and Gallagher, 1996].

A PLS method has been widely used as a powerful tool for constructing empirical models from lab and field measurement data. It typically provides more robust and reliable models than ordinary least squares methods, particularly when the data are noisy and highly correlated with each other [Eriksson et al., 1995; Neogi and Schlags, 1998]. The basic concept of the PLS method is to project the high dimensional spaces of the input and output data obtained from a process onto the low dimensional feature (latent) spaces and then to find the best relation between the feature vectors. It is capable of dealing with singular and highly correlated regression problems which the traditional multiple linear regression methods cannot handle. In addition, it enables the modeling results to be easily interpreted by providing helpful information in the form of scores, loadings, and regression coefficients.

The first step in a PLS modeling is to arrange the measurements on k process variables and quality variables at n different sampling times into an n×k process data matrix **X** and an n×m output matrix **Y**, respectively. Then, after being scaled and mean-centered, each **X** and **Y** matrix is decomposed as a sum of series of rank-one matrices according to the following outer relations:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{P}_a^T + \mathbf{E} \tag{1}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} = \sum_{a=1}^{A} \mathbf{u}_a \mathbf{q}_a^T + \mathbf{F} \tag{2}$$

Finally, the inner relation is described by the following equation:

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F} = \mathbf{X}\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{Q}^T + \mathbf{F} \tag{3}$$

In the above, **T** and **U** represent the score matrices that summarize **X** and **Y** variables, respectively, while **P** and **Q** the loading matrices that show the influences of **X** and **Y**, respectively. The score vectors $\mathbf{t}_a$ are calculated sequentially from the data for each latent variable a (also called the PLS dimension) such that the linear combinations of the **X** and **Y** variables defined by $\mathbf{t}_a=\mathbf{X}\mathbf{p}_a$ and $\mathbf{u}_a=\mathbf{Y}\mathbf{q}_a$ maximize the covariance between **X** and **Y** that is explained at each latent variable. The total number of latent variables A is typically much lower than the number of process variables k and is usually determined by means of cross-validations [Geladi and Kowalski, 1986]. If one uses all the latent variables (A=k) to describe **X** and **Y** variables, the residual matrices **E** and **F** will be zeros. The PLS regression coefficients **B** in Eq. (3) are determined from the underlying PLS model and can be used to interpret how the process variables **X** are correlated to the quality variables **Y**. Typically, the most instructive method to calculate the PLS model parameters including the scores (**T** and **U**), the loadings (**P** and **Q**), and the weights (**W**) is known as the nonlinear iterative partial least squares (NIPALS) algorithm in which the PLS parameters are computed sequentially for each latent variable. Details on the general concepts of the PLS modeling and the NIPALS algorithm are shown in the literature [Geladi and Kowalski, 1986].

PROCESS DESCRIPTION

A commercial process capable of processing *p*-xylene of about 430,000 tons per year is currently operated to manufacture TPA in Korea; the name of the process is concealed not to disclose proprietary information. Fig. 1 shows a simplified process flow diagram of this process consisting of the major six unit processes: oxidation, centrifuging, digestion, filtering and drying, catalyst purification, and solvent separation. In the oxidation process, pressurized-air directly oxidizes *p*-xylene to TPA and other by-products in the presence of a combination of cobalt, manganese, and bromine as

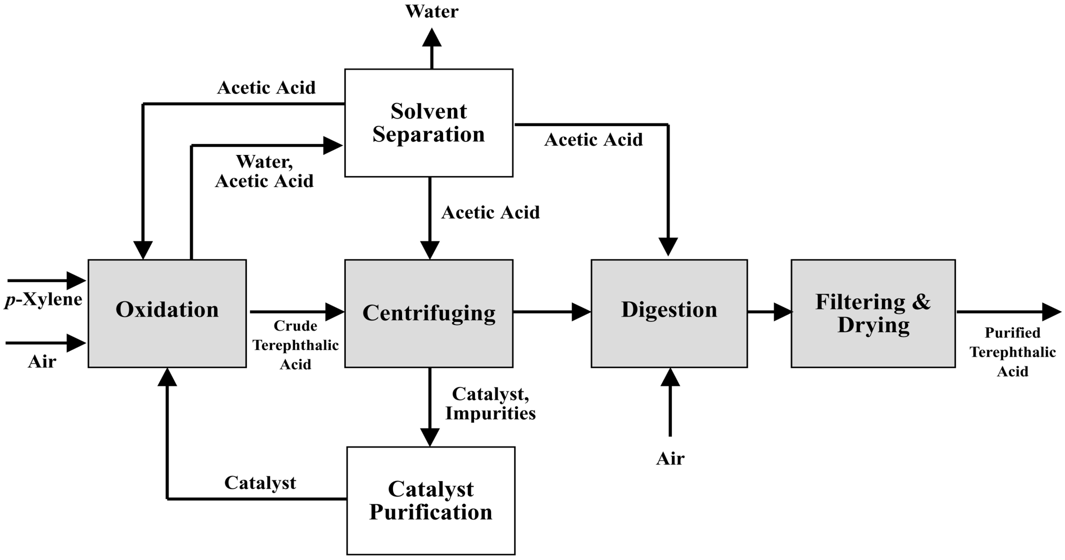

Fig. 1. Simplified process flow diagram of the terephthalic acid manufacturing process.

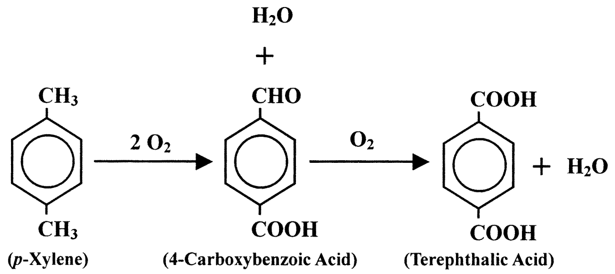

Fig. 2. Oxidation of *p*-xylene to terephthalic acid and 4-carboxybenzaldehyde.

catalyst and acetic acid as solvent. Crude TPA is produced from the oxidation process according to the overall reaction scheme shown in Fig. 2. The actual reaction mechanisms are more complex than shown in Fig. 2 and would include the generation mechanisms for such minute products as *p*-toluic acid, tolualdehyde, terephthaldehyde, and carbon dioxide [Cincotti et al., 1999]. As can be seen in the overall reaction, TPA is inevitably accompanied by 4-CBA, which is the undesirable by-product whose concentration in the crude TPA typically ranges from 3,000 and 8,000 ppm. Over 98% conversion of *p*-xylene is achieved in the oxidation process where the major controlled variables are oxidation temperature, residence time, catalyst to *p*-xylene ratio, and so forth. A portion of the crude TPA coming from the oxidation process is sold as a product after going through the filtering and drying process while the rest is sent in a slurry state to the centrifuging process. In the centrifuging process, catalyst liquid is separated from the slurry and then recycled to the catalyst purification process. The crude TPA is sent along with a small amount of catalyst from the centrifuging process to the digestion process where 4-CBA is further oxidized to TPA by pressurized air, resulting in lowered concentration below 150 ppm. The major controlled variables of the digestion process are temperature, residence time, and air flow-rate. The filtering and drying process is responsible for filtering solid-type impurities and for drying wet TPA to yield final dried-powder product. The solvent separation process consisting of several distillation columns serves to separate the water generated in the oxidation step from the acid-water mixture, and the resulting highly concentrated acetic acid is then recycled to various unit processes. Since heavy organic compounds generated during oxidation contaminate the catalyst in liquid solvent, they are separated from the catalyst solvent in the catalyst purification process comprising several distillation columns.

MODELING AND ANALYSIS

The concentration of 4-CBA in the purified TPA (final product) is the major quality variable in TPA manufacturing processes. Hence, the key process variables affecting the oxidations of *p*-xylene and 4-CBA should be identified first to control the 4-CBA within a desired level. In the TPA manufacturing process shown in Fig. 1, the average concentration of 4-CBA over a certain period of production is quite low enough to satisfy the PET manufacturers to whom the purified TPA is supplied. However, the concentration of the 4-CBA shows large variations that render the quality control difficult in the PET manufacturing processes. In this section, the major causes of variations in the 4-CBA concentration are identified first, and then the effects of eliminating these causes on the variability are examined by simulating the product quality on the basis of the derived PLS models.

1. Data Preparation and Preprocessing

The real-time database (RTDB) system and the laboratory information management system (LIMS) have been running to collect the measurements on the process and quality variables for the whole process shown in Fig. 1. The RTDB system gathers over 1500 measurements every minute on the process operating variables, and the LIMS stores the measurements on the quality variables and catalyst concentrations which are sampled and analyzed once to eight times a day. Since the 4-CBA concentration only varies in the oxidation step, only the data measured in the oxidation and digestion processes are required for the PLS modeling and analysis. Tables 1 and 2 summarize the process and quality variables for the oxidation process and for the digestion process, respectively. All the PLS

Table 1. Process and quality variables used for the PLS modeling and analysis of the oxidation process

Variable	Description
V1	Solid in the slurry from the oxidation process, [wt%]
V2	Bromine in the catalyst liquid into the oxidation process, [ppm]
V3	Cobalt in the catalyst liquid into the oxidation process, [ppm]
V4	Manganese in the catalyst liquid into the oxidation process, [ppm]
V5	Iron in the catalyst liquid into the oxidation process, [ppm]
V6	Sprayed acetic acid into the oxidation process, [kg/min]
V7	Catalyst liquid into the oxidation process, [kg/min]
V8	*p*-xylene fed into the oxidation process, [kg/min]
V9	Acetic acid recycled into the oxidation process, [kg/min]
V10	Air into the oxidation process, [kg/min]
V11	Pressure of the air into the oxidation process, [kg/cm^2]
V12	Temperature of the air into the oxidation process, [C]
V13	Oxidizer level, [%]
V14	Oxidizer pressure, [kg/cm^2]
V15	Oxidizer temperature, [C]
V16	Ratio of V7 to V8
V17	Ratio of V10 to V8
V18	Ratio of V23 to V8
V19	Level of the reflux tank, [%]
V20	Level of the catalyst liquid tank 1, [%]
V21	Level of the catalyst liquid tank 2, [%]
V22	Acetic acid into the reflux tank, [kg/min]
V23	Total crude TPA from the oxidation process, [kg/min]
V24	Crude TPA into the digestion process, [kg/min]
V25	Excess oxygen in the oxidation process, [vol%]
V26	Ratio of V24 to V23
V27	Ambient temperature, [C]
V28	Relative humidity, [%]
V29	Off-gas Absorber pressure, [kg/cm^2]
Q1	4-CBA concentration in the crude TPA, [ppm]

Table 2. Process and quality variables used for the PLS modeling and analysis of the digestion process

Variable	Description
Q1	4-CBA concentration in the crude TPA, [ppm]
W1	Solid in the slurry from the digestion process, [wt%]
W2	Bromine in the catalyst liquid into the digestion process, [ppm]
W3	Cobalt in the catalyst liquid into the digestion process, [ppm]
W4	Manganese in the catalyst liquid into the digestion process, [ppm]
W5	Iron in the catalyst liquid into the digestion process, [ppm]
W6	Slurry from the oxidation process to the digestion process, [kg/min]
W7	Air into the digestion process, [kg/min]
W8	Acid vapor into the digestion process, [kg/min]
W9	Excess oxygen in the 1st digester, [vol%]
W10	Temperature of the gaseous phase in the 1st digester, [C]
W11	Inventory of the 1st digester, [ton]
W12	Pressure of the 1st digester, [kg/cm^2]
W13	Temperature of the reaction zone of the 1st digester, [C]
W14	Temperature of the slurry exiting the 1st digester, [C]
W15	Slurry circulation into the 2nd digester, [kg/min]
W16	Temperature of the acid vapor into the digestion process, [C]
W17	Temperature of the slurry feed into the digestion process, [C]
W18	Temperature of the gaseous phase in the 2nd digester, [C]
W19	Excess oxygen in the 2nd digester, [vol%]
W20	Inventory of the 2nd digester, [ton]
W21	Pressure of the 2nd digester, [kg/cm^2]
W22	Temperature of the reaction zone of the 2nd digester, [C]
W23	Slurry circulation into the slurry flash tank, [kg/min]
W24	Acetic acid into the slurry flash tank, [kg/min]
W25	Acetic acid vapor into the reflux tank, [kg/min]
W26	Acetic acid into the acetic acid tank, [kg/min]
W27	Product slurry into the product tank, [kg/min]
W28	Level of the slurry flash tank, [%]
W29	Temperature of the slurry flash tank, [C]
W30	Ambient temperature, [C]
W31	Relative humidity, [%]
Q2	4-CBA concentration in the purified TPA, [ppm]

models built here use the historical data measured during the last 10 month long operation of the TPA manufacturing process.

A problem arises here because the quality variables are measured much less frequently than the process variables and further because the time stamps marked on each measurement would not exactly match between the two types of variables due to time-delays. To resolve the problem, the original process data matrix with size n×k is rearranged into a new matrix with size l×(k×s) by selecting s rows located around the time stamps of the quality variables in the original process data matrix as time-shifted variables that accounts for time-delays [Wise and Gallagher, 1996; Dayal and MacGregor, 1996]. Fig. 3 depicts the arrangement of a process data matrix that accounts for time-delays as well as conforms the observations for the process variables to those for the quality variables.

After the data matrices were rearranged, statistical outliers that might appear due to measurement errors or abnormal operations were removed from the data set on the basis of principal component analysis [Wold et al., 1987]. Then, the following two data matrices (each of which includes the observations both for the process variables and for the quality variable) were prepared for the PLS modeling: the data matrix 1 with the size of 306 observations by 206 variables (205 input variables and 1 output variable) for the oxidation process, and the data matrix 2 with the size of 621 observations by 163 variables (162 input variables and 1 output variable) for the digestion process. Performing a PLS modeling to such an arranged data matrix is similar to identifying an ARX (Auto Regressive eXogeneous) model; details can be found in the literature [Wise and Gallagher, 1996; Dayal and MacGregor, 1996; Shi and MacGregor, 2000].

2. PLS Modeling and Analysis

To identify the correlations between the 4-CBA concentrations and the process variables including catalyst concentrations, PLS models for the oxidation and digestion processes are built from the data matrix 1 and 2, respectively. After cross-validations using the data matrices are performed, the final PLS model for the oxidation process has seven latent variables and explains 62.2% of the variance in the 4-CBA concentration of crude TPA (Q1). The PLS model for the digestion process needs eight latent variables with which 55.3% of the variability in the 4-CBA concentration of purified TPA (Q2) is captured. Note that the PLS modeling results are quite good if we consider the fact that there should exist numerous sensor faults and measurement errors in the process and quality variables of actual processes and that the PLS models attempt to explain the "common-cause" variations [MacGregor et al., 1994] in the process data and to exclude the random variations and measurement errors that are uncorrelated with other process and quality variables.

Fig. 4 compares the measured 4-CBA concentrations with the predicted ones in the crude TPA from the oxidation process. As shown in the figure, the predicted values agree well with the measured values in trend, and most of the deviations from the average value of the total observations can be predicted by the PLS model for the oxidation process. Fig. 5 compares the measured 4-CBA concentrations with the predicted ones in the purified TPA from the digestion process. Though the variance captured by the PLS model for the digestion process is relatively low compared with that for the oxidation process, the major trend (showing long-term variations) of the measured 4-CBA concentrations is well predicted as illustrated in the figure. Since the quality variations (Q1 and Q2) for a long operation period are rather undesirable than small fluctuations for a short operation period, the predictions of the long-term variations are more important than those of the short-term variations. Because the major trends are well predicted by using the PLS models as shown in Figs. 4 and 5, both the PLS models can be reliably used for analyzing the causes of variability in the 4-CBA concentrations.

Fig. 6 shows the regression coefficients of the PLS model for the oxidation process. The higher a bar in the figure is, the more

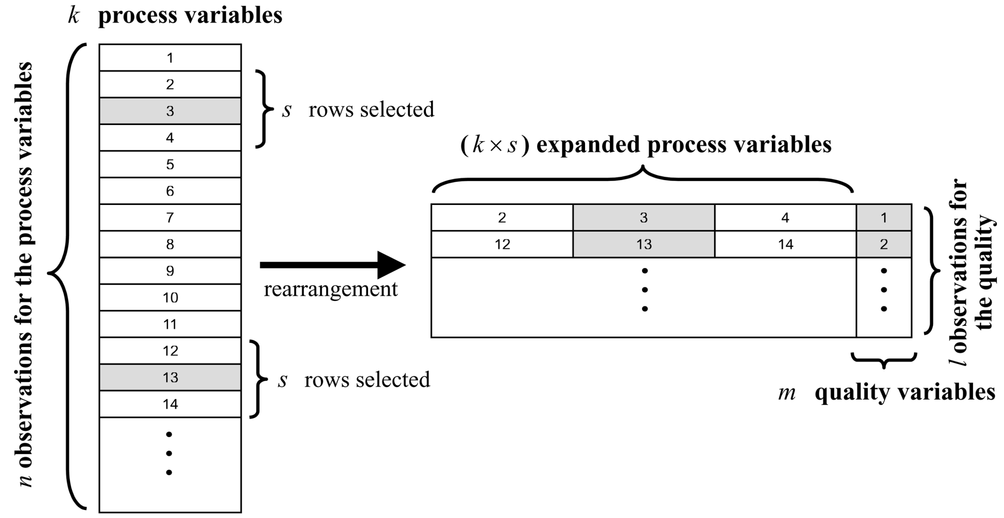

Fig. 3. Arrangement of an original process data matrix into a new process data matrix; the gray-colored rows indicate the observations for the process variables corresponding to the time stamps at which the quality variables are measured.

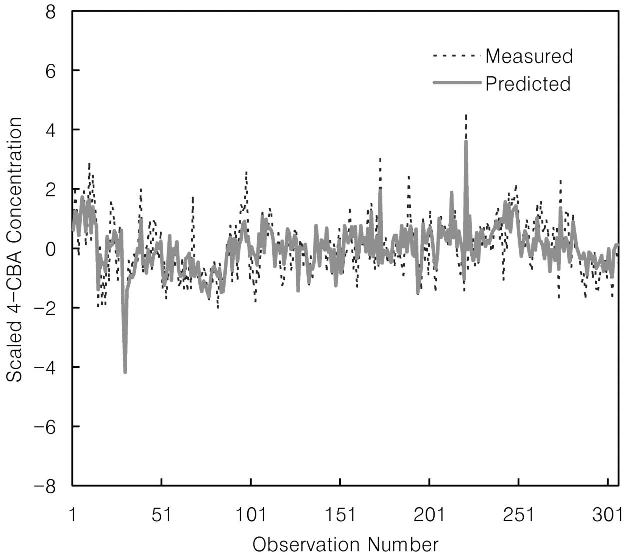

Fig. 4. 4-CBA concentrations of crude TPA predicted by the PLS model for the oxidation process.

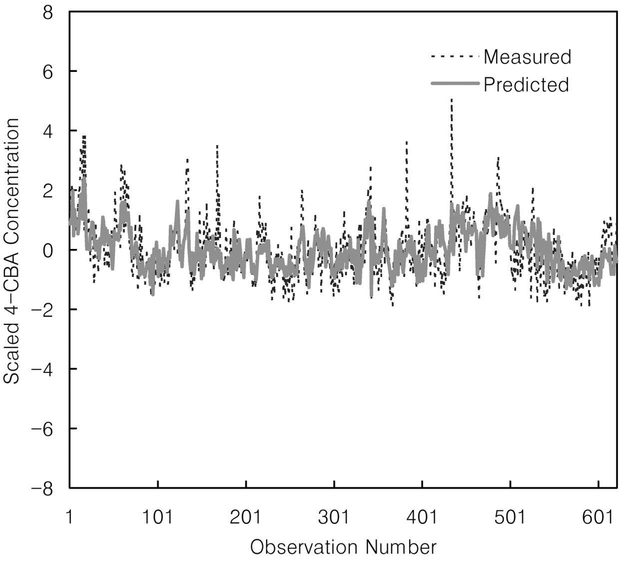

Fig. 5. 4-CBA concentrations of purified TPA predicted by the PLS model for the digestion process.

effect the corresponding variable has on the quality variable. And the sign of a regression coefficient denotes the direction of the effect on the quality variable. As shown in the figure, the concentrations of the catalysts (V2 and V4) into the oxidation process are strongly correlated to the 4-CBA concentration of crude TPA. Especially, the manganese concentration has a positive correlation to the concentration of the 4-CBA in the oxidation process, but the bromine has a strong negative correlation contrary to the manganese concentration. It implies that small fluctuations in the catalyst concentrations lead to large variability in the concentration of the 4-CBA in the crude TPA and thus that the catalyst concentrations should be elaborately controlled to reduce the quality variations. The process variables (V8 and V15) associated with the process throughput also highly affect the variability in the 4-CBA concentration of crude TPA. Therefore, we can reach the conclusion that the variations in the catalyst concentration and the changes in the process throughput are the major sources of variability in the 4-CBA concentration of crude TPA.

Fig. 7 shows the regression coefficients of the PLS model for the digestion process. As can be seen in the figure, the 4-CBA concentration of crude TPA (Q1) is strongly correlated with the 4-CBA concentration of purified TPA (Q2). It means that most of the variations in the purified TPA quality are propagated from the oxidation process, and consequently that the stabilization of oxidation process is very important to reduce the variability in the 4-CBA concentration of purified TPA. In the digestion process, the catalyst concentrations (W3 and W4) and the flow rate of air (W7) also affect the 4-CBA concentration of purified TPA as shown in Fig. 7.

3. Estimation of the Variability in the Product Quality

The PLS models are used to investigate the effects of the stabili-

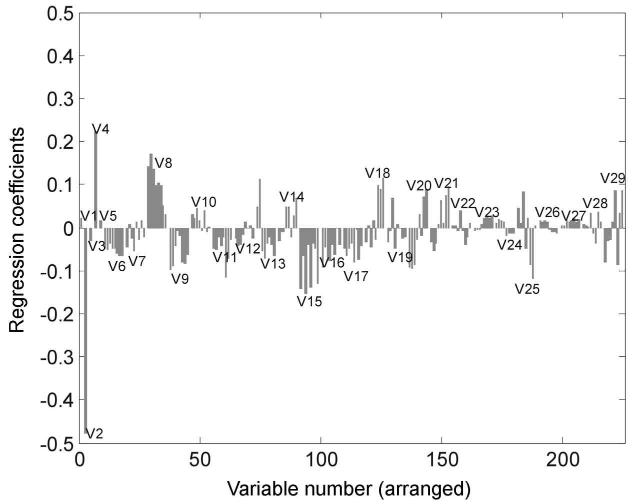

Fig. 6. Regression coefficients of the 4-CBA concentration of crude TPA (Q1) for the PLS model for the oxidation process (data matrix 1).

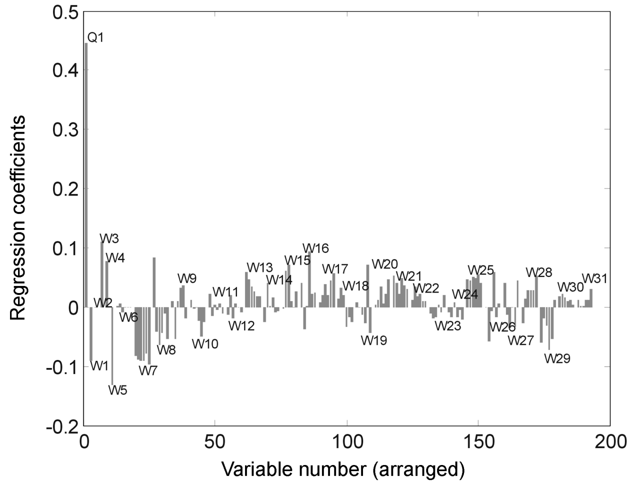

Fig. 7. Regression coefficients of the 4-CBA concentration of purified TPA (Q2) for the PLS model for the digestion process (data matrix 2).

zation of variability both in the catalyst concentrations and in the process throughput on the variations in the 4-CBA concentrations (Q1 and Q2). Using the PLS models for the oxidation and digestion processes, the simulation results are obtained at two different operating conditions: one is the actual condition during the period of the data sampling, and the other is an assumed scenario in which the catalyst concentrations (V2-4 and W2-4) and the process variables (V7-8, V10, V15-18, V23-24, V26, W6-7, W10, W13, W18, W22, and W17) associated with the process throughput are stabilized by intentionally reducing their standard deviations to halves the actual values. In all the simulations, the concentrations of 4-CBA (Q1) predicted by the PLS model for the oxidation process are used as an input vector of the PLS model for the digestion process. Then, the effect of the stabilization can be roughly estimated by comparing the simulation results obtained at the two different operating conditions under the assumption that all the process variables are independent of each other, hence freely adjustable.

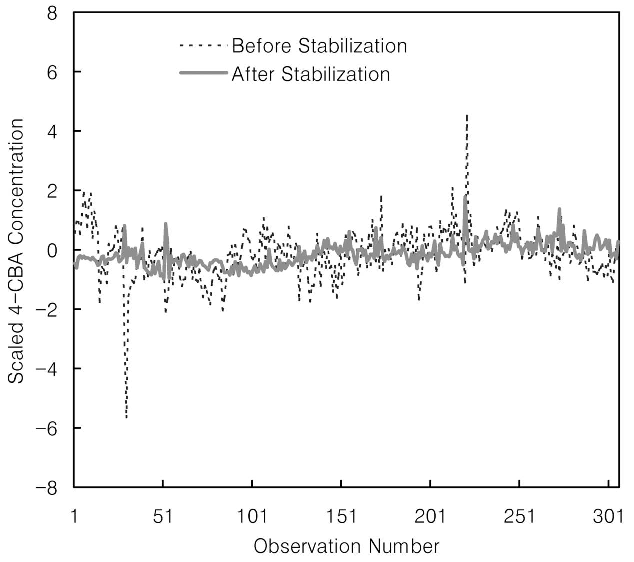

Fig. 8. Simulated 4-CBA concentrations of crude TPA before and after the stabilization.

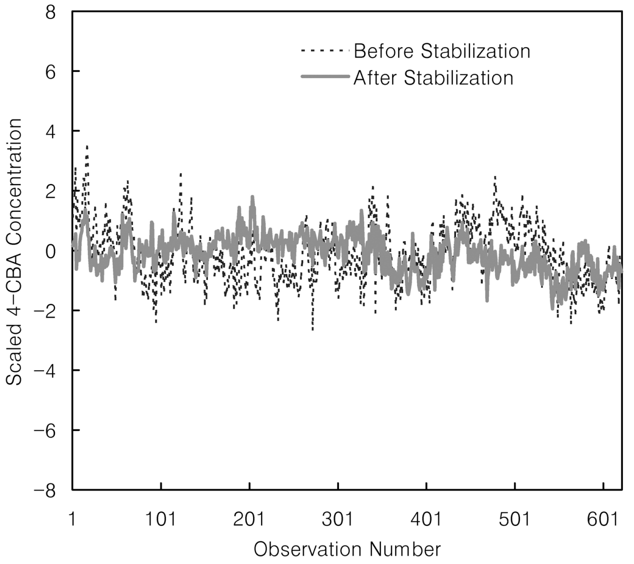

Fig. 9. Simulated 4-CBA concentrations of purified TPA before and after the stabilization.

Fig. 8 compares the 4-CBA concentrations of crude TPA after both the catalyst concentrations and the process throughput are stabilized with those before stabilization. As shown in the figure, the standard deviation of the 4-CBA concentrations of crude TPA considerably decreases by 51.5% after stabilization. Fig. 9 shows the 4-CBA concentrations of purified TPA after the catalyst concentrations and the process throughput are stabilized. After the simulations for the oxidation process and the digestion process are performed in succession, the standard deviation is reduced by 36.6%, which is rather lower than that for the crude TPA. Although the estimated reduction in variations obtained by using the PLS models would differ from the actual values due to that portion of the variations treated as noises in the model, the simulation approach presented here can be taken to provide a useful evaluation of the effect of stabilizing the major causes of variations.

PRACTICAL APPROACHES TO PRODUCT QUALITY CONTROL

As one can see from the PLS regression coefficients and the simulation results shown in the previous section, the major causes of variability in the 4-CBA concentrations are the fluctuations in the catalyst concentrations and changes in the process throughput. Thus, we have to get rid of these causes in order to achieve more uniform concentrations of 4-CBA both in the oxidation process and in the digestion process. In practice, to stabilize the catalyst concentrations in the TPA manufacturing process, the concentrations of the catalysts into the oxidation process should be measured online by using a near-infrared spectroscopy. Then, a control system that uses the measured concentrations can be devised to keep the ratio of catalyst to *p*-xylene at a constant value in the oxidation and digestion processes. A simple proportional-integral-derivative (PID) control system can be adopted for adjusting the flow rate of the catalyst liquid into the oxidation process if the only control objective is to regulate the total catalyst to *p*-xylene ratio. However, to regulate both the catalyst composition (cobalt, bromine, and manganese) and the total catalyst to *p*-xylene ratio, a more sophisticated control system such as a model predictive controller [Jaisinghani et al., 1997; Hur et al., 2003] should be employed for the TPA manufacturing process shown in Fig. 1.

Typically, the process throughput varies depending on the production schedule and on a sudden or scheduled maintenance for the unit processes. Thus, it is hard to directly reduce the variability in the process throughput. Since a change in the process throughput would inevitably make the 4-CBA concentrations in the oxidation and digestion processes deviate from the target values, we have to look for key process variables to adjust for the purpose of regulating the 4-CBA concentrations. Such variables should fulfill the requirement that they not only have strong and consistent effect on the 4-CBA concentrations but also be easily adjustable without violating the operating conditions. Figs. 6 and 7 can be used to screen the candidates of the key process variables that satisfy the requirement. As a result, we chose the temperature of the oxidizer (V15) as the most suitable process variable for regulating the 4-CBA in the crude TPA in the oxidization process. For controlling the 4-CBA in the purified TPA in the digestion process, we selected the flow rate of the air into the digestion process (W7) and the temperature of the 1st digester (W13).

If there is a change in the process throughput, we have to adjust the set-point of the oxidizer temperature to prevent the 4-CBA concentration of crude TPA from drifting from a target value. In this case, the following empirical relation can be used to set up a new set-point of the temperature:

$$C_{4CBA,\,CTA}=\alpha_1 F_f+\alpha_2 T_{OXD}+\alpha_3 \tag{4}$$

Similarly, the 4-CBA concentration of purified TPA can also be regulated by adjusting the set-points of both the air flow rate and the digester temperature, based on the following equation:

$$C_{4CBA,\,PTA}=\beta_1 F_f+\beta_2 T_{DIG}+\beta_3 F_{aDIG}+\beta_4 \tag{5}$$

In the equations above, the model coefficients (α_1, α_2, α_3 and β_1, β_2, β_3, β_4) are determined by designing suitable experiments [Coleman and Montgomery, 1993; Montgomery, 2001] for the oxidation and digestion processes. However, the determined coefficients are only valid within the normal operating regions where the experiments are performed.

CONCLUSIONS

PLS methods were applied to an industrial TPA manufacturing process to investigate the major causes of variability in the major quality variable (4-CBA concentration in the TPA product). Multivariate statistical analyses were performed by using the PLS models built from historical data measured on the process and quality variables in the TPA manufacturing process. The major causes of variability in the product quality were successfully found from the PLS analyses, and the effects of eliminating these causes were investigated via the simulations using the derived PLS models. Practical approaches to getting rid of the major causes, which were revealed through the PLS analyses, were proposed to design control systems for stabilizing the catalyst concentrations and to set up operating conditions on the basis of a simple linear equation when there are changes in the process throughput of the TPA manufacturing process.

ACKNOWLEDGMENT

This work was supported in part by the Brain Korea 21 Project and the Ministry of Commerce, Industry and Energy (MOICE) (grant No. IMT2000-00015993).

NOMENCLATURE

B : k×m matrix of regression coefficients
$C_{4CBA,\,CTA}$: 4-CBA concentration of crude TPA
$C_{4CBA,\,PTA}$: 4-CBA concentration of purified TPA
E : n×k residual matrix for **X**
F : n×m residual matrix for **Y**
F_{aDIG} : flow rate of the air entering the digestion process
F_f : throughput of *p*-xylene feed
P : k×A loading matrix for **X**
$\mathbf{p}_a$: loading vector with the size of k×1 for **X** in a latent variable a
Q : m×A loading matrix for **Y**
$\mathbf{q}_a$: loading vector with the size of m×1 for **Y** in a latent variable a
T : n×A score matrix for **X**
T_{DIG} : digester temperature
T_{OXD} : oxidizer temperature
$\mathbf{t}_a$: score vector with the size of n×1 for **X** in a latent variable a
U : n×A score matrix for **Y**
$\mathbf{u}_a$: score vector with the size of n×1 for **Y** in a latent variable a
W : k×A matrix of weight vector
$\mathbf{w}_a$: vector of weights with the size of k×1 for **X** variables in a latent variable a
X : matrix of process data with the size of n×k
Y : matrix of quality data with the size of n×m

Subscripts

A : total number of latent variables

a	: index of latent variable
k	: number of process variables for **X**
l	: number of sampling times for the quality variables
m	: number of quality variables for **Y**
n	: number of sampling times for **X** or **Y**
s	: number of rows selected in an original data matrix for rearrangement to a final data matrix

Greek Letters

α_1, α_2, α_3 : model coefficients of Eq. (4)

β_1, β_2, β_3, β_4 : model coefficients of Eq. (5)

REFERENCES

Cincotti, A., Orru, R. and Cao, G., "Kinetics and Related Engineering Aspects of Catalyst Liquid-Phase Oxidation of *p*-Xylene to Terephthalic Acid," *Catalyst Today*, **52**, 331 (1999).

Coleman, D. E. and Montgomery, D. C., "A Systematic Approach to Planning for a Designed Industrial Experiment," *Technometrics*, **35**, 1 (1993).

Dayal, B. S. and MacGregor, J. F., "Identification of Finite Impulse Response Models: Methods and Robustness Issues," *Ind. Eng. Chem. Res.*, **35**, 4078 (1996).

Eriksson, L., Hermens, J. L. M., Johansson, E., Verhaar, H. J. M. and Wold, S., "Multivariate Analysis of Aquatic Toxicity Data with PLS," *Aquatic Sciences*, **57**, 217 (1995).

Fujii, H., Lakshminarayanan, S. and Shah, S. L., "Application of the PLS Technique to the Estimation of Distillation Tower Top Composition," Preprint of IFAC Symposium on Advanced Control of Chemical Processes, 529 (1997).

Geladi, P. and Kowalski, B. R., "Partial Least-Squares Regression: A Tutorial," *Analytica Chimica Acta*, **185**, 1 (1986).

Han, I.-S. and Han, C., "Modeling of Multistage Air-Compression Systems in Chemical Processes," *Ind. Eng. Chem. Res.*, **42**, 2209 (2003).

Hong, S. J., Hua, C. K. and Han, C., "Local Composition Soft Sensor in a Distillation Column using PLS," *HWAHAK KONGHAK*, **37**, 445 (1999).

Hur, S. M., Park, M. J. and Rhee, H. K., "Polymer Property Control in a Continuous Styrene Polymerization Reactor Using Model-on-Demand Predictive Controller," *Korean J. Chem. Eng.*, **20**, 14 (2003).

Jaisinghani, R., Sims, R. and Lamshing, W., "APC Improves TA/PTA Plant Profits," *Hydrocarbon Processing*, **Oct**., 99 (1997).

Kim, J. Y., Kim, H. Y. and Yeo, Y. K., "Identification of Kinetics of Direct Esterification Reactions for PET Synthesis Based on a Genetic Algorithm," *Korean J. Chem. Eng.*, **18**, 432 (2001).

Kroschwitz, J. I.,"Encyclopedia of Chemical Technology," John Wiley & Sons, New York, USA (1991).

Liu, J., Min, K., Han, C. and Chang, K. S., "Robust Nonlinear PLS Based on Neural Networks and Application to Composition Estimator for High-Purity Distillation Colums," *Korean J. Chem. Eng.*, **17**, 184 (2000).

MacGregor, J. F., Jaeckle, C., Kiparissides, C. and Koutoudi, M., "Process Monitoring and Diagnosis by Multiblock PLS Methods," *AIChE J.*, **40**, 826 (1994).

MacGregor, J. F. and Kourti, T., "Statistical Process Control of Multivariate Processes," *Control Eng. Practice*, **3**, 403 (1995).

Montgomery, D. C., "Introduction to Statistical Quality Control," John Wiley & Sons, New York (2001)

Neogi, D. and Schlags, C. E., "Multivariate Statistical Analysis of an Emulsion Batch Process," *Ind. Eng. Chem. Res.*, **37**, 3971 (1998).

Shi, R. and MacGregor, J. F., "Modeling of Dynamic Systems using Latent Variable and Subspace Methods," *J. Chemometrics*, **14**, 423 (2000).

Wise, B. M. and Gallagher, N. B., "The Process Chemometrics Approach to Process Monitoring and Fault Detection," *J. Proc. Cont.*, **6**, 329 (1996).

Wold, S., Esbensen, K. and Geladi, P., "Principal Component Analysis," *Chemometrics and Intelligent Laboratory Systems*, **2**, 37 (1987).