AIChE

Robust Processes Through Latent Variable Modeling and Optimization

Francois Yacoub and John F. MacGregor
McMaster Advanced Control Consortium, Dept. of Chemical Engineering, McMaster University, Hamilton, ON, Canada

DOI 10.1002/aic.12352
Published online August 16, 2010 in Wiley Online Library (wileyonlinelibrary.com).

A data-based approach for developing robust processes is presented and illustrated with an application to an industrial membrane manufacturing process. Using historical process data, principal component analysis and partial least squares are used to extract models of the process and of the sensitivities of the process to various disturbances, including raw material variations, environmental conditions, and process equipment differences. Robustness measures are presented to quantify the robustness of the process to each of these disturbances. The process is then made robust (insensitive) to the disturbances over which one has some control (e.g., by modifying the equipment units to which the process is sensitive and imposing specification regions on sensitive raw materials). It is also made robust to disturbances over which one has little control (e.g., environmental variations) by optimizing the process operating conditions with respect to performance and robustness measures. The optimization is easily performed in the low-dimensional space of the latent variables even though the number of process variables involved is very large. After applying the methodology to historical data from the membrane manufacturing process, results from several months of subsequent operation are used to demonstrate the large improvement achieved in the robustness of the process. © 2010 American Institute of Chemical Engineers *AIChE J,* 57: 1278–1287, 2011
Keywords: robustness, latent variables, PCA, PLS, optimization, disturbances, membranes

Introduction

In general, robustness can be defined as the ability of a system to maintain performance in the face of disturbances and uncertainties that are likely to occur in practice. In biological systems, cellular functions are maintained in the presence of diverse disturbances arising from environmental changes. It has long been recognized that this robustness is an inherent property of all biological systems and is strongly favored by evolution.

Robustness is a relative property because no system can maintain robustness for all its functions in the face of any kind of perturbation. Hence, we have to specify (i) which characteristic behaviors should remain unchanged and (ii) for what type of disturbances or uncertainties this invariance property should hold. In biological systems, the primary function of a system is usually robust to a wide range of perturbations; yet, for some functions these systems can show extreme fragility toward other (even seemingly much smaller) perturbations. This coexistence of extremes in robustness and fragility, ‘‘robust yet fragile’’ in effect constitutes a very fundamental phenomenon that needs to be explored and clearly defined for any system.

Industrial processes share the same type of challenges as biological systems. Processes are subjected to external disturbances in the form of raw material variations, the presence of impurities, environmental changes and internal disturbances such as changing catalyst activities, and differences in equipment. The

Correspondence concerning this article should be addressed to J. F. MacGregor at macgreg@mcmaster.ca.

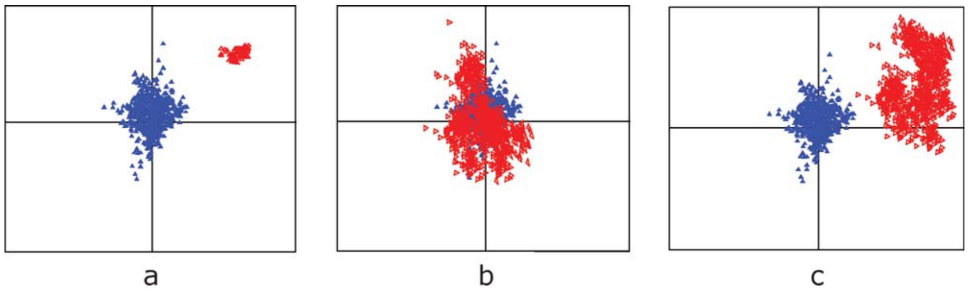

Figure 1. Effects of different disturbances on the quality variables.

(a) Shift in the mean of the distribution; (b) increase in the variance; and (c) shift in the mean and increase in the variance. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

major quality attributes of the manufactured products need to be robust to the most important of these disturbances. Clearly, these industrial systems, like the biological systems discussed above, have feedback controllers that make them more robust. However, they may still be fragile to certain disturbances. The reality of "robust yet fragile" processes is that one should try to ensure that the important product quality attributes are insensitive to certain key disturbances, but at the same time one should also develop specifications regions to limit the range of other measured disturbances to which the process is sensitive less these disturbances become too large.

The methodology for achieving robust industrial processes proposed in this article is inspired by the robust performance of living biological cells, often perceived as small factories, and by the pioneering works on robustness by Box and Andersen[1,2] and Taguchi.[3–6] It is based on learning from process data, quantifying robustness and optimizing the process operation conditions to achieve both acceptable performance and a high degree of robustness to important disturbances. The methodology is entirely data driven and uses latent variable (LV) models and optimization in the identified LV space.

Measures of Robustness

The lack of robustness of processes to disturbances can be observed as a shift in the mean of a response, as an increase in the variability around the mean, or both, as illustrated in Figure 1. To develop a quantity that reflects the degree of robustness, one can formulate this problem as a binary classification with one class representing the process with acceptable performance under common-cause variation (***c***) and the second class representing the process under the additional effect of a disturbance (***d***). The deviation or shift in the mean of the distribution of the responses can then be calculated as a squared distance $D^2_{d/c}$ between the two classes and a shift in relative variance as $\mathrm{VR}_{d/c}$ (both defined below).

Structure of the data

In this article, it is assumed that historical data are available on the process when subject to a range of disturbances against which robustness is desired. The data matrix consists of measured values ($\mathbf{i}$ = 1,…,$\mathbf{N}$) on process variables $\boldsymbol{x}_{ik}$ (k = 1,…,K) and quality variables $\boldsymbol{y}_{im}$ (m = 1,…,M) as illustrated in Figure 2. The data are arranged in the observation matrices $\mathbf{X}$ with dimensions ($N*K$) and $\mathbf{Y}$ with dimensions ($N*M$). The observations made on a single object, at a given point in time (x_i), form a K-dimensional vector and can consequently be represented as a point in a K-dimensional space. Data are assumed to be available for both conditions of acceptable performance when only common-cause variation[7] (inherent variation with no abnormal disturbances present) and under the influence of different disturbances (Figure 2). This type of data is almost universally available on processes. As the quality of any product cannot be evaluated univariately, the objective then is to assess the similarity or dissimilarity of objects as in Figure 1 in a multivariate sense.

Dissimilarity measures for disturbance robustness

Dissimilarity in the Means. As illustrated in Figure 1, lack of robustness to a disturbance can be manifested as a shift in the mean of the process. To quantify such shifts, a measure of dissimilarity is needed. In the present approach, a model is first build from data from the class when acceptable performance is achieved and only "common-cause" variation is present. The data within this class can be described

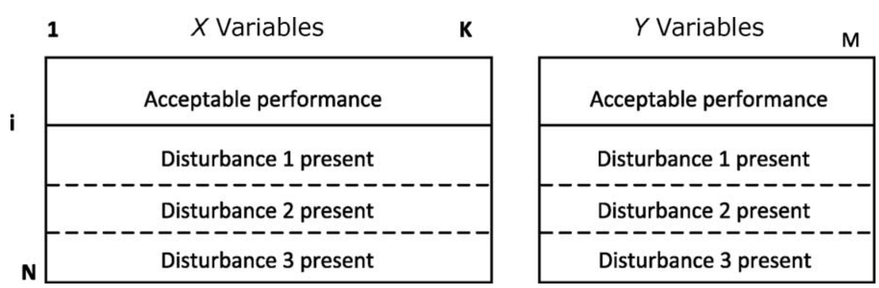

Figure 2. Data matrix structure for the classification problem.

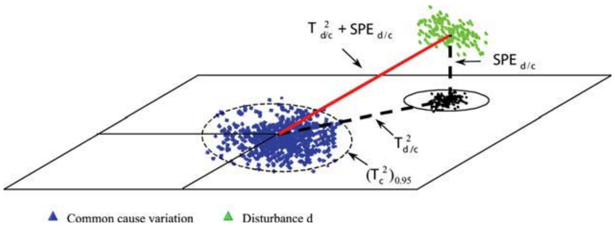

Figure 3. Latent variable distance measures ($T^2_{d/c}$ and SPE$_{d/c}$) of a disturbance class of observations (*d*) from an acceptable class (*c*) of observations.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

by the LV model built using principal component analysis (PCA)[8,9]:

$$\mathrm{Y_c} = \mathrm{T_c P_c^T} + \mathrm{E_c} \tag{1}$$

where $\mathbf{Y_c}$ is a matrix containing observations on the response variables of interest collected under that class, $\mathbf{T_c}$ is a ($N*A$) matrix of LV scores ($A < M$) that summarizes the $\mathbf{Y_c}$ matrix, $\mathbf{P_c}$ is a matrix of loadings, and $\mathbf{E_c}$ is a matrix of residuals. This model will be referred to as $\mathrm{PCA_{Yc}}$.

The dissimilarity function is constructed based on the SIMCA method[10] for classification in which a series of PCA models are built, one for each class of data shown in Figure 2. A set of observations under a given disturbance can then be accessed against the reference data class (***c***) where there is acceptable operation and only "common-cause" variation by using a distance measure as a criterion for association. For a set of observations under the influence of a disturbance, a multivariate distance measure, $D^2_{Yd/c}$, for the distance of the responses (Y_d) under the disturbance class (***d***) from the responses (Y_c) under the reference class (***c***) is defined as described in Eq. 2. The distance $D^2_{Yd/c}$ provides equal weighting to the distance in the LV space of model (Hotelling's T^2) and to the distance in the orthogonal residual space (squared prediction error or SPE).[10,11] Similar combined distance measures have also been used in other contexts.[12,13] A graphic illustration of the concept of the distance measure is shown in Figure 3. The 95% confidence limits for class ***c*** in Eq. 2 can be seen as normalization factors for the two terms.

$$D^2_{Y,d/c} = \frac{T^2_{d/c}}{(T^2_c)_{0.05}} + \frac{\mathrm{SPE_{d/c}}}{(\mathrm{SPE}_c)_{0.05}}. \tag{2}$$

In Eq. 2, overall the square prediction error (SPE$_{d/c}$) of class ***d*** observations from the model built on class ***c*** ($\mathrm{PCA_{Yc}}$) is calculated as shown in Eqs. 3 and 4. For the i-th observation or object of class ***d***, SPE$_{d/c,i}$ can be defined as:

$$\mathrm{SPE}_{d/c,i} = \sum_{m=1}^{M} e^2_{d,m,i}, \tag{3}$$

where the $e_{d,m,i}$ are the components of the residual vector $(y_{d,m,i} - \hat{y}^c_{d,m,i})$ for observations under disturbance ***d*** relative to the model for the acceptable performance class ***c***.

The notation d/c implies the point in class ***d*** (under the influence of the disturbance) is projected onto the model for class ***c***.

The average of SPE$_{d/c,i}$ for the n_d observations in class ***d*** can be computed as:

$$\mathrm{SPE}_{d/c} = {}^1\!/_{n_d} \sum_{i=1}^{n_d} \mathrm{SPE}_{d/c,i}, \tag{4}$$

where SPE$_{d/c}$ is the average squared distance of the observations in class ***d*** from the plane defined by the PCA model for the reference or "common-cause" class ***c*** and $(\mathrm{SPE}_c)_{0.95}$ is the 95% confidence limit for the residuals of the samples of the PCA model built on class ***c***.[9,11]

The average Hotelling's distance for observations in class ***d*** from the center of the common-cause model $T^2_{d/c}$ is calculated as in Eq. 5.

$$T^2_{d/c} = \frac{1}{n_d}\sum_{i=1}^{n_d} T^2_{d/c,i} = {}^1\!/_{n_d} \sum_{i=1}^{n_d}\sum_{a=1}^{A} \frac{t^2_{d/c,a}}{s^2_{t_a}}, \tag{5}$$

where $T^2_{d/c,i}$ is Hotelling's T^2 statistic for the i-th sample in class ***d*** based on the model for class ***c***. $(T^2_c)_{0.95}$ is the 95% confidence limit of the Hotelling T^2 for the $\mathrm{PCA_{Yc}}$ model[9,11] built on class ***c***. s^2_{ta} is the variance of the principal component score, t_a, according to the $\mathrm{PCA_{Yc}}$ model of class ***c***.

Intuitively, $D^2_{Yd/c}$ in Eq. 2 provides a quantitative measure of process robustness to a shift in mean of the response variables (Y). From Figure 3, this can be seen as the combined "deviation in the LV space" $T^2_{d/c}$ and the "deviation from the plane" defined by the SPE$_{d/c}$ caused by the presence of a disturbance. A small value of $D^2_{Yd/c}$ is an indication of a robust process. Based on the normalization used, $D^2_{Yd/c}$ values much above 2.0 would tend to imply lack of robustness to that disturbance class ***d***.

However, the distance measure in the response space (**Y**) will also depend upon the magnitude of the disturbance. To normalize the distance measure with respect to the magnitude of the disturbance, one can also compute, in a similar fashion to $D^2_{Yd/c}$, the distance $D^2_{X_d d/c}$ that the disturbance variables (X_d) have moved during that disturbance.

A normalized measure of robustness to any disturbance or group of disturbances can then be defined by Eq. (6)

$$\text{Measure of robustness} = \bar{D}^2_{Yd/c} = \frac{D^2_{Y_{d/c}}}{D^2_{X_d d/c}}. \tag{6}$$

Dissimilarity in the Variance. As illustrated in Figure 1, a lack of robustness of a process to a disturbance may also manifest itself as an increase in variance of the responses. A dissimilarity measure is therefore also needed to measure this change in variance. One can compute the variance–covariance matrices for the data with operation under both the acceptable (c) and disturbance (d) conditions. For $\mathbf{n}_d$ observations ($\boldsymbol{y}_d$) in the disturbance class (with $\boldsymbol{y_d}$ scaled in the same way as the preferred class, c), the variance–covariance matrix for the M y's is computed as

$$\textstyle\sum_d^2 = \frac{1}{n_d}\sum_{i=1}^{n_d} (y_{d,i} - \bar{y}_d)(y_{d,i} - \bar{y}_d)^T. \tag{7}$$

One can then compute the trace $\mathrm{Tr}(\sum_d^2)$ (the sum of the diagonal variances of $\sum_d^2$) and compare this against the $\mathrm{Tr}(\sum_d^2)$ of the variance–covariance matrix under the preferred operation with only common-cause variation (c). However, the data collected when the process is operating under the preferred mode are usually much more extensive and should contain most common-cause disturbances over time, whereas the disturbance, class $\boldsymbol{d}$, is a more limited set of data collected over a limited duration where only some of the common-cause variations may be present. To deal with this problem, one can divide the preferred operation dataset to a number of subgroups (G) of n_d successive observations (where n_d is the size of the disturbance class) and hence provide an estimate of the equivalent short-term variance in the reference class. The ratio between the variance under preferred operation to the variance under the influence of disturbance can then be calculated as:

$$\text{Variance ratio} = \mathrm{VR}_{d/c} = \frac{\mathrm{Tr}\left(\sum_d^2\right)}{\frac{1}{G}\sum_{j=1}^{G}\mathrm{Tr}\left(\sum_{c,j}^2\right)}, \tag{8}$$

where the denominator is the average trace of the variance—covariance matrix for the G sets of n_d successive observations taken from the preferred class $\boldsymbol{c}$ of observations.

This value, being a ratio of mean squares should follow approximately an F distribution with numerator degrees of freedom equal to $K(n_d - 1)$ and the denominator with degrees of freedom equal to $\mathrm{GK}(n_d - 1)$. If these degrees of freedom are relatively large, then a variance ratio (8) much above ~ 2 would indicate a significant change in variance due to the disturbance.

The measures of robustness (6) and (8) are used in this work to provide relative measures of robustness to different disturbances. The magnitude guidelines are only provided as rules of thumb to assess the importance of any disturbance.

Optimization for Robustness

Many issues have to be taken into account when formulating the robust design problem as an optimization problem. We have to select first the final product quality or response variables (Y) for which the process must be robust, and the design variables or system parameters (x_p) that can be manipulated in order to achieve robust performance, define robust performance in the form of an objective function, and develop models that can define both the responses and their sensitivities in terms of the manipulated variables, raw materials, and disturbance variations. To achieve a robust process, one has to simultaneously achieve the following two goals: (i) find in the X-space a new set of process conditions (**x**p) that will make the Y-space insensitive to specified disturbances; (ii) identify the disturbances that can create a "fragile" condition (i.e., raw materials properties) and properly set specification limits on these disturbances.

This section discusses the development of LV regression models relating the response variables Y to the process variables X, the extraction of sensitivity estimates of the responses to the disturbances of interest, and the optimization in the space of the LVs to obtain operating conditions and raw material specifications that ensure robust performance. The whole procedure is based on the off-line analysis of historical data and is not being suggested as an on-line control or real-time optimization problem.

Partial least squares regression model

Nonlinear LV models built using partial least squares (PLS)[14] are used to relate the key product responses of interest (Y) to the manipulated variables (X_p), raw material variables (X_m), and all measured process variables (all of these are combined in X). The LV model provides models for both the X space and Y space and the relationship between these spaces.

$$\begin{aligned}\hat{\mathbf{X}} &= \mathbf{T_A P_A^T}\\ \hat{\mathbf{Y}} &= \mathbf{T_A C_A^T}\end{aligned}. \tag{9}$$

$\mathbf{T_A}$ is a matrix of LV scores, where (a = 1, …, A) are estimated as linear combinations of the $\mathbf{X}$ matrix

$$\mathbf{T_A} = \mathbf{XW}^*, \tag{10}$$

where the columns of $\mathbf{W}$* are the weights from PLS that maximize the covariance of (X Y).

The LV score matrix can be re-expressed in normalized form as

$$\mathbf{T_A} = \mathbf{U_A}.\sum\nolimits_A, \tag{11}$$

$\mathbf{U_A}$ is a matrix with orthonormal columns and $\sum_A$ is a diagonal matrix containing the standard deviations of the t_a scores, (s_a) (a = 1,2,..A) on the diagonals.

Using a quadratic PLS algorithm,[14] the predicted response variables can be expressed as

$$\hat{y}_j(u_{\text{new}}) = \sum_{a=1}^{A}\left(b_{0a}+b_{1a}(u_{\text{new,a}}\cdot s_a) + b_{2a}(u_{\text{new,a}}\cdot s_a)^2\right).c_{ja}, \tag{12}$$

where the c_{ja}s are the PLS weights for the j-th response (Y) in the a-th LV (t_a) and the bs are the regression coefficients for the quadratic relationship between y and the LVs of the X space ($t_a = u_a.s_a$). Although a quadratic inner relation was used to model the nonlinearities, any other form of nonlinearity can be introduced in a similar way.[14] This model will be referred to as the NPLS model.

Objective function for robust performance

The objective of the optimization problem is to find new process conditions $\hat{x}_{p,\text{new}}$ and raw material specifications $x_{m,\text{new}}$ that satisfy several conditions: (i) they are consistent with the PLS model from past operating conditions; (ii) they yield desired values for the quality variables y_{des}; and (iii) they minimize the sensitivity of the quality variables to disturbances $\frac{\partial y}{\partial d}$. Here, the partial derivative symbol is used to denote the sensitivity, but, as shown below in the application of the methodology, these sensitivities are estimated as coefficients of a PLS model fitted to specific subsets of historical data. The optimization is thus formulated to minimize the following quadratic objective function.

$$\underset{\substack{\mathbf{u_{new}},a\\ a=1,2,\ldots A}}{\mathbf{Min}}\left[(y_{\text{des}} - \hat{y}(u_{\text{new}}))^T Q_1 (y_{\text{des}} - \hat{y}(u_{\text{new}})) + \left(\frac{\partial y^T}{\partial d} Q_2 \frac{\partial y}{\partial d}\right)\right], \tag{13}$$

where y_{des} is a vector of desired values of final product quality characteristics, and $\frac{\partial y}{\partial d}$ is the vector of sensitivities of the product qualities with respect to specified disturbances. The first term in Eq. 13 minimizes the deviation from the desired product quality and the second term penalizes the effect of disturbances on the quality. This objective thus attempts to simultaneously achieve conditions that will yield good performance and robustness (robust performance). ***Q*1** and ***Q*2** are selected weighting matrices that balance the emphasis placed on performance vs. robustness (sensitivity to disturbances). In this context, to increase the robustness of a process, one would increase the weight $\boldsymbol{Q_2}$ relative to $\boldsymbol{Q_1}$.

To ensure that the solution to the optimization u_{new} will be feasible, i.e., lie on the plane defining the X space and lie in the region of that plane spanned by past data. The following three sets of constraints are applied during the optimization.

i The projected solution t_{new},a ($a = 1, 2, ..A$) must fall within the region of historical t-vectors that is.

$$T^2 = \left(\frac{t_1^2}{s_1^2} + \frac{t_2^2}{s_2^2} + ...\right) \leq \text{constant} \tag{14}$$

The value of the constant can be taken as the historical 99% confidence contour[9,11] for data from the preferred process.

ii To be consistent with past operating conditions, the solution $x$new must lie close to the model plane and so must have a small residual. This implies the constraint

$$\mathbf{SPE} = \sum_{l=1}^{k} (x_{\text{desired},i} - \hat{x}_{\text{desired},i})^2 \leq e \tag{15}$$

where, **e** is a small value.

iii Additional constraints may also be added for all the disturbances variables that are measured and hence known to within measurement error. In the optimization, these are subjected to bounded constraints.

$$\text{X}_{\text{measured},i} - \varepsilon_i \leq \hat{\text{X}}_{\text{new},i} \leq \text{X}_{\text{measured},i} + \varepsilon_i \tag{16}$$

Values of εi are determined from the estimates of measurement errors.

The optimization Eq. 13 is solved using trust-region method of the sequential quadratic programing (SQP) algorithm.[15] Note that the optimization in (13) is performed in the LV space $\mathbf{u_a}$ ($\mathbf{a = 1,2,..A}$) of the PLS model. This is essential because this is the only space (i.e., the plane of the LV model) in which one has causal information on the process.[16] Once the optimal values of $u_{new,a}$ have been obtained, the corresponding optimal values of the process variables can be obtained from the PLS model of the X-space (9) as:

$$\begin{array}{cccc} \hat{\text{x}}_{\text{new}}^{\text{T}} = & \hat{\text{u}}_{\text{new}}^{\text{T}} \cdot & \sum_A & \text{P}^{\text{T}} \\ 1 \times k & 1 \times A & (A \times A) & A \times k\ . \end{array} \tag{17}$$

This $\hat{\text{x}}_{\text{new}}^{T}$ will (by Eq. 17) fall on the plane of the LV model and in the region of existing data and hence will provide a feasible solution, consistent with past operation. Equation 17 provides the vector of process conditions ($\hat{\text{x}}_{\text{new}}^{T}$), consisting of both the new values for the manipulated variables (x_{p}) as well as the estimated values that all the other process variables will have at these optimal conditions.

Sensitivity functions

To provide an expression for the second term in Eq. 13 that describes the sensitivities of the final product quality to specified disturbances, one needs to extract this information from the historical data. This is a difficult problem, and we present a procedure for estimating these sensitivities as a function of the LVs. It is accomplished in three steps:

i Using only process manipulated variables (x_{p}) (excluding all disturbances), build a PCA model (PCA_{Xp}) that cluster all the observations into regions that have a similar operational mode. As the observations within each cluster have similar process conditions, one can assume that the variation within each cluster is due mainly to changes in disturbances.

ii Extract all the observations from each cluster and build a regression model, one for each cluster, relating the disturbances (d) to the responses (y_j) ($\text{REG}_{y/d}$). The regression coefficient (β_k) relating the effect of each disturbance (d_k) on y_{j} is an estimate of the sensitivity of y_j to that disturbance ($\frac{\partial y_j}{\partial \text{d}_k}$) for each cluster. The sensitivity estimates for each cluster are then assigned to all observations in that cluster.

iii A regression model ($\text{REG}_{\text{u}/\beta}$) is then built to relate the sensitivities from all clusters to the LVs (u_{a}, $a = 1,..,A$) of the NPLS model as required in the optimization of (13).

This procedure of estimating the sensitivities from historical data is a key element of this robust problem. However, further details on it are best illustrated with real process data and so are left to the industrial example section to follow.

Achieving Robust Performance in a Membrane Manufacturing Process

This industrial hollow fiber membrane manufacturing process has been described in another article on the diagnosis of equipment problems.[17] Therefore, it will be described only briefly here. In that article,[17] it was shown by an indicator variable PLS approach that the major sources of variations in product quality in this complex series-parallel process were equipment-to-equipment variations, raw material variations, and environmental variations (ambient temperature and humidity). In this section, we will show again, but by means of the dissimilarity measures in the section on Measures of Robustness, that the process is not generally robust over its entire range of operation to these disturbances (equipment-to-equipment variations will be broadly interpreted as disturbances in this article). The LV modeling and optimization procedures outlined above will be used to find a region of the operating space where the quality responses are much more robust to these disturbances. Actual data from several months of operation of the robustified process will then be used to verify the success of these methods.

The process for the manufacturing of hollow fiber membranes used in water purification consists of a complex series-parallel arrangement of processing units as illustrated in Figure 4.

Batches of membrane are produced by processing raw materials (polymer, solvent, additives) through any 1 of 11 batch mixing units, followed by passing the dope through

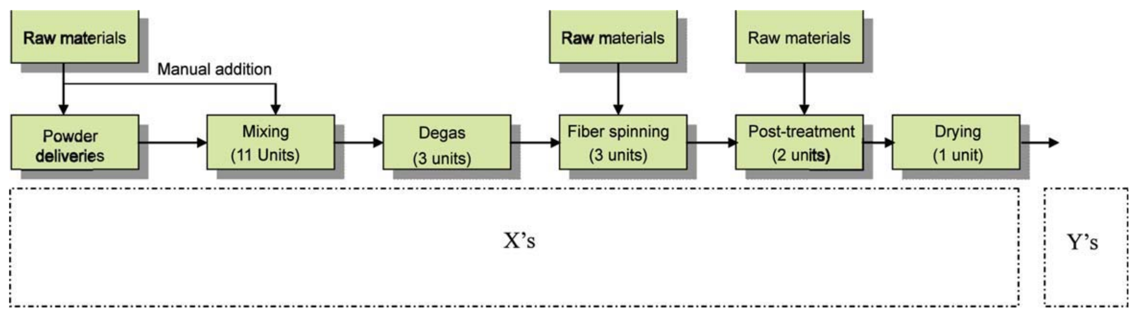

Figure 4. Schematic diagram of the overall membrane manufacturing process showing the structure of the data.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

any one of the three degassing units, spinning the hollow fibers in one of three spin-lines, post-treating in one of two post-treatment units, followed by drying in the single drier unit. Almost 3 years of data were available consisting of 200 batches of membrane. Final assessment of each of the batches of membrane was made in a quality control laboratory by measuring 12 responses related to the structure and permeation properties of the resulting hollow fiber membranes (**Y**s). A plot of the first two principal components (t_1, t_2) that explain most of the variation in all the historical quality data (**Y**s) is shown in Figure 5. One can see that there is a considerable variation in quality, and some of the areas in the score plot where certain problems occur are noted qualitatively by ellipses. (The colors in the plot show different periods of operation and are shown only to help visualization.) The problem is to find out what disturbances and equipment variations are leading to this sensitivity or lack of robustness of the existing process and then define conditions and an operating region under which one can achieve robust performance.

Data were available on measurements from variables characterizing the raw materials (polymers and solvents), on initial conditions and process trajectory variables for the batch mixers, on the degassing unit variables, on fiber spinning variables (polymer delivery system, bore fluid properties, spin jets and gap control, coagulation, and rinse tank), on post-treatment variables, and on dryer variables. In total, for the manufacture of each batch of membrane, 8025 measurements were collected on process and raw material variables.

Analysis of robustness to disturbances

The measures of robustness $\overline{D}^2_{Yd/c}$ (Eq. 6) and variance ratio $VR_{d/c}$ (Eq. 8) were calculated and summarized in Table 1 for all raw materials (two sources), all process units, and all environmental conditions for which clear changes had occurred. The measure of robustness $\overline{D}^2_{Yd/c}$ is calculated as described by Eq. 6 by projecting observations under the influence of disturbances (class d) (i.e., a change to a different polymer raw material source, equipment changes (mixers, powder delivery system, and degas system), and environmental conditions (relative humidity and ambient temperature) onto the PCA model (PCA_{Yc}) built using data from the desired space at the center of the score plot in Figure 6 with only "common-cause variation" (class c). The computed robustness measures are shown in Table 1 for a set of equipment units and disturbances.

The process units and disturbances that exhibit high dissimilarity distance ($\overline{D}^2_{Yd/c}$) and/or high variance ratio ($VR_{d/c}$) are shown in the top part of Table 1. This implies that the process is not robust to the use of mixers 1, 7, and 8 to spin-line 2 and degas system 2 nor to using the alternative raw material (polymer cluster 1) nor to changes in the relative humidity and temperature of the environment (these latter two were correlated). Most of the other equipment units and disturbances had relatively low dissimilarity distances and variance ratios; a few of these are shown for comparison in the bottom part of Table 1. The process therefore appears robust to these other equipment units and disturbances.

Optimization for robust performance

The objective is to search for a good operational space that is also robust to the key disturbances revealed by the robustness measures of the previous section. It was shown in Table 1 that the process was not robust to many disturbances. These could be classified into those disturbances over which the company had little control (relative humidity and ambient temperature) and those over which some control was possible (source of polymeric materials and equipment units that could be used). It was therefore decided to robustify the process against the ambient disturbances (humidity and temperature) through optimization of Eq. 13 using this as disturbance (d) and to set specifications on the other disturbances (polymer material properties and equipment units to employ) to which the process would still be "fragile."

Latent Variable Model. A nonlinear (quadratic) PLS model (NPLS – Eqs. 9–12) was built for historical data between the final product quality variables (Y) and the process variables (X) from all units including: the initial conditions and trajectories in the batch mixing, the degas variables, the spin-line variables, and the post-treatment variables. The model used historical data that spanned all conditions of the environmental disturbances (environmental humidity and temperature), but these environmental variables were not included in the model because the eventual use of the model is to find process conditions at which the process is insensitive to any changes in these disturbances. The nonlinear PLS model provided a good fit of the data (the fraction of the sum of squares explained by the fit R^2 was 0.92) and was quite predictive of new data (the fraction of the prediction error sum of squares from cross-validation Q^2 was 0.88). Linear PLS gave a slightly lower R^2 and Q^2 but could equally well have been used in this method.

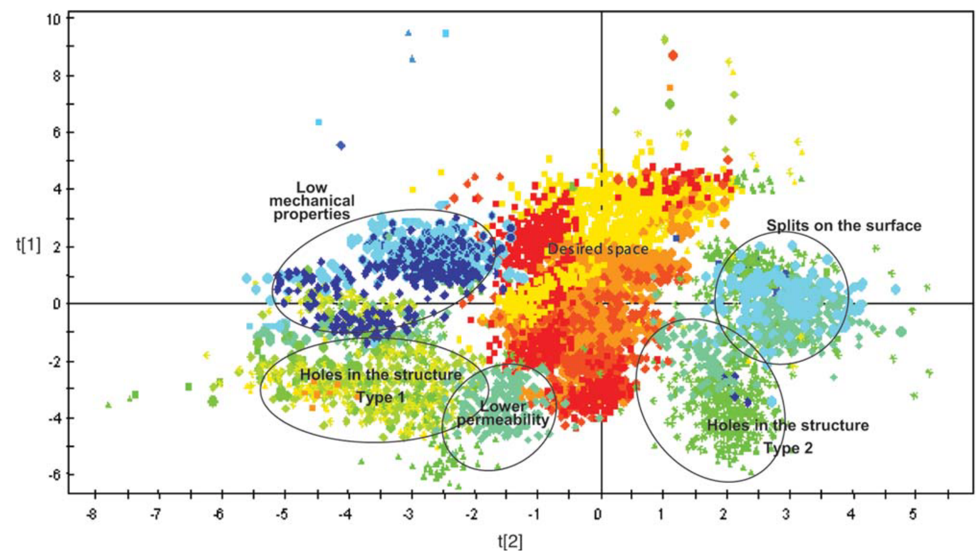

Figure 5. Variation of the final quality variables in the PCA score space of Y.

Results shown for all historical data on Y. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Sensitivities to Disturbances. As discussed in section Optimization for Robustness, crude estimates of the sensitivities of the *Y*s to the major disturbances can be obtained by a three-step procedure. First, a PCA model is fitted to all the historical process manipulated variable data (PCA_{Xp}). This will extract the major variations related to process operating data and cluster the observations in the data set into clusters with similar operating values. Figure 6 shows five clusters in the score space of the first two principal components (t_1, t_2). A regression model is then built on the data within each cluster relating the product quality variables (Y) to the main disturbance variables for which robustness is desired. (In this study, relative humidity and ambient temperature of the environment around the process were the disturbances considered.) As the observations within each cluster have similar process operating conditions, it is assumed that the main variations in the Y space are due to the effect of disturbances occurring within that cluster. The regression model ($\text{REG}_{y/d}$) is of the form:

$$Y_i = \beta_0 + \beta_1(\text{relative humidity})_i + \beta_2(\text{ambient temperature})_i \quad . \tag{18}$$

The coefficients in the regression model are then taken as estimates of the sensitivities of the disturbances common to all points within that cluster. The same sensitivities are assigned to each observation in that cluster. A similar procedure is followed with each of the other clusters. One is then left with a matrix of sensitivities for different regions in the LV space. These sensitivities are then regressed against the LV score vectors ($\mathbf{u}_a$, $a = 1,..,A$) from the PLS model to obtain a model for the variation of the sensitivities in the space of the NPLS LVs ($\text{REG}_{u/\beta}$). This provides a model for the sensitivities in the second term in the objective function (Eq. 13), while the NPLS model provides the $\hat{y}$ predictions in the first term, both of these as a function of the LVs ($\mathbf{u}_a$, $a = 1,..,A$).

Optimization. The optimization of the objective function (Eq. 13), subject to the constraints discussed in section

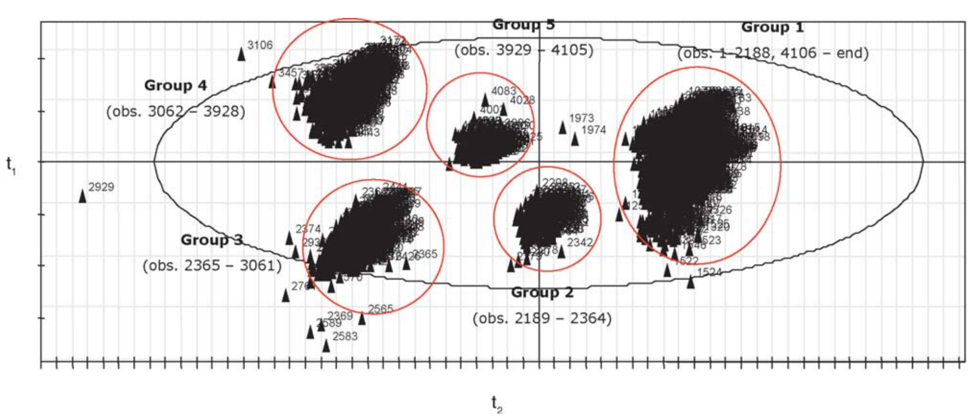

Figure 6. Score plot (t_1 vs. t_2) of the PCA_{Xp} model of the process conditions (X_p) using all historical data.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 1. Value of the Distance Measure ($\overline{D}^2_{Yd/c}$) Calculated as in Eq. 6 and the Variance Ratio ($VR_{d/c}$) Calculated as in Eq. 8 for Selected Disturbances

	$\overline{D}^2_{Yd/c}$	Variance Ratio $VR_{d/c}$
Disturbances		
Polymer (cluster 2)	8.70	6.30
Mixer 8	5.80	5.45
Spin-line 2	5.23	1.09
Mixer 7	5.70	5.55
Powder delivery system	7.18	1.20
Relative humidity	8.23	1.81
Ambient temperature	7.12	1.10
Mixer 1	6.10	4.65
Degas system 2	5.20	1.97
Examples of raw material and equipment that did not cause a shift in the mean or variance of the cluster distribution		
Polymer (cluster 1)	2.11	1.11
Mixer 11	2.02	1.32
Degas system 3	2.30	1.80

Optimization for Robust Performance was carried out for a range of weighting matrices ($\boldsymbol{Q_1}$ and $\boldsymbol{Q_2}$) in the objective function (Eq. 13). This provided a range of solutions shown in the LV score space in Figure 7, ranging from solution 1 that emphasizes performance only ($\boldsymbol{Q_2} = \mathbf{0}$) to increasing emphasis on robustness as $\boldsymbol{Q_2}$ increases in solutions 2 through 4.

By projecting the results back to the PCA space of the Y data, one can display the four solutions. Although solution 1 provides an optimum performance (high permeability, high BP, high mechanical properties, and low defect rate), it is sensitive to environmental changes as reflected in the sensitivity values $\frac{\partial y}{\partial d}$, where $\mathbf{y}$ is the defect rate (%), and $\mathbf{d}$ is the relative humidity and temperature disturbances. On the other hand, solution 4 yields operating conditions with less performance (82% of the optimal performance) but with a higher degree of robustness to environmental changes. Conditions corresponding to solution 3 were selected as those to be implemented on the process.

Reconstruction of the Optimal Desired Space. From the knowledge developed from Table 1, a multivariate raw material specification region[18] was established to use only raw materials with properties consistent with those from cluster 1. New process conditions that correspond to the optimized values of the LVs ($\mathbf{u}_{a,\mathrm{new}}$ $a = 1,\ldots,A$) are easily obtained from the LV model of the X space (Eq. 17). This $\hat{x}^T_{\mathrm{new}}$ vector contains the optimal steady-state setting for all the variables of the continuous process units (spin-line, post-treatment, and drying), for the fixed and initial conditions of the batch processes, and for the optimal time varying trajectories of the process variables in the batch process units (mixing). The batch trajectory data were unfolded in a batch-wise manner[19] and included in the rows of X. Figure 8 shows the optimal trajectory of the dope temperature for the mixing process (heavy red line) together with some typical trajectories from the historical data. One trajectory from each of mixers 7 and 8 is also shown. These are two of the units to which the process was found to be not robust (Table 1).

Manufacturing Results from Implementation of the New Robust Performance Conditions. The optimal robust performance conditions corresponding to the selected solution 3 in Figure 7 were implemented on the manufacturing process. In addition, a multivariate raw material specification region was developed and used to exclude polymer raw materials with properties significantly outside those consistent with polymer source 1. In addition, mixers 1, 7, and 8 were modified and introduced back into production after acceptable performance was demonstrated. The results of 80 new membrane batches produced under these new conditions are shown in Figure 9 as a projection of the new product quality results (Y) onto the LV space of the PCA model for Y based on the original data (Figure 5). The historical Y data from Figure 5 are also shown to illustrate the much more consistent (robust) operation of the process under these conditions. The process at these new conditions was still subject to essentially the same environmental variations as was the data from the historical process operation. Clearly, the process is much more robust under the new process conditions,

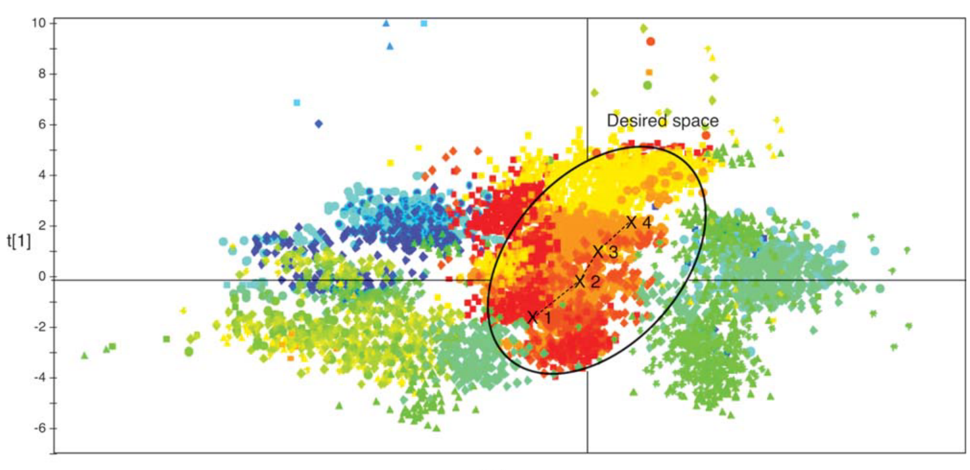

Figure 7. Projection of different optimization solutions into the latent variable space (t_1, t_2) of the PCA model for Y (see Figure 5).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

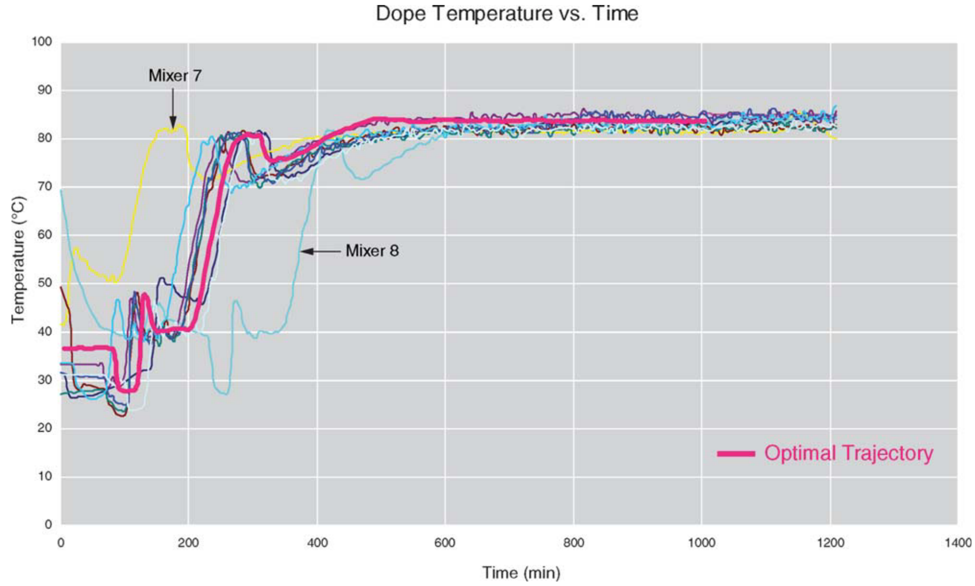

Figure 8. Trajectory of dope temperature (°C) vs. time (min) showing the optimal trajectory computed from Eq. 17 and several other trajectories from good and poorly performing mixers.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with the new raw material specifications and the repaired mixers. This improved robustness is also illustrated in Table 2, where the calculated robustness measures of dissimilarity and variance ratio are again calculated using the new data. These robustness measures, shown for the disturbances that previously resulted in lack of robustness (see Table 1), also confirm that the new process is much more robust.

Some idea of the relative impact of the various actions on the performance and robustness of the process could also be assessed due to a staged implementation of the improvements. The first stage involved removing all the equipment to which the process was not robust and implementing specification regions on the raw materials. The process was improved, but the yield of good membranes increased to only 80%. By also optimizing the process operating conditions to make the process more robust to humidity and ambient temperature, the yield increased to 96%. The final step then involved introducing the modified equipment (mixers 1, 7, 8 and spin-line/degas system 2) back into service. The yield then stabilized around 94%.

Conclusions

The concept of "robust but fragile" processes is used to develop an approach to robustifying processes with respect to important disturbance and equipment variations. The approach is based on using historical plant data to identify those disturbances and equipment units to which the process responses are not robust. Although the methodology is dependent upon the availability of historical data and on a

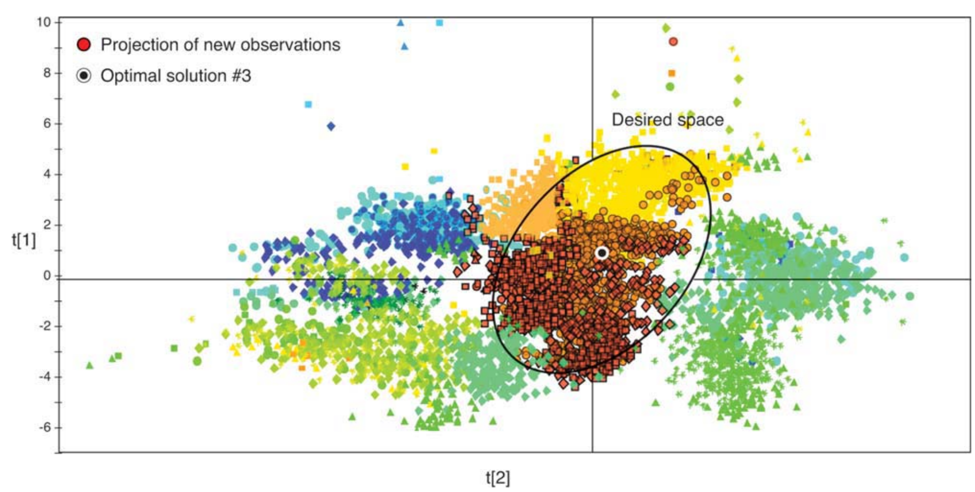

Figure 9. Projection of new observations onto the PCA latent variable space of Y.

Projection of old observations is also shown from Figure 5. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 2. Value of the Robustness Distance Measure ($\overline{D}^2_{Yd/c}$) Calculated as in Eq. 6 and the Variance Ratio ($VR_{d/c}$) Calculated as in Eq. 8 for Identified Disturbances

	$\overline{D}^2_{Yd/c}$	Ratio of the Variances $VR_{d/c}$
Disturbances		
Raw materials	2.23	1.20
Mixer 8	2.16	1.45
Spin-line 2	2.30	1.09
Mixer 7	2.11	1.55
Relative humidity	2.3	1.52
Ambient temperature	2.14	1.21
Mixer 1	2.3	1.65
Degas system 2	2.51	1.55

detailed analysis of those data, such data are usually available from historical databases on most processes. Multivariate robustness measures for the effect of the disturbances on the mean of the process responses and on the variance of the responses were developed for this purpose. LV models based on PCA and PLS and built from the historical data are then used to model the process responses and sensitivities with respect to these disturbances and to the process variables. For those disturbances over which one has some control, such as poorly performing equipment units and raw material supply, direct action can be taken. In the membrane process, the identified poorly performing equipment units were taken out of operation, modified, and returned to service, and multivariate specification regions were imposed on the raw material properties from the suppliers. For those disturbances over which one has little control optimization can be performed in the LV space to find process operating conditions that make the process responses insensitive to them. In the membrane process, these latter disturbances were the environmental conditions (humidity and ambient temperature). Application of the methodology to a hollow fiber membrane manufacturing process was presented to demonstrate the methodology and to illustrate the improved performance and robustness achievable through its application.

Literature Cited

1. Box GEP. Non-normality and tests on variances. *Biometrika*. 1953;40:318–335.
2. Box GEP, Andersen SL. Permutation theory in the derivation of robust criteria and the study of departures from assumption. *J R Stat Soc B*. 1955;17:1–34.
3. Taguchi G. *Introduction to Quality Engineering—Designing Quality into Products and Processes*. Tokyo: Asian Productivity Organization, 1986:122.
4. Nair VN. Taguchi's parameter design: a panel discussion. *Technometrics*. 1992;34:127–161.
5. Leon RV, Shoemaker AC, Kackar RN. Performance measure independent of adjustment: an explanation and extension of Taguchi's signal to ratio. *Technometrics*. 1987;29:253–285.
6. Ross P. *Taguchi Techniques for Quality Engineering*. London: McGraw-Hill, 1996.
7. Montgomery D. *Introduction to Quality Control*. New York: Wiley, 1991.
8. Eriksson L, Johansson E, Kettaneh-Wold N, Wold S. *Multi- and Megavariate Data Analysis Principles and Applications*. Sweden: Umetrics AB,2001.
9. Jackson JE. *A User's Guide to Principal Components*. New York: Wiley, 1991.
10. Wold S. Pattern recognition by means of disjoint principal components models. *Pattern Recognit*. 1976;8:127–139.
11. Kourti T, MacGregor JF. Multivariate statistical process control methods for monitoring and diagnosing process and product performance. *J Qual Technol*. 1996;28:409–428.
12. Raich A, Cinar A. Diagnosis of process disturbances by statistical distance and angle measures. *Comput Chem Eng*. 1997;21:661–673.
13. Yue H, Qin SJ. Reconstruction based fault identification using a combined index. *Ind Eng Chem Res*. 2001;40:4403–4414.
14. Wold S, Kettaneh-Wold N, Skagerberg B. Nonlinear PLS modelling. *Chemom Intell Lab Syst*. 1989;7:53–65.
15. Nocedal J, Wright SJ. *Numerical Optimization*. New York: Springer, 1999.
16. Yacoub F, MacGregor JF. Product optimization and control in the latent variable space of nonlinear PLS models. *Chemom Intell Lab Syst*. 2004;70:63–74.
17. Yacoub F, MacGregor JF. Analysis of processes with complex serial-parallel structures. *Ind Eng Chem Res*. 2009;48:7181–7185.
18. Duchesne C, MacGregor JF. Establishing multivariate specification regions for incoming materials. *J Qual Technol*. 2004;36:78–94.
19. Nomikos P, MacGregor JF. Monitoring of batch processes using multi-way principal components analysis. *Am Inst Chem Eng J*. 1994;40:1361–1375.

Manuscript received July 9, 2009, revision received Jan. 6, 2010, and final revision received Jun. 22, 2010.