Concurrent Monitoring of Operating Condition Deviations and Process Dynamics Anomalies with Slow Feature Analysis

Chao Shang, Fan Yang and Xinqing Gao
Tsinghua National Laboratory for Information Science and Technology (TNList),
Dept. of Automation, Tsinghua University, Beijing 100084, P.R. China
Xiaolin Huang, Johan A.K. Suykens
Dept. of Electrical Engineering (ESAT-STADIUS), KU Leuven, B-3001 Leuven, Belgium
Dexian Huang*
Tsinghua National Laboratory for Information Science and Technology (TNList),
Dept. of Automation, Tsinghua University, Beijing 100084, P.R. China

Latent variable (LV) models have been widely used in multivariate statistical process monitoring. However, whatever deviation from nominal operating condition is detected, an alarm is triggered based on classical monitoring methods. Therefore they fail to distinguish real faults incurring dynamics anomalies from normal deviations in operating conditions. In this article, a new process monitoring strategy based on slow feature analysis (SFA) is proposed for the concurrent monitoring of operating point deviations and process dynamics anomalies. Slow features as LVs are developed to describe slowly varying dynamics, yielding improved physical interpretation. In addition to classical statistics for monitoring deviation from design conditions, two novel indices are proposed to detect anomalies in process dynamics through the slowness of LVs. The proposed approach can distinguish whether the changes in operating conditions are normal or real faults occur. Two case studies show the validity of the SFA-based process monitoring approach.

Keywords: latent variable models, slow feature analysis, process monitoring, alarm removal, fault diagnosis

Introduction

In industrial processes, it is crucial to detect and diagnose faults, process upsets and other abnormal events to achieve safe and efficient operations. In the last two decades, multivariate statistical process monitoring (MSPM) approaches have been extensively studied and implemented in the process control community.[1-3] In general, they resort to massive process data measured in industrial processes rather than first-principle knowledge. Most industrial processes operate around certain well-controlled conditions that are designed in advance. Under design operation conditions, abundant routine data are archived and further analyzed to establish monitoring models. In general, the development of basic monitoring models includes the

* Corresponding concerning this article should be addressed to D. Huang at huangdx@tsinghua.edu.cn

following steps. First, some statistical distribution type $\{P(\boldsymbol{x})\}$ of nominal process data $\boldsymbol{x}$ is assumed. Parameters of $P(\boldsymbol{x})$ are then empirically estimated using the data measured under nominal operating conditions. Finally, confidence limits are set by hypothesis tests with given significance levels. At present, a variety of basic MSPM studies (without adaptation) belong to this category, which utilize different characteristics of process data to design the steady distribution $P(\boldsymbol{x})$ under nominal operating conditions. For example, significant correlation is a unique characteristic of process data because major variations are driven by a smaller number of essential factors, and the effective dimension of process variables is far lower than their actual dimension. In this context, latent variable (LV) models are desirable and popular in modeling the steady nominal distribution $P(\boldsymbol{x})$, thereby gaining extensive attentions in data-driven process monitoring.[4] They provide reduced dimensional subspaces to describe the essential sources of variations, and allow direct visualization and monitoring statistics design. Their intrinsic complexity could be highly reduced and hence generalization be enhanced. Typical data-driven approaches such as principal component analysis (PCA)[5,6] and independent component analysis (ICA)[7-9] can be applied to derive LV subspaces and calculate monitoring statistics, e.g., the Hotelling's T^2 statistic.

In practical scenarios, variations in operating conditions are likely and common, which primarily arise from two aspects. First, the process objectives could be oriented actively according to planning and scheduling with the aim to meet changing demands for its products. For example, the set-point of the melt index in a propylene production process may be adjusted to meet the changing demand of the market. Second, some variations or disturbances, for instance, temperature changes in the ambient environment or the operating condition changes in the upstream unit, might as well exert influence on the operating condition of the process. In both cases, the steady distribution $P(\boldsymbol{x})$ under normal operating condition is disrupted, and therefore alarms would be consistently triggered by basic MSPM methods; however, if the process remains well controlled owing to compensation of controllers, further actions become no longer necessary, and hence alarms in this context should be of not only less importance but also greater distinctions than those detecting real faults.[10] Most hazardously, industrial practitioners cannot distinguish real

faults from such nominal changes in operating conditions simply based on monitoring results.

For process monitoring systems, the key to a correct discrimination between normal changes in operating conditions and real faults is to rely on the process dynamics because different dynamic behaviors are manifested in these two scenarios. In the case of normal changes in operating conditions, processes have a similar control performance to that in reference conditions, since all controlled variables are still well manipulated around their set-points by control loops. In the cases of real faults, processes are problematic because controllers essentially fail to attenuate or reject process anomalies, resulting in unusual dynamic behaviors. To deal with process dynamics, classical LV models have been extended to dynamic versions thereof, such as dynamic PCA (DPCA),[11,12] dynamic PLS (DPLS)[12,13] and dynamic ICA (DICA)[14]. They account for temporal similarities in measurements due to process settling time by means of low-dimensional LVs. Alternatively, state-space models based on subspace identification techniques[15,16] have been applied to process monitoring tasks[17.18]. The latent states in the state-space formulation evolve over time in a Markovian fashion, thereby being temporally correlated. They can be deemed as abstractions of historical time series observations with dynamic information embodied.

However, existing monitoring models cannot provide explicit representations for temporal behaviors of process data. More deeply, the steady-state operating condition of processes could be abstracted as the steady distribution $P(\boldsymbol{x})$, whereas process dynamics could be perceived as the distribution of temporal variations $P(\dot{\boldsymbol{x}})$, which ought to be explained by temporal behaviors of LVs that are *completely irrelevant* with their steady states. They carry different information, analogous to the concepts of ‘position’ and ‘velocity’ in physics, respectively. In order to make a correct evaluation of process dynamics, temporal information should be separated from the steady counterpart such that the temporal dynamics and operating condition can be individually monitored. In this way, an accurate assessment of process dynamics that avoids the effect of steady-state operating condition can be made when the process shifts to a new working point but has nominal dynamics thanks to feedback control. Unfortunately, existing monitoring models fail to isolate the temporal information of $P(\dot{\boldsymbol{x}})$ completely from the steady-state distribution $P(\boldsymbol{x})$. Firstly, DPCA and DICA are inadequate in mining dynamic information from process data

because their LVs on behalf of essential driving factors of processes, are still assumed as statistically independent rather than temporally distributed *a priori*. Secondly, for state-space models, both temporal information of $P(\dot{\boldsymbol{x}})$ and the static information of $P(\boldsymbol{x})$ are coupled in the hidden states, based on which monitoring statistics are established. Consequently, when the process deviates its operating condition, alarms will be raised and valuable temporal information will be buried.

Slow feature analysis (SFA) is a prevailing unsupervised dimension reduction methodology that extracts slowly varying LVs from temporal data.[19] As a biologically inspired approach, SFA was first applied to the analysis of self-organization of complex-cell receptive fields using synthetic image sequences.[20,21] In recent years, SFA has received increasing attentions and its applications in various fields have emerged, such as nonlinear blind source separation[22,23], human action recognition[24], and remote sensing[25,26]. Different from classical LV models, the LVs in SFA, which are termed as slow features, are assumed to be slowly varying a priori, and could be sequentially ordered according to their slowness levels that are statistically measurable. Mathematically, SFA enables separate descriptions to the steady-state distribution $P(\boldsymbol{x})$ and the distribution of temporal variations $P(\dot{\boldsymbol{x}})$. Therefore, SFA has improved interpretation ability than classical LV models in terms of temporal coherence. In the present work, a new statistical process monitoring scheme is proposed based on SFA. In virtue of the interpretable slowness levels of slow features, two novel metrics, namely S^2 and S_e^2, that characterize $P(\dot{\boldsymbol{x}})$ are proposed to monitor dynamic characteristics of processes, in addition to the T^2 and T_e^2 statistics that are responsible for abstracting $P(\boldsymbol{x})$ and detecting steady deviations from the design working point. With these two distinct groups of monitoring statistics, we can extract more meaningful information from routine data as an important reference to effectively distinguish between normal changes in operating status and real faults with dynamics disruptions. Particularly, when the process deviates its nominal operating condition due to set-point switches, the SFA-based monitoring method can not only recognize the operating condition deviation by monitoring $P(\boldsymbol{x})$ with the T^2 and T_e^2 statistics, but also effectively reveal whether or not the process is well

controlled by monitoring $P(\dot{\boldsymbol{x}})$ with the S^2 and S_e^2 statistics. To investigate the feasibility of the proposed approach, its monitoring performance is evaluated and compared with those of classical MSPM methods via comprehensive case studies of a simple continuous stirred tank reactor (CSTR) process and the Tennessee Eastman (TE) benchmark process.

The remaining part of this article proceeds as follows. In the next section, SFA is first introduced and reviewed with detail. Its geometric interpretation is also revealed that both the steady-state distribution $P(\boldsymbol{x})$ and the temporal distribution $P(\dot{\boldsymbol{x}})$ can be individually approximated. Similarities and differences between PCA, ICA and SFA are discussed in Section "Comparisons between PCA, ICA and SFA." The SFA-based process monitoring scheme and associated fault detection indices are developed in Section "Process Monitoring with Slow Features." In Section "CSTR Case Studies", the effectiveness of the proposed method is illustrated through a simple CSTR case. Section "Tennessee Eastman Benchmark Process Case Studies" presents performance of the SFA-based monitoring scheme on the TE benchmark process and comparisons with classical monitoring approaches, followed by concluding remarks in the final section.

Overview of Slow Feature Analysis

Optimization problem

Given a multidimensional input signal, the goal of SFA is to find instantaneous scalar input-output mapping functions of which output signals have as slow variations as possible. The output signals, conceived as LVs, would carry major information provided that a temporal structure is embodied in input data. Mathematically, the optimization problem of SFA can be defined as follows.[19] Assuming that there exists an *m*-dimensional temporal input signal $\boldsymbol{x}(t)=\left[x_1(t),\cdots,x_m(t)\right]^{\mathrm{T}}$, SFA aims to find a set of feature functions $\{g_j(\cdot)\}_{j=1}^{m}$ such that for the LVs $\left\{s_j(t)=g_j(\boldsymbol{x}(t))\right\}_{j=1}^{m}$,

$$\min_{g_j(\cdot)} \left\langle \dot{s}_j^2 \right\rangle_t \tag{1}$$

under the constraints

$$\left\langle s_j \right\rangle_t = 0 \qquad \text{(zero mean)} \tag{2}$$

$$\left\langle s_j^2 \right\rangle_t = 1 \qquad \text{(unit variance)} \tag{3}$$

$$\forall i \neq j: \quad \left\langle s_i s_j \right\rangle_t = 0 \qquad \text{(decorrelation and order)} \tag{4}$$

where $\Delta(s_j) \doteq \left\langle \dot{s}_j^2 \right\rangle_t$ can be seen as a measure of the slowness of $s_j(t)$, $\dot{s}$ indicates the first-order derivative of the LV with respect to time, and $\langle \cdot \rangle_t$ denotes temporal averaging, defined as

$$\langle f \rangle_t \doteq \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} f(t) dt. \tag{5}$$

The objective in (1) explicitly aims at minimizing the temporal variation of LVs, which are referred to as *slow features* (SFs) in this work. In other words, their variations should be as slow as possible. Constraint (2) is added merely to simplify the problem without loss of generality. Constraint (3) helps avoiding the trivial solution $s_j(t) \equiv \text{const}$ and enforces SFs to incorporate some information. Constraint (4) guarantees that SFs carry different information and do not simply reproduce each other. In addition, a descending order can be assigned to $\{s_j(t)\}_{j=1}^m$ such that the most slowly varying SF has the lowest index. The first SF $s_1(t)$ is the slowest, and $s_2(t)$ is the second slowest, and so forth. The number of slow features is assumed as equal to that of input variables.

In general, the form of feature functions can be chosen as linear combinations of some basis functions: $g_j(\boldsymbol{x}) := \sum_i W_{ji} f_i(\boldsymbol{x})$, where $f_i(\boldsymbol{x})$ is the basis function selected in advance. In linear cases, the feature function can be simply expressed as

$$s_j = g_j(\boldsymbol{x}) = \boldsymbol{w}_j^{\mathrm{T}} \boldsymbol{x}, \tag{6}$$

where $\left\{\boldsymbol{w}_j\right\}_{j=1}^m$ denote coefficient vectors. It is equivalent to finding a matrix $\mathbf{W} = \left[\boldsymbol{w}_1 \quad \cdots \quad \boldsymbol{w}_m\right]^{\mathrm{T}}$ such that

$$\boldsymbol{s} = \mathbf{W}\boldsymbol{x}. \tag{7}$$

In this work only the linear case is considered. Substituting (6) into Constraint (2), we obtain:

$$\left\langle s_j \right\rangle_t = \boldsymbol{w}_j^{\mathrm{T}} \left\langle \boldsymbol{x} \right\rangle_t = 0. \tag{8}$$

It can be seen that Constraint (2) is automatically fulfilled if the input variables $\boldsymbol{x}$ have been scaled to zero mean in advance.

The SFA algorithm

A simple algorithm for performing linear SFA has been developed in Ref. 19. Assume that the input data $\boldsymbol{x}$ have zero mean. Similar to ICA, the initial step in linear SFA is whitening, also known as sphering, which eliminates all cross-correlations between input variables. The whitening can be readily accomplished via classical PCA. Assuming that the covariance matrix $\left\langle \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} \right\rangle_t$ has the following singular value decomposition (SVD):

$$\left\langle \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} \right\rangle_t = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}}, \tag{9}$$

the whitening transformation can then be expressed as

$$\boldsymbol{z} = \boldsymbol{\Lambda}^{-1/2}\mathbf{U}^{\mathrm{T}}\boldsymbol{x} \doteq \mathbf{Q}\boldsymbol{x} \tag{10}$$

where $\mathbf{Q} = \boldsymbol{\Lambda}^{-1/2}\mathbf{U}^{\mathrm{T}}$ is the whitening matrix. It obviously holds that $\left\langle \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}} \right\rangle_t = \mathbf{Q}\left\langle \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} \right\rangle_t \mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ and $\left\langle \boldsymbol{z} \right\rangle_t = \mathbf{0}$. Hence the objective of linear SFA is identical to finding a matrix $\mathbf{P} = \mathbf{W}\mathbf{Q}^{-1}$ that satisfies $\boldsymbol{s} = \mathbf{P}\boldsymbol{z}$ because

$$\boldsymbol{s} = \mathbf{W}\boldsymbol{x} = \mathbf{W}\mathbf{Q}^{-1}\boldsymbol{z} \doteq \mathbf{P}\boldsymbol{z}. \tag{11}$$

Then notice that Constraints (3) and (4) can be compactly rewritten in a covariance matrix form:

$$\left\langle \boldsymbol{s}\boldsymbol{s}^{\mathrm{T}} \right\rangle_t = \mathbf{I}. \tag{12}$$

Substituting (11) into (12), we have

$$\left\langle \boldsymbol{s}\boldsymbol{s}^{\mathrm{T}} \right\rangle_t = \mathbf{P}\left\langle \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}} \right\rangle_t \mathbf{P}^{\mathrm{T}} = \mathbf{P}\mathbf{P}^{\mathrm{T}} = \mathbf{I} \tag{13}$$

which implies that $\mathbf{P}$ must be an orthogonal matrix. Therefore the optimization problem of linear SFA reduces to finding an orthogonal matrix $\mathbf{P}$ such that $\left\langle \dot{s}_j^2 \right\rangle_t$ is minimized with $\boldsymbol{s} = \mathbf{P}\boldsymbol{z}$. This can be easily resolved by applying SVD to the covariance matrix $\left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}} \right\rangle_t = \mathbf{P}\boldsymbol{\Omega}\mathbf{P}^{\mathrm{T}}$ and obtaining orthogonal eigenvectors $\left\{ \boldsymbol{p}_j \right\}_{j=1}^{m}$ and the corresponding eigenvalues $\left\{ \omega_j \right\}_{j=1}^{m}$. Assuming that

eigenvalues $\left\{\omega_j\right\}_{j=1}^{m}$ are organized in an ascending order, we can easily verify that

$$\begin{aligned}\left\langle \boldsymbol{s}\boldsymbol{s}^{\mathrm{T}}\right\rangle_t &= \mathbf{P}\left\langle \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}}\right\rangle_t \mathbf{P}^{\mathrm{T}} = \mathbf{I},\\ \left\langle \dot{s}_j^2\right\rangle_t &= \boldsymbol{p}_j^{\mathrm{T}}\left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}}\right\rangle_t \boldsymbol{p}_j = \omega_j.\end{aligned} \tag{14}$$

In addition, it holds that $\forall i < j$, $\left\langle \dot{s}_i\, \dot{s}_j\right\rangle = 0$. The proof is given in Appendix A. Notice that linear SFA is able to reach the global optimal solution because the global minimum of $\left\langle \dot{s}_j^2\right\rangle_t$ can be found by means of SVD. Finally the transformation matrix **W** can be calculated as

$$\mathbf{W} = \mathbf{P}\mathbf{Q} = \mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{U}^{\mathrm{T}}. \tag{15}$$

In practical scenarios, process data $\boldsymbol{x}(t)$ are often collected with a discrete sampling interval Δt. Therefore, the temporal derivative can be empirically approximated as

$$\dot{z}_j(t) \approx \frac{z_j(t) - z_j(t-\Delta t)}{\Delta t}. \tag{16}$$

Geometric properties of SFA

An elegant geometric interpretation for SFA has been provided by Turner and Sahani[27], which is introduced here to illustrate that SFA can simultaneously model both the steady-state distribution $P(\boldsymbol{x})$ and the temporal distribution $P(\dot{\boldsymbol{x}})$. Figure 1 gives a two-dimensional example of geometric properties of SFA. In the first step of the SFA algorithm, original inputs are first sphered, as indicated by (10). Figure 1(a) shows the covariance $\left\langle \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}}\right\rangle_t$ of the sphered transformations $\{z_1, z_2\}$, conceptually plotted as an isotropic blue circle. Figure 1(b) depicts the covariance $\left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}}\right\rangle_t$ of the temporal derivatives $\{\dot{z}_1, \dot{z}_2\}$, plotted as a red dotted ellipsoid. Then in the second step, the slow feature directions, which are expressed as arrows in Figures 1(a) and 1(b), are obtained by eigen-decomposition of the covariance $\left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}}\right\rangle_t$ in Figure 1(b). Notice that the direction of s_1 coincides with the minor axis of the ellipsoid because the slowest feature corresponds to the smallest eigenvalue, which indicates the optimal value in the minimization problem of (1). This differs from PCA since dominant PCs tend to capture *the most* variances in input data and result in a maximization problem. It can be clearly seen that the steady-state distribution $P(\boldsymbol{x})$ is characterized by the transformed slow features ***s*** in Figure 1(a), which are

distributed as isotropic Gaussian; and the dynamic behavior $P(\dot{\boldsymbol{x}})$ is characterized by the distribution of temporal variations $\dot{\boldsymbol{s}}$, which is assumed as non-isotropic Gaussian. In this sense, SFA provides exclusive allowance for both the steady-state distribution $P(\boldsymbol{x})$ and the temporal distribution $P(\dot{\boldsymbol{x}})$.

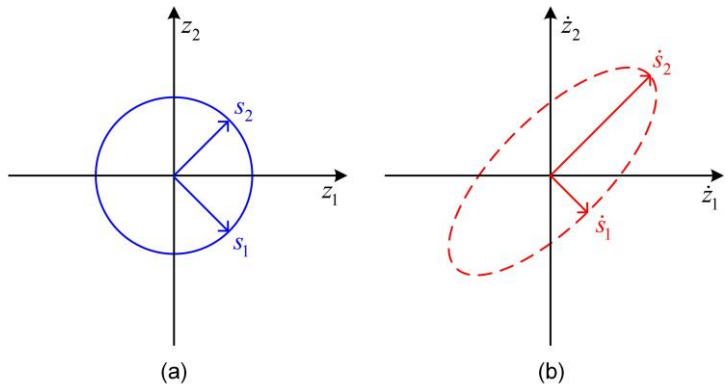


Figure 1. A two-dimensional example of the geometric interpretation of SFA. (a) Covariance $\left\langle \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}} \right\rangle_t$ of the sphered space (blue circle) that characterizes $P(\boldsymbol{x})$. (b) Covariance $\left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}} \right\rangle_t$ of temporal difference in the sphered space (red dotted ellipsoid) that describes $P(\dot{\boldsymbol{x}})$, and slow feature directions (red arrows).

Statistical Properties of SFA Assuming that all diagonal entries of $\boldsymbol{\Lambda}$ are non-zero, matrix $\mathbf{W}$ will be invertible and thus $\boldsymbol{x}$ can be represented as a linear combination of SFs $\boldsymbol{s}$ as $\boldsymbol{x} = \mathbf{R}\boldsymbol{s}$, where

$$\mathbf{R} = \mathbf{W}^{-1} = \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{P}^{\mathrm{T}}. \tag{17}$$

And the statistical properties of SFA can be finally summarized as

$$\boldsymbol{x} = \mathbf{R}\boldsymbol{s},\ \mathbb{E}\{\boldsymbol{s}\} = \mathbf{0},\ \mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{\mathrm{T}}\} = \mathbf{I},\ \mathbb{E}\{\dot{\boldsymbol{s}}\} = \mathbf{0},\ \mathbb{E}\{\dot{\boldsymbol{s}}\dot{\boldsymbol{s}}^{\mathrm{T}}\} = \boldsymbol{\Omega}, \tag{18}$$

such that

$$\mathbb{E}\{\boldsymbol{x}\} = \mathbf{0},\ \mathbb{E}\{\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}},\ \mathbb{E}\{\dot{\boldsymbol{x}}\} = \mathbf{0},\ \mathbb{E}\{\dot{\boldsymbol{x}}\dot{\boldsymbol{x}}^{\mathrm{T}}\} = \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Omega}\mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{U}^{\mathrm{T}}. \tag{19}$$

Dynamic SFA

Notice that SFs $\{s_j(t)\}_{j=1}^{m}$ are calculated by using process data $\boldsymbol{x}(t)$ measured at the current snapshot *t*. It implies that SFs cannot be derived by filtering time series data, being irrelevant with observations at past time. It violates the physical truth of chemical processes that

underlying factors are related to historical process data in a period of time due to process dynamics. In this regard, SFA can be extended to its dynamic version by augmenting each input vector with d lagged samples and stacking the data matrix as follows:

$$\boldsymbol{X}(d)=\begin{bmatrix} \boldsymbol{x}(t)^{\mathrm{T}} & \boldsymbol{x}(t-1)^{\mathrm{T}} & \cdots & \boldsymbol{x}(t-d)^{\mathrm{T}} \\ \boldsymbol{x}(t+1)^{\mathrm{T}} & \boldsymbol{x}(t)^{\mathrm{T}} & \cdots & \boldsymbol{x}(t-d+1)^{\mathrm{T}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}(t+N-1)^{\mathrm{T}} & \boldsymbol{x}(t+N-2)^{\mathrm{T}} & \cdots & \boldsymbol{x}(t+N-d-1)^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{m(d+1)\times N}, \qquad (20)$$

where N denotes the number of available process samples. Therefore, the input dimension of the SFA model becomes $m(d+1)$.

Comparisons between PCA, ICA and SFA

One similarity shared by PCA, ICA and SFA is that a low-dimensional LV subspace exists. In linear cases, their LVs are obtained via a linear combination of high-dimensional inputs. However, they are characterized by different physical meanings, which are summarized in Table 1. The LVs in PCA, named principal components (PCs), are expected to have as much variances as possible. The LVs in ICA, named independent components (ICs), are expected to be as statistically independent as possible. SFs in SFA are assumed to be as slowly varying as possible.

Table 1. A Brief Comparison of Different Latent Variable Models

Model	**Name of LV**	**Properties of LV**	**Ordering Criterion for LV**	**Global Optimality**	**Computation Complexity**
PCA	Principal Component (PC)	Having as much variance as possible	Captured Variance	√	1 SVD for optimization
ICA	Independent Component (IC)	Being as statistically independent as possible	No Standard Criterion	×	1 SVD for sphering + iterative optimization
SFA	Slow Feature (SF)	Being as slowly varying as possible	Temporal Slowness	√	1 SVD for sphering + 1 SVD for optimization

Both PCA and SFA provide explicit criteria for ordering LVs, as shown in the fourth column of Table 1. The ordering of LVs, however, is very difficult in ICA and there is no standard criterion.[28] In addition, PCA and SFA can reach global optimum. The above limitations add some

difficulties in applying ICA to process monitoring.

As for the computational complexity, PCA is the simplest since only one step of SVD is needed. The computational cost of SFA is relatively low because only an additional SVD helps deriving all parameters, which provides implementation convenience for applying SFA in practical scenarios. It can be seen from the last column of Table 1 that ICA and SFA share one similarity that PCA is performed to sphere data in the first step. However, ICA has the most tedious algorithm because it involves iterative optimization in subsequent steps.

Formal equivalence between ICA and linear SFA

At first glance, ICA and SFA are different dimension reduction approaches because they have their own optimization objectives. The objective in SFA can be explicitly expressed as (1), while the statistical independence in ICA is difficult to optimize and necessitates approximate solutions. Ref. 28 proposes to use degrees of non-Gaussianity such as kurtosis and negentropy as an alternate for statistical independence. It is worth mentioning that ICA can be performed as well by using other criteria to measure independence. In fact, if the second-order correlation with time delay one, i.e.,

$$C(\Delta t) = \frac{1}{2}\left[\left\langle \boldsymbol{s}(t)\boldsymbol{s}(t+\Delta t)^{\mathrm{T}}\right\rangle_t + \left\langle \boldsymbol{s}(t+\Delta t)\boldsymbol{s}(t)^{\mathrm{T}}\right\rangle_t\right] \tag{21}$$

is used to measure statistical independence, ICA becomes formally identical to linear SFA.[29] Therefore, both classical ICA and linear SFA could be conceived to pursue statistical independence, only with different criteria to measure independence. The SFA utilizes time shift correlation to cover dynamic patterns, leading to a better interpretation for temporal structures. In this sense, SFA is superior to ICA in extracting dynamic information from time series data. An inevitable limitation of SFA is that correlation is used instead of independence, which might cause loss of information.

Process Monitoring with Slow Features

Dimension reduction of SFA

In the aforementioned SFA algorithm, the number of acquired SFs is the same as that of process variables. In order to carry out dimension reduction and denoising, only a small portion of SFs which carry significant information should be selected to reduce the model complexity,

whereas the remaining ones correspond to unimportant residuals. If SFA is applied to chemical process data, the derived slowest features tend to catch the general trend of process variations, whereas the fastest ones can be seen as short-term noises. For dimension reduction purposes, it is therefore reasonable to select $M\left(<m(d+1)\right)$ slowest features $\{s_1,\cdots,s_M\}$ as the dominant components.

The choice of *M*, i.e., the number of SFs is an important issue in process monitoring. Including too many SFs may put unnecessary emphasis on noise and further deteriorate monitoring performance. Unfortunately, there are few objective guidelines for choosing the number of SFs. Here we suggest a criterion from a reconstruction standpoint. A process variable x_j can be exactly recovered from slow features via the following linear mapping:

$$x_j = \boldsymbol{r}_j^{\mathrm{T}} \boldsymbol{s} \tag{22}$$

where $\boldsymbol{r}_j^{\mathrm{T}}$ denotes the *j*th row of $\mathbf{R} = \mathbf{W}^{-1}$. For dimension reduction purposes, a few SFs are selected to derive the following denoised reconstruction

$$x_j^{\mathrm{rec}} = \hat{\boldsymbol{r}}_j^{\mathrm{T}} \boldsymbol{s}, \tag{23}$$

where $\hat{\boldsymbol{r}}_j$ is the reconstruction vector formed by replacing some components of $\boldsymbol{r}_j$ with zeros.

The slowness $\Delta(x_j^{\mathrm{rec}})$ is an objective criterion to evaluate the importance of information retained in the reconstruction x_j^{rec}. Including too much fast varying noise would inevitably entail a high value in $\Delta(x_j^{\mathrm{rec}})$. Because of the denoising effect, an appropriate reconstruction x_j^{rec} is expected to have slower variations than the original input x_j, i.e., $\Delta(x_j^{\mathrm{rec}}) < \Delta(x_j)$. The following proposition reveals that $\Delta(x_j)$ is linearly dependent on $\Delta(s_i)$.

Proposition 1: The slowness of x_j is a weighted sum of the slowness levels of all SFs. That is,

$$\Delta(x_j) = \sum\nolimits_i \alpha_i \Delta(s_i) \tag{24}$$

where $\alpha_i = r_{ji}^2 / \sum_i r_{ji}^2$ and $\sum_i \alpha_i = 1$.

The proof of Proposition 1 is given in Appendix B. From (24), we see by inspection that the SFs with $\Delta(s_i) > \Delta(x_j)$ that vary faster than x_j tend to drive x_j to be faster, whereas the

SFs with $\Delta(s_i) < \Delta(x_j)$ that vary slower than x_j tend to drive x_j to be slower. To remove the noise effect, an intuitive choice is to remove those faster SFs with $\Delta(s_i) > \Delta(x_j)$ because they make x_j faster. Interestingly, it can be proved that such an empirical strategy renders reconstruction slower according to the following proposition.

Proposition 2: If $\hat{\boldsymbol{r}}_j = \boldsymbol{r}_j - r_{jk}\boldsymbol{e}_k$ where $\boldsymbol{e}_k$ stands for the *j*th unit vector, of which the *j*th component is one and the others equal zero, along with the index *k* satisfying $\Delta(s_k) > \Delta(x_j)$, a slower reconstruction would be obtained such that $\Delta(x_j^{\text{rec}}) < \Delta(x_j)$.

The proof of Proposition 2 is given in Appendix C. In this way, we can sequentially preclude the fastest feature from SFs that are kept, thereby continuously driving the reconstruction slower. As discussed in Proposition 1, we suggest removing SFs that are faster than all input variables, the number of which is given by

$$M_e = \text{card}\left\{ s_i \mid \Delta(s_i) > \max_j \left\{ \Delta(x_j) \right\} \right\}, \tag{25}$$

where $\text{card}\{\cdot\}$ denotes the number of elements in a certain set. The number of dominant SFs can then be calculated as

$$M = m(d+1) - M_e, \tag{26}$$

However, the estimation in (25) might be not robust because the maximum function is sensitive to outliers. Process variables used for monitoring are usually selected manually, and thus some process-irrelevant but fast varying signals might be involved, resulting in an inadequate estimation of M_e. In this regard, we use the quantile statistic instead. The estimation for M_e can then be rectified as

$$M_e = \text{card}\left\{ s_i \mid \Delta(s_i) > \max_j^q \left\{ \Delta(x_j) \right\} \right\}, \tag{27}$$

where we can set $q = 0.1$ and $\max_j^q \left\{ \Delta(x_j) \right\}$ stands for the *q*-upper quantile of the set $\left\{ \Delta(x_j) \right\}$. Finally, the dominant slowest SFs with reduced dimension can be expressed as $\boldsymbol{s}_d = [s_1, \cdots, s_M]^{\mathrm{T}} \in \mathbb{R}^M$, whereas the remaining fastest ones are denoted as $\boldsymbol{s}_e = [s_{M+1}, \cdots, s_{m(d+1)}]^{\mathrm{T}} \in \mathbb{R}^{M_e}$.

Monitoring statistics with SFs

Based on the low-dimensional SFA model given in the previous subsection, monitoring indices are established with reference to both distributions $P(\boldsymbol{x})$ and $P(\dot{\boldsymbol{x}})$. Dominant SFs $\boldsymbol{s}_d$ and residual SFs $\boldsymbol{s}_e$ form an orthogonal decomposition to the data space of $\boldsymbol{x}$. First, the Hotelling's T^2 statistic is applied to both $\boldsymbol{s}_d$ and $\boldsymbol{s}_e$ to characterize $P(\boldsymbol{x})$ for monitoring the systematic change of a common cause, similar to that of PCA- and ICA-based monitoring methods. Noticing that SFs have unit variances, the Hotelling's T^2 statistics are defined as follows:

$$T^2 = \boldsymbol{s}_d^{\mathrm{T}} \boldsymbol{s}_d, \tag{28}$$

$$T_e^2 = \boldsymbol{s}_e^{\mathrm{T}} \boldsymbol{s}_e. \tag{29}$$

According to (12), SFs $\boldsymbol{s} = [s_1, \cdots, s_{m(d+1)}]^{\mathrm{T}} = [\boldsymbol{s}_d^{\mathrm{T}}, \boldsymbol{s}_e^{\mathrm{T}}]^{\mathrm{T}}$ are assumed to be independently Gaussian distributed. Therefore, the T^2 statistic follows a χ^2 distribution with M degrees of freedom, and the T_e^2 statistic follows a χ^2 distribution with M_e degrees of freedom,[30] i.e.,

$$T^2 = \boldsymbol{s}_d^{\mathrm{T}} \boldsymbol{s}_d \sim \chi_M^2, \tag{30}$$

$$T_e^2 = \boldsymbol{s}_e^{\mathrm{T}} \boldsymbol{s}_e \sim \chi_{M_e}^2. \tag{31}$$

The T^2 statistic measures the static variations inside the subspace spanned by the dominant slowest SFs, whereas the T_e^2 measures the static variations of the remaining fastest SFs. These two statistics are responsible for evaluating the correspondence of a data point with the steady-state distribution $P(\boldsymbol{x})$. Besides, *temporal slowness*, i.e. $\dot{\boldsymbol{s}} = [\dot{s}_1, \cdots, \dot{s}_{m(d+1)}]^{\mathrm{T}} = [\dot{\boldsymbol{s}}_d^{\mathrm{T}}, \dot{\boldsymbol{s}}_e^{\mathrm{T}}]^{\mathrm{T}}$, is also an important attribute of SFs, which actually describes the temporal distribution $P(\dot{\boldsymbol{x}})$, i.e. temporal dynamics within processes. As indicated by Appendix A, $\dot{\boldsymbol{s}}_d$ and $\dot{\boldsymbol{s}}_e$ also form an orthogonal decomposition to the data space of $\dot{\boldsymbol{x}}$. It is then straightforward to define analogous monitoring indices for $\dot{\boldsymbol{s}}_d$ and $\dot{\boldsymbol{s}}_e$ as S^2 and S_e^2 statistics, to describe the distribution of dynamic variations in nominal operation. These two statistics are calculated as follows:

$$S^2 = \dot{\boldsymbol{s}}_d^{\mathrm{T}} \boldsymbol{\Omega}_d^{-1} \dot{\boldsymbol{s}}_d , \tag{32}$$

$$S_e^2 = \dot{\boldsymbol{s}}_e^{\mathrm{T}} \boldsymbol{\Omega}_e^{-1} \dot{\boldsymbol{s}}_e . \tag{33}$$

where $\boldsymbol{\Omega}_d$ and $\boldsymbol{\Omega}_e$ denote the empirical covariance matrix of $\dot{\boldsymbol{s}}_d$ and $\dot{\boldsymbol{s}}_e$, respectively, which are calculated as $\boldsymbol{\Omega}_d = \left\langle \dot{\boldsymbol{s}}_d \dot{\boldsymbol{s}}_d^{\mathrm{T}} \right\rangle_t = \mathrm{diag}\{\omega_1, \cdots, \omega_M\}$ and $\boldsymbol{\Omega}_e = \left\langle \dot{\boldsymbol{s}}_e \dot{\boldsymbol{s}}_e^{\mathrm{T}} \right\rangle_t = \mathrm{diag}\{\omega_{M+1}, \cdots, \omega_{m(d+1)}\}$. With postulation that both $\dot{\boldsymbol{s}}_d$ and $\dot{\boldsymbol{s}}_e$ follow multivariate Gaussian distributions, the S^2 statistic follows a scaled F-distribution with M and $N-M-1$ degrees of freedom, and the S_e^2 statistic follows a scaled F-distribution with M_e and $N-M_e-1$ degrees of freedom:[2,31]

$$S^2 \sim gF_{M,N-M-1} , \tag{34}$$

$$S_e^2 \sim g_e F_{M_e,N-M_e-1} , \tag{35}$$

where constants g and g_e are computed as

$$g = \frac{M(N^2-2N)}{(N-1)(N-M-1)} , \tag{36}$$

$$g_e = \frac{M_e(N^2-2N)}{(N-1)(N-M_e-1)} . \tag{37}$$

To use the above indices for monitoring, control limits should be estimated using routine data. With $(1-\alpha)$ confidence level, the monitoring policy can be summarized as follows:

1. If $T^2 > \chi^2_{M,\alpha}$ or $T_e^2 > \chi^2_{M_e,\alpha}$, a steady deviation from the design operating condition is detected.
2. If $S^2 > gF_{M,N-M-1,\alpha}$ or $S_e^2 > g_e F_{M_e,N-M_e-1,\alpha}$, a potential anomaly in process dynamics is detected.

In summary, four indices of the SFA-based process monitoring method are used in together with different physical meanings. The T^2 and T_e^2 statistics measure the consistency of a data point $\boldsymbol{x}$ with the steady-state distribution $P(\boldsymbol{x})$, whereas the S^2 and S_e^2 statistics evaluate the

its consistency with the temporal distribution $P(\dot{\boldsymbol{x}})$. In practice, the proposed monitoring scheme recognizes steady deviation from design operating point using the T^2 and T_e^2 indices, and detects variations in dynamic characteristics of processes using the S^2 and S_e^2 indices. If either S^2 or S_e^2 goes beyond the control limit after T^2 or T_e^2 detects a steady deviation, process dynamics is considered to be affected by a certain fault and further measurements should be taken. If both S^2 and S_e^2 return normal after either T^2 or T_e^2 exceeds the threshold, the disturbance is considered as having no influence on process dynamics and all controlled variables are under control although a new operating status has been created.

CSTR Case Studies

In this section, the proposed method for process monitoring is illustrated through a simple non-isothermal CSTR process[32] with clear first-principle knowledge incorporated in typical fault scenarios. The process has one feed stream that contains the reactant *A*, one product stream, and a cooling water stream. It is assumed that the tank is well mixed and its physical properties keep constant. The CSTR process can be described by the component material balance and the energy balance, which are formulated as a set of differential equations:

$$\frac{dC_A}{dt} = \frac{q}{V}\left(C_{Af} - C_A\right) - k_0 \exp\left\{-\frac{E}{RT}\right\} C_A + v_1 \tag{38}$$

$$\frac{dT}{dt} = \frac{q}{V}\left(T_f - T\right) - \frac{\Delta H}{\rho C_p} k_0 \exp\left\{-\frac{E}{RT}\right\} C_A + \frac{UA}{V\rho C_p}\left(T_c - T\right) + v_2 \tag{39}$$

where C_A is the outlet concentration, T is the reactor temperature, T_c is the temperature of cooling water, q is the feed flow rate, C_{Af} is the feed concentration, T_f is the feed temperature, and $[v_1, v_2]^{\mathrm{T}}$ are independent system noises. In the simulation, $[C_A, T]^{\mathrm{T}}$ are controlled variables with nominal values, and $[q, T_c]^{\mathrm{T}}$ are chosen as manipulated variables with feedback from control errors. A sampling interval $\Delta t = 1\ \mathrm{min}$ was used to collect the measured process variable $[C_A, T, T_c, q]^{\mathrm{T}}$, to which measurement noise e is added. In addition, random disturbances, which are modeled as first-order auto-regressive processes, were added to the system. Negative feedback inputs were added to $[q, T_c]^{\mathrm{T}}$ with PID controllers as $\mathbf{K}_2(K_1 + T_d s + T_I / s)\varepsilon$, where

$\varepsilon = [C_A - C_A^*, T - T^*]^{\mathrm{T}}$ is the residual vector. All the system parameters and conditions are reported in Table 2.

Table 2. Simulation Condition of CSTR

Simulation parameters	$-\Delta H = 17835.821$ J/mol, $\rho = 1000$ g/L, $E/R = 5360$ K, $UA = 11950$ J/(min·K), $C_p = 0.239$ J/(g·K)
Initial conditions	$q = 100$ L/min, $T_c = 419$ K, $C_{Af} = 1$ mol/L
Nominal values	$C_A^* = 0.2$ mol/L, $T^* = 446$ K
Controller setting	$K_1 = 1, T_d = 0.1, T_I = 10, \mathbf{K}_2 = [5,1;1,2]$
Measurement noise	$e_T \sim N(0,0.005), e_{C_A} \sim N(0,0.005)$, $e_{T_c} \sim N(0,10^{-6}), e_q \sim N(0,10^{-6})$
System fluctuations	$v_1(t) = 0.999v_1(t-1) + 0.001\sigma_e e_1(t)$, $v_2(t) = 0.999v_2(t-1) + 0.001\sigma_e e_2(t)$, $e_i \sim N(0,1)\ (i = 1,2), \sigma_e = 0.1$

There are 2000 normal samples to build four monitoring models, i.e., DPCA, DICA, canonical variate analysis (CVA) and linear SFA. In this work, CVA is chosen as a representative state-space model widely applied to process monitoring[15,18] for comparisons. The auto-correlation plots are shown in Figure 2, in which we can observe that x_1 shows high correlations with x_3 and x_4 with one time lag. Therefore, one lagged variable of each measurement is sufficient to include relevant dynamic information and is hence adopted for all methods. Two PCs are selected for DPCA according to cross-validation based on the PRESS statistic[33,34], and two ICs are selected for DICA according to Ref. 9. For CVA, the order of state-space is chosen as 2 according to dominant singular values[35]. For SFA, the coefficient matrix **W** is calculated as:

$$\mathbf{W} = \begin{bmatrix} 0.0756 & 0.0698 & -0.0045 & -0.0214 & 3.5149 & 3.1707 & -3.7276 & -3.6521 \\ 0.9695 & 0.9758 & 0.2282 & 0.2127 & 4.1806 & 6.4189 & -3.3021 & -5.9049 \\ 0.2681 & 0.2363 & -0.6882 & -0.6782 & 0.6354 & 2.2702 & -0.2697 & -2.1950 \\ -0.0955 & -0.1574 & 0.0533 & 0.0588 & 8.5809 & 62.1334 & 0.3721 & -62.5218 \\ 0.0035 & -0.0400 & -0.0239 & -0.0540 & -3.3829 & 11.6869 & 5.2223 & -12.5184 \\ -0.8848 & 0.9070 & 0.3629 & -0.3480 & -3.5271 & 1.7527 & 2.9357 & -1.3636 \\ 0.5450 & -0.5165 & 0.6223 & -0.6345 & 3.1600 & -2.7806 & -2.7625 & 2.4435 \\ 0.0453 & -0.0111 & 0.0295 & -0.0367 & -3.4371 & 3.0933 & 3.1743 & -2.6397 \end{bmatrix}$$

(40)

and the optimal number of slow features is determined as 5 based on the proposed criterion. Two test datasets are elaborated with abnormal conditions with the aim to demonstrate the monitoring capability of the SFA-based approach. In the first dataset, the disturbance is set to change the operating condition only, while in the second dataset, both the operating condition and process dynamics are affected. There are 2000 samples in each test dataset, and the corresponding disturbance is introduced from the 300th sample and removed at the 1300th sample. Detailed settings of abnormal datasets are tabulated in Table 3. The 99% confidence limits are obtained from data in nominal operating condition for DPCA, DICA, CVA and SFA.

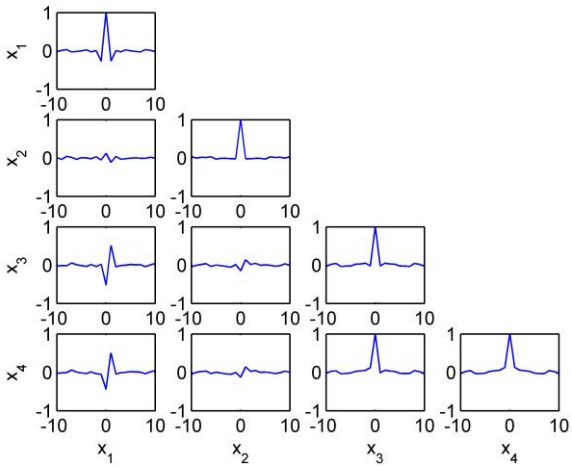


Figure 2. Auto-correlation plots for four input variables in the CSTR example

Table 3. Two Cases with Different Process Disturbances in the CSTR Example

Case	Description
Disturbance changes the operating condition only	Sample 1-300: T_f = 400 K Sample 301-1300: T_f = 401 K Sample 1301-2000: T_f = 400 K
Disturbance simultaneously affects the operating condition and dynamic performances	Sample 1-300: $k_0 = 6.6\times10^5$ min^{-1} Sample 301-500: k_0 linearly decreases to 5.0×10^5 min^{-1} Sample 501-1300: $k_0 = 5.0\times10^5$ min^{-1} Sample 1301-2000: $k_0 = 6.6\times10^5$ min^{-1}

Disturbance changes the operating condition only

In this scenario, a minor increase of 1 K occurs in the feed temperature T_f. Because the reactor temperature is controlled by manipulating the coolant temperature, the entire process could become under control again after the compensation. Notice that dynamic behaviors of process in

this case remain unchanged. This can be deduced from the fact that time constants of $[C_A, T]^T$ in (38) and (39) are irrelevant with $[T_f, T]^T$, and the temperature of cooling water T_c would be adjusted to compensate the change in the feed temperature T_f , rendering the constant term in (39) unchanged. Therefore, when the disturbance occurs, the reactor temperature can still be well controlled around 446K (its set-point value), as shown in Figure 3. The monitoring results are shown in Figure 4. For this tiny external disturbance, DPCA fails to detect it in both indices, while both DICA and CVA successfully detect it. However, we cannot distinguish whether or not this disturbance actually deteriorates the control performance from results of DICA and CVA. Note that although the canonical variates (CVs) of CVA are temporally correlated, they contain information about both temporal variations and steady states, and thus cannot give an explicit assessment of temporal dynamics. In contrast, the SFA-based approach delivers a reasonable monitoring result. On one hand, the T^2 index is obviously magnified, indicating that the operating condition has changed. On the other hand, the S^2 and S_e^2 indices stay below control limits thereof after the process enters a new operating condition, indicating that process dynamics is still normal. A simultaneous utilization of these four indices is able to not only monitor this disturbance sensitively but also recognize that the process still operates normally. In particular, the S^2 and S_e^2 indices provide additional insight into the control performance. Nominal values of the S^2 and S_e^2 indices indicate that all controlled variables are under good control. Considering the case that controlled variables fail to return to set-points thereof, controllers will tend to make unusual responses that differ from the nominal condition, thereby affecting process dynamics. In this example, the reactor temperature is under control as shown in Figure 3, which is consistent with trends of the S^2 and S_e^2 indices. In addition, temporary effects of controllers when T_f is changed are exhibited as two significant peaks in the S^2 chart around the 300th and the 1300th samples in Figure 4(d).

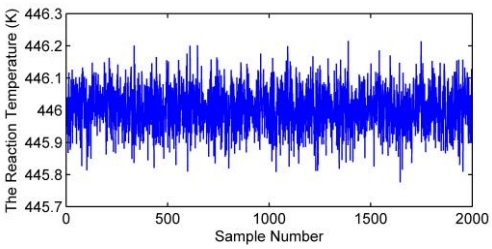
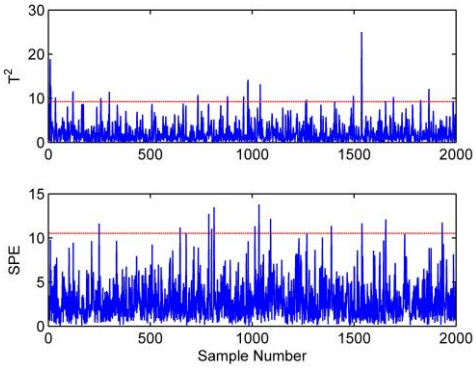
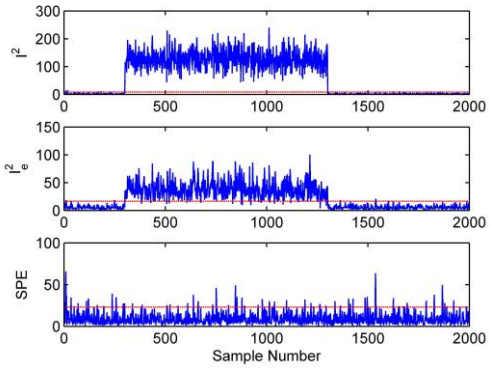


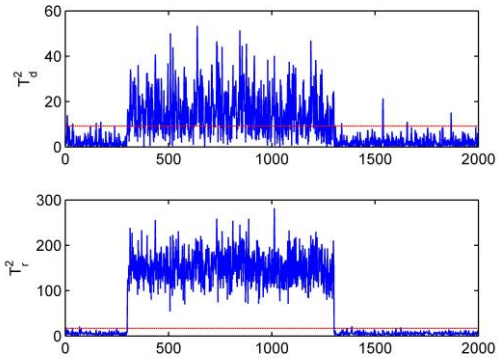
Figure 3. The reactor temperature in the presence of step changes in the input reactant temperature



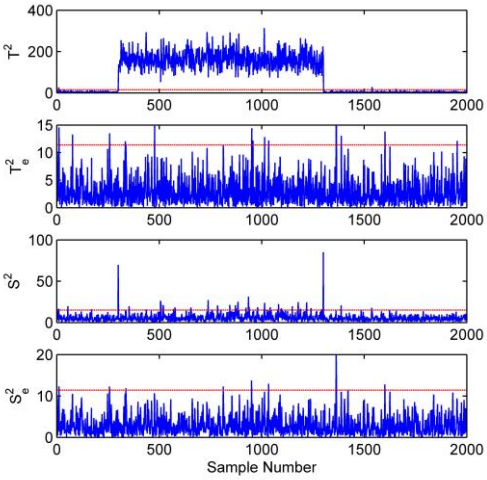
(a)



(b)



(c)

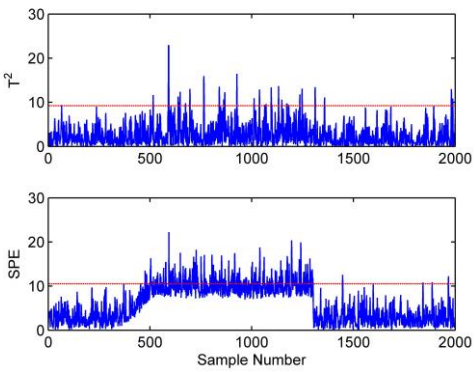


(d)

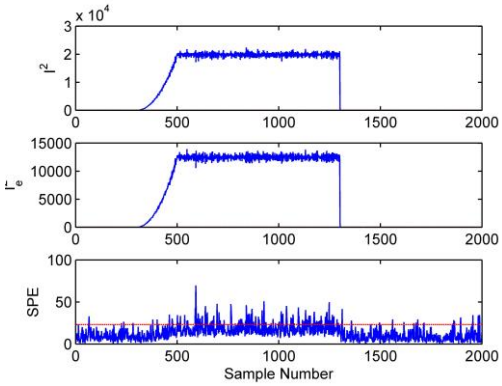
Figure 4. Monitoring results for an increase in the feed temperature. (a) DPCA, (b) DICA, (c) CVA, (d) SFA.

Disturbance simultaneously affects the operating condition and dynamic performances

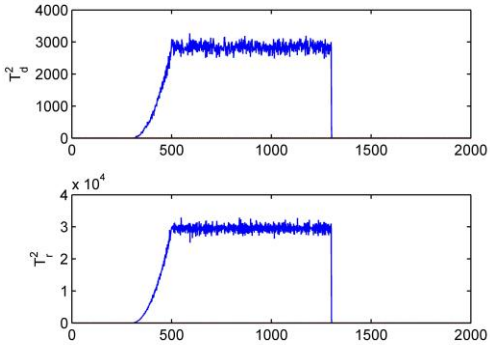
In this scenario, a deactivation of catalyst causes changes in process dynamics. Because k_0 is directly related to the time constant of C_A (see (38) and (39)), the process dynamics would be affected as expected. Therefore the control performance would deteriorate after the deactivation of catalyst. The monitoring results are shown in Figure 5. In the DPCA-based monitoring, the fault can be detected by the SPE index; however, there are still many samples below the control limit. All charts in DICA and CVA are able to detect this fault, but in a similar manner to the previous example; however, a clear interpretation of temporal dynamics which distinguishes from normal changes in operating conditions is absent. Hence this crucial fault would be masked in massive alarms of DICA and CVA. In contrast, the T^2 and T_e^2 indices of SFA also manage to detect this fault, and subsequently S^2 implies that this fault does affect process dynamics.



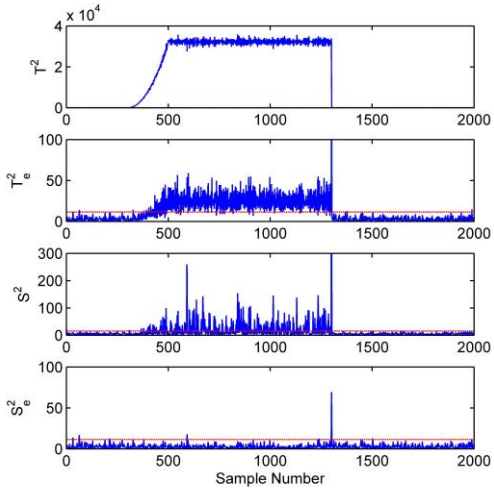
(a)



(b)



(c)



(d)

Figure 5. Monitoring results for a deactivation of catalyst. (a) DPCA, (b) DICA, (c) CVA, (d) SFA.

False alarm rates with out-of-sample normal data

Generalization ability is a major concern about monitoring models. From a viewpoint of process monitoring, weak generalization ability inevitably incurs a high false alarm rate (FAR) because the model could be easily violated by out-of-sample normal observations. In this regard,

we calculate FARs of diverse models using an independent test dataset with 2000 routine samples, which is not used for model training. The results are given in Table 4, in which we can observe that four statistics of the SFA-based monitoring model give satisfactory FARs comparable to those of DPCA, DICA and CVA. For four models, their false alarms can be further reduced using a more flexible rule that an alarm is indicated only when points in a moving window consecutively fall out of the control limit, as shown in the last column of Table 4. Using a window length $l = 2$, which indicates that an alarm is issued only when two consecutive violations in the control limit are detected, effectively helps pruning out all unnecessary false alarms. In this respect, SFA can fit process data desirably and the related monitoring scheme is practically applicable.

Table 4. False Alarm Rates in the CSTR Example (%)

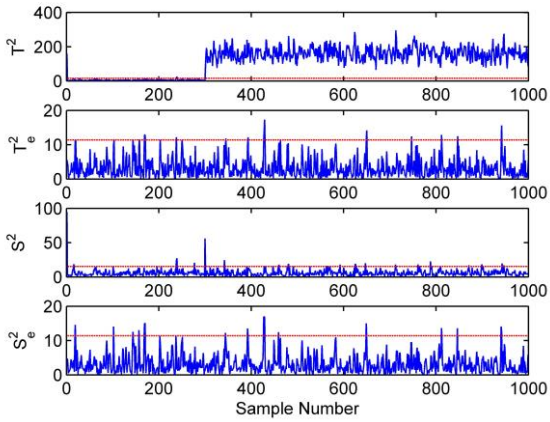
		Window Length $l = 1$	**Window Length** $l = 2$
DPCA	T^2	0.85	0
	SPE	0.55	0
	Overall	1.40	0
DICA	I^2	1.10	0
	I_e^2	1.05	0
	SPE	1.95	0
	Overall	3.55	0
CVA	T_d^2	1.35	0
	T_r^2	0.55	0
	Overall	1.90	0
SFA	T^2	0.95	0
	T_e^2	0.60	0
	S^2	0.40	0
	S_e^2	0.85	0
	Overall	2.45	0

Alarm patterns of typical disturbances

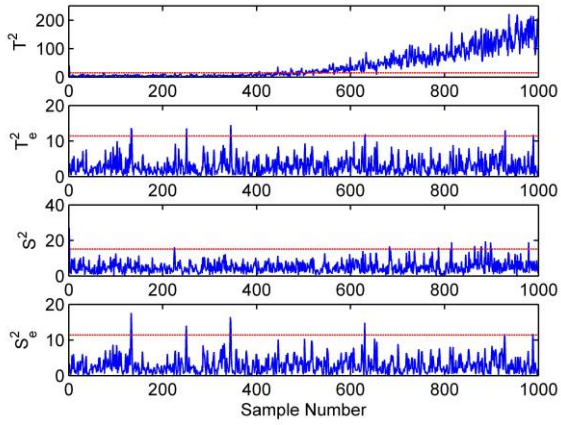
In the CSTR example, we have shown that although various disturbances lead to operating condition deviations, they are characterized by significantly different dynamic behaviors, thereby making alarms of the SFA-based model manifest discrepant patterns. Step changes, drifts, and random variations are typical types of disturbances that are of primary interests in process monitoring tasks. Therefore, in order to further assess the performance of the SFA-based monitoring approach, we examine the alarm patterns in the context of idealized disturbance. Table 5 lists the descriptions to three different disturbances, and monitoring results of SFA are shown in Figure 6.

Table 5. Different Typical Disturbances in the CSTR Example

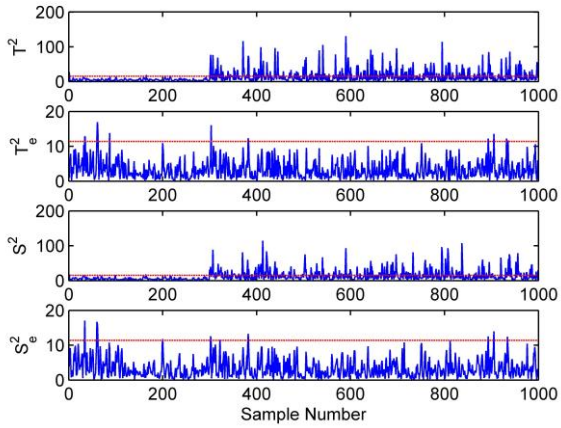
Disturbance	**Description**
Step change	Sample 1-300: T_f = 400 K Sample 301-1000: T_f = 401 K
Drift	Sample 1-300: T_f = 400 K Sample 301-1000: T_f linearly decreases to 399 K
Random variation	Sample 1-300: T_f = 400 K Sample 301-1000: T_f has random variations following $N(400, 0.1)$



(a)



(b)



(c)

Figure 6. Typical alarm patterns in ideal disturbance cases. (a) Step change, (b) drift, (c) random variation.

Next, we carry out a comprehensive discussion based on the alarm patterns presented in Figure 6. In the ideal case of step changes, the controller is considered to take effect and thus process dynamics remains normal, in spite of evident operating condition deviations. This mechanism well matches the alarm pattern in Figure 6(a), and has been discussed previously as an example of disturbance changing the operating condition only. In the ideal case of drift disturbances, assuming the drift to be sufficiently slow, the controller can still compensate the disturbance continuously. As shown in Figure 6(b), the alarm pattern of drifts is that, a slow deviation in the T^2 statistic occurs, implying that the process is deviating from its original operating condition, whereas the S^2 and S_e^2 statistics are below their thresholds, indicating no disruptions in process dynamics. For random variations, it takes some time for the controller to respond to such abrupt fluctuations, and hence there are always abnormal variations along with

operating condition deviations. This is in accordance with the alarm pattern in Figure 6(c) in which both T^2 and S^2 give persistent alarms. Table 6 provides a summary of typical alarm patterns of three ideal disturbances. Therefore, three types of disturbance have specific alarm patterns of their own, which are potentially useful for further identifying the fault type after the monitoring system issues an alarm.

Table 6. Typical Alarm Patterns of Some Idealized Disturbances

Monitoring Indices	**Step Changes**	**Drifts**	**Random Variations**
T^2 & T_e^2	either one jumps to a high value out of the control limit, or both	either one exceeds the control limit gradually, or both	either one exceeds the control limit with abrupt variations, or both
S^2 & S_e^2	both remain normal	both remain normal	either one exceeds the control limit with abrupt variations, or both

Tennessee Eastman Benchmark Process Case Studies

The Tennessee Eastman (TE) process[36] proposed by Eastman Chemical Company has been used as a well-established benchmark for evaluating different process monitoring and fault diagnosis approaches.[35,37,38] In this study, the plant-wide control strategy proposed by Lyman and Georgakis was adopted.[39] The simulation code as well as datasets can be downloaded from the website of Prof. Richard Braatz.

There are two blocks of variables in the TE process: the XMV block with 12 manipulated variables XMV(1-12) and the XMEAS block with 41 measured variables XMEAS(1-41). In this study, the input variables are chosen as XMV(1-11) and XMEAS(1-22), which are collected in a sampling interval of 3 min. The agitation speed is excluded from input variables because it is not manipulated.[39] To facilitate analysis of its feedback control mechanism in the sequel, Figure 7 depicts the plant-wide control structure of the TE process. Monitoring models based on DPCA, DICA, CVA and SFA are built using 500 normal samples. For DPCA and DICA, the lag order is selected as $d = 2$ according to the method described in Ref. 11. For a fair comparison, we use

the same number of lags $d = 2$ for SFA. Cross-validation is used to choose the number of PCs and ICs, and the criterion in (26) and (27) is used to choose the number of SFs. For CVA, parameters are determined based on the standing results in Ref. 18. Optimal numbers of LVs for different models are summarized in Table 7. For all methods, the corresponding 99% control limits are calculated. Next, three concrete examples of various disturbances are provided, in which disturbances are introduced at the 160th sample.

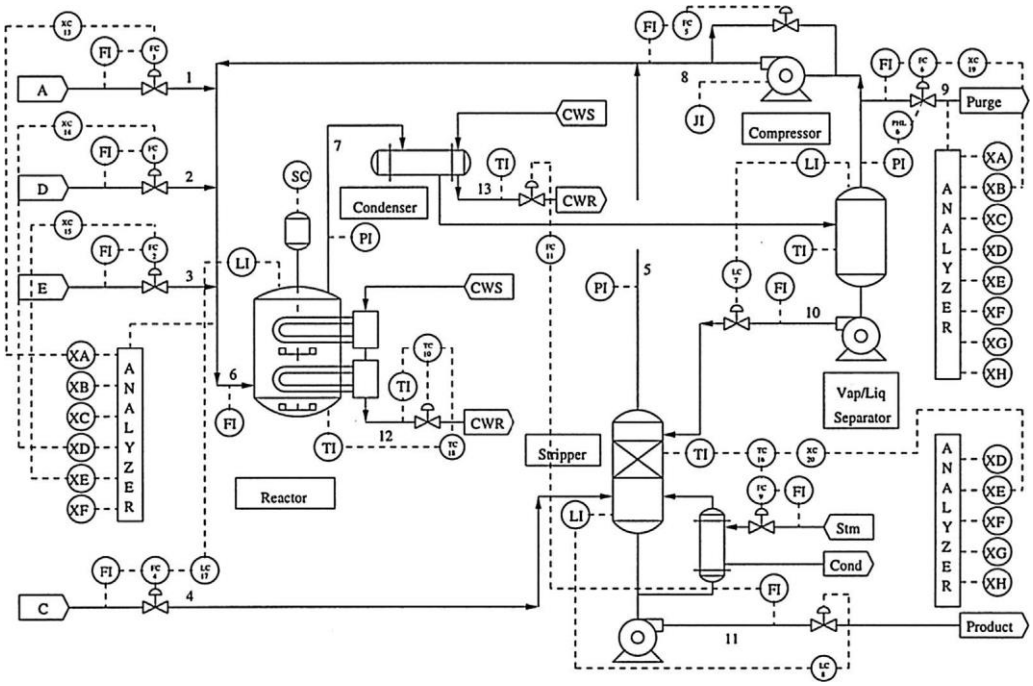


Figure 7. Flowsheet for the TE Process[39]

Table 7. Design Parameter Selection for DPCA, DICA, CVA and SFA

Model	DPCA	DICA	CVA	SFA
Design parameter	#PCs = 22	#ICs = 22	#CVs = 29	#SFs = 55

IDV(4): Step disturbance in reactor cooling water inlet temperature

In this case, a step change occurs in reactor cooling water inlet temperature. As shown in Figure 7, the reactor temperature (XMEAS(9)) is controlled by manipulating the reactor cooling water flow (XMV(10)) via a cascade controller. Their variation trends in this disturbance case are shown in Figure 8. When the disturbance occurs, the reactor temperature suddenly rises and the control loop makes a rapid response by increasing the reactor cooling water flow, which drives the reactor temperature to its set-point shortly afterwards. Although the increase in the reactor cooling water flow indicates an operating condition deviation, this disturbance could be well compensated.

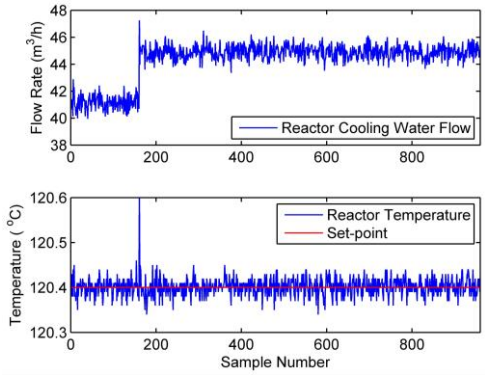
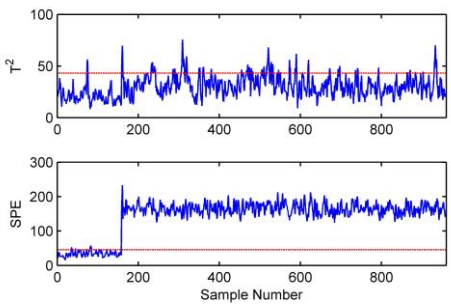
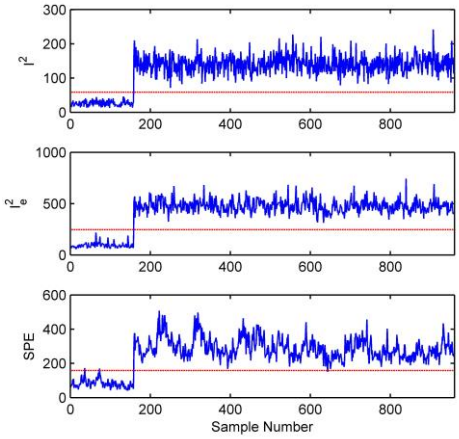


Figure 8. Reactor cooling water flow and reactor temperature in IDV(4)

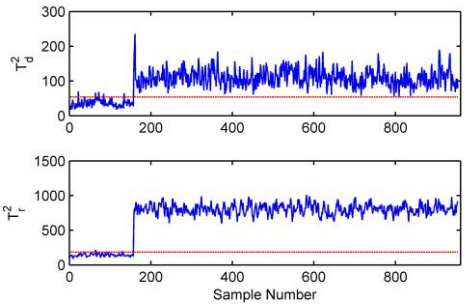
The process monitoring results for DPCA, DICA, CVA and SFA are shown in Figure 9. In the DPCA monitoring results, only the SPE statistic enables a successful detection, while all statistics of DICA and CVA take effect. For SFA, the T^2 statistic remains at a high value, while both S^2 and S_e^2 statistics reduce sharply toward normal values after the step disturbance. This implies that although the operating condition has changed after the step disturbance, the process dynamics is not affected because the cascade controller reduces the effect of the disturbance timely. This is in line with the physical analysis. In this sense, practitioners should be aware of the fact that this disturbance is essentially a normal change in the operating condition that deserves less attention. Nevertheless, DPCA-, DICA- and CVA-based monitoring approaches yield incomplete information about the control performance. Therefore, SFA is superior to the other three methods in filtering out this normal change from various process disturbances.



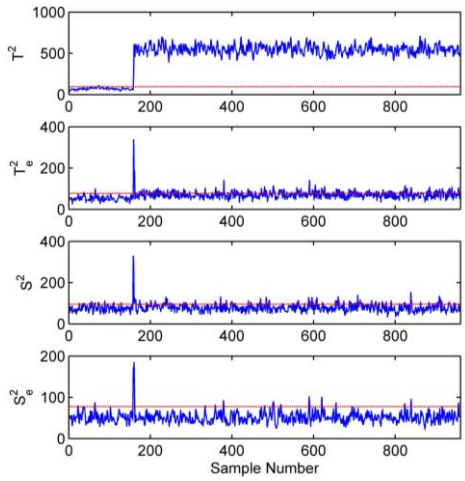
(a)



(b)



(c)



(d)

Figure 9. Monitoring results for a step disturbance in reactor cooling water inlet temperature. (a) DPCA, (b) DICA, (c) CVA, (d) SFA.

IDV(5): Step disturbance in condenser cooling water inlet temperature

In this case, a step disturbance occurs in condenser cooling water inlet temperature. The mechanism of this disturbance is slightly complicated compared to that of IDV(5), because the condenser locates in the major flow path and hence influences the final product flow. When the disturbance occurs at the 160th sample, the temperature of condenser underflow rises and a less

amount of products G and H are condensed, which brings a decrease of the stripper level and eventually a reduction in the stripper underflow, i.e., the final product flow. Therefore, the condenser cooling water flow (XMV(11)) is chosen as the manipulator of the stripper underflow (XMEAS(17)) through a PI controller.[39] As indicated by two dotted ellipsoids in Figure 10, after the disturbance happens, the production flow rate tends to decrease as a large number of measurements in the dotted ellipsoid stay below the set-point. The controller then makes compensating behaviors by increasing the condenser cooling water flow accordingly. However, the process is not stabilized rapidly. As shown in two solid ellipsoids in Figure 10, the majority of stripper underflow measurements between 250th sample and 300th sample drop below the set-point and the controller makes further adjustments. After 350th sample, the controlled variable gets eventually stabilized.

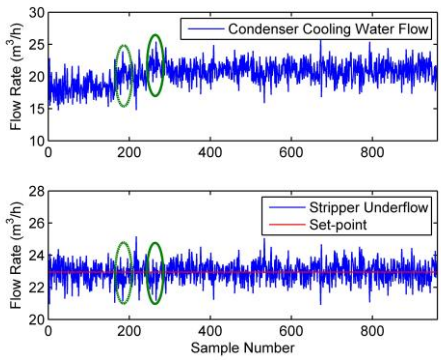
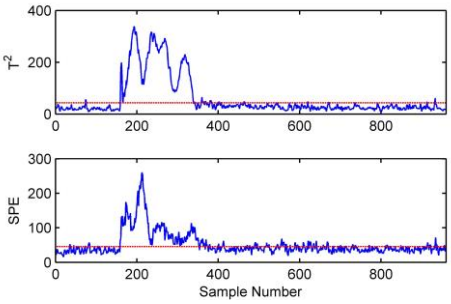
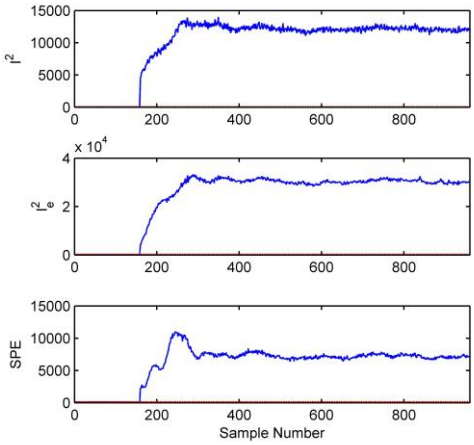


Figure 10. Condenser cooling water flow and stripper underflow in IDV(5)

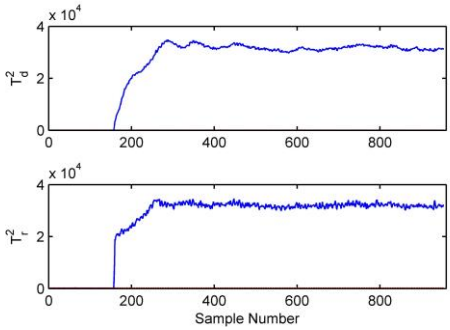
The process monitoring results are reported in Figure 11. DPCA makes a correct detection in both PC and residual subspaces at the beginning stage of the disturbance; however, it fails to reflect the deviation of condenser cooling water flow after the 350th sample, as already revealed in Figure 10. All indices of DICA and CVA keep violating the control limits after the 160th sample, indicating that an abnormal operating condition occurs all the time. In contrast to the other three approaches, the proposed method achieves more reliable monitoring results that are in consensus with the physical truth, since both T^2 and T_e^2 statistics sensitively detect the change of operating condition whereas the S^2 and S_e^2 statistics show that the control strategy is effective to stabilize the process after the 350th sample and there is no need for operators to take actions.


T2
SPE
Sample Number

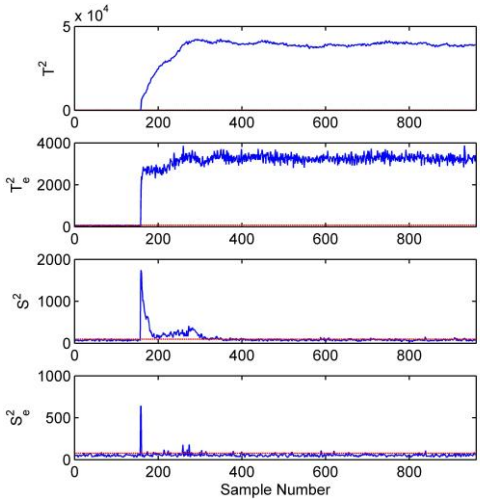
(a)


I2
x 10^4
I2e
SPE
Sample Number

(b)


x 10^4
T2d
x 10^4
T2r
Sample Number

(c)


x 10^4
T2
T2e
S2
S2e
Sample Number

(d)

Figure 11. Monitoring results for a step disturbance in condenser cooling water inlet temperature. (a) DPCA, (b) DICA, (c) CVA, (d) SFA.

IDV(10): Random variation in C feed temperature

In this scenario, the C feed temperature is changed with a random variation. Because C feed directly enters into the stripper, the random variation in C feed temperature results in a persistent fluctuation of stripper temperature after the 160th sample. Consequently, the stripper steam valve (XMV(9)) is manipulated in a different manner throughout to alleviate the changes of the stripper temperature (XMEAS(18)), as shown in Figure 12, thereby affecting both operating condition and process dynamics.

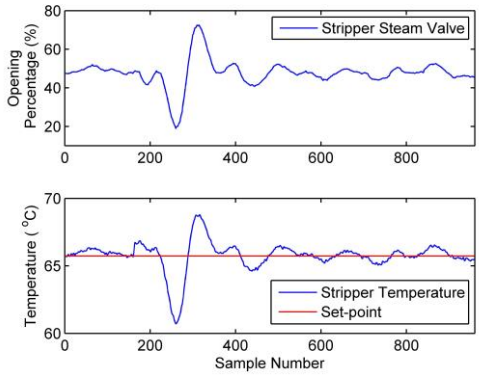
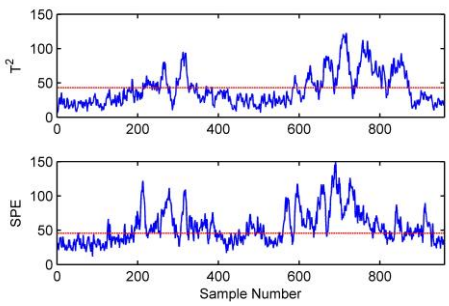
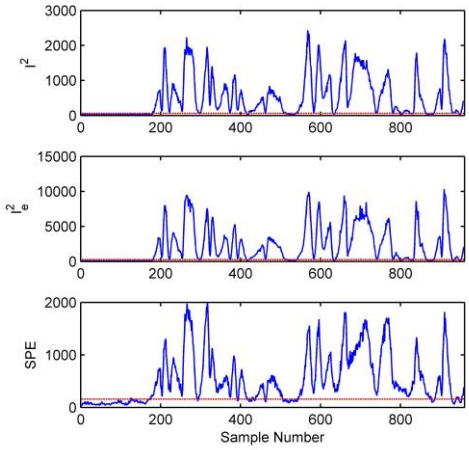


Figure 12. Stripper steam valve opening percentage and stripper temperature in IDV(10)

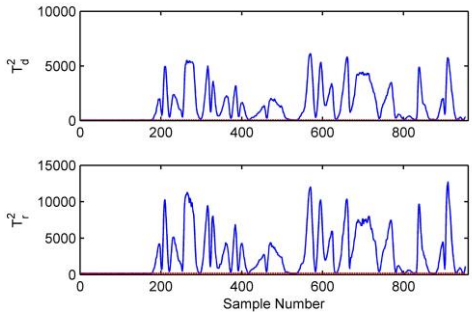
Monitoring results are given in Figure 13. DPCA can detect this disturbance from about the 160th sample. However, there are still a large number of samples below the control limits, thereby delivering confusing messages. In comparison with the DPCA-based monitoring, DICA, CVA and SFA give alarms correctly up to the end. Their statistics well reveal the variation pattern induced by this random disturbance. In addition, the S^2 index of SFA signals this disturbance as an alarm that practically influences the control performance. This fault is noteworthy for industrial practitioners, which differs from the previous two cases.



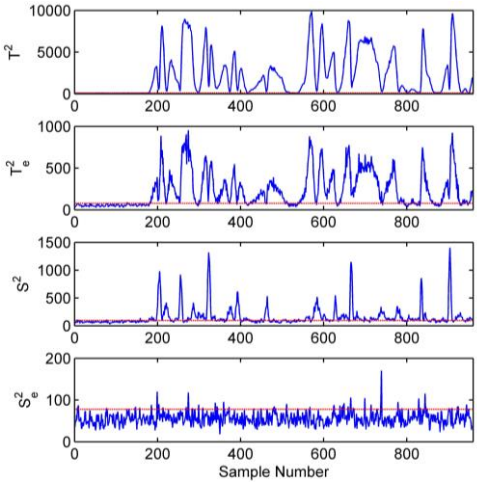
(a)



(b)



(c)



(d)

Figure 13. Monitoring results for random variation in C feed temperature. (a) DPCA, (b) DICA, (c) CVA, (d) SFA.

Process dynamics in all 21 disturbance cases

In 21 disturbance cases of the TE process, it is known that IDV(3), IDV(9) and IDV(15) lead to insignificant deviations from the operating condition, and hence are difficult to identify by generic monitoring methods.[37] However, dynamic behaviors in 21 disturbance cases are unknown. Therefore, we examine the system dynamics in 21 disturbance cases according to the monitoring

results provided by SFA. The disturbance is introduced at the 160th sample in each case. According to the monitoring indices provided by SFA[1], 21 disturbances can be qualitatively categorized into three classes, as reported in Table 8.

Table 8. Dynamic Properties of All 21 Disturbances of TE Process

Category	**Disturbance Number**	**Disturbance Variable**	**Type**
Disturbances without evident dynamics anomalies	IDV(3)	D feed temperature	Step
	IDV(9)	D feed temperature	Random variation
	IDV(15)	Condenser cooling water valve	Sticking
	IDV(21)	The valve fixed at steady position	Constant position
Disturbances with short-lived dynamics anomalies	IDV(1)	A/C feed ratio, B composition constant	Step
	IDV(2)	B composition, A/C ratio constant	Step
	IDV(4)	Reactor cooling water inlet temperature	Step
	IDV(5)	Condenser cooling water inlet temperature	Step
	IDV(7)	C header pressure loss-reduced availability	Step
Disturbances with persistent dynamics anomalies	IDV(6)	A feed loss	Step
	IDV(8)	A, B, and C feed composition	Random variation
	IDV(10)	C feed temperature	Random variation
	IDV(11)	Reactor cooling water inlet temperature	Random variation
	IDV(12)	Condenser cooling water inlet temperature	Random variation
	IDV(13)	Reaction kinetics	Slow drift
	IDV(14)	Reactor cooling water valve	Sticking
	IDV(16)	Unknown	Unknown
	IDV(17)	Unknown	Unknown
	IDV(18)	Unknown	Unknown
	IDV(19)	Unknown	Unknown
	IDV(20)	Unknown	Unknown

IDV(3), IDV(9), IDV(15) and IDV(21) belong to the first category without obvious dynamics

[1] Due to the limited space as well as the ease of implementation of the SFA-based method, figures are omitted and only results are given.

up to the end. The others cause dynamics anomalies to various extent, in which IDV(1), IDV(2), IDV(4), IDV(5) and IDV(7) lead to short-lived abrupt variations that can be well mitigated later on. In these occasions, the approximated numbers of samples required for compensating short-lived dynamics anomalies are tabulated in Table 9. And the remaining disturbances yield consistent disruptions in process dynamics. It can be seen that, the plant-wide control strategy proposed by Lyman and Georgakis[35] can resist the effects of most step disturbances after a period of time, whereas most random variations still impose some negative impacts on the process dynamics.

Table 9. Approximated Numbers of Samples Required for Recovery from Short-lived Dynamics Anomalies

	IDV(1)	**IDV(2)**	**IDV(4)**	**IDV(5)**	**IDV(7)**
Sample No.	290	240	5	190	290

Conclusions and Future Considerations

In this article, we have proposed an SFA-based approach for the simultaneous monitoring of operating conditions and process dynamics. The slow features are captured using the SFA algorithm with clear temporal interpretations. A new strategy to select the number of slow features is also proposed by analyzing the slowness of input reconstructions. Four process monitoring indices are developed based on the derived slow features and their physical meanings. The proposed method can well distinguish real faults that affect control performances from normal changes of operating conditions. It is of great importance in practical scenarios because unnecessary alarms can be removed and monitoring schemes are able to convey more useful information to process operators. Two simulated experiments show the applicability of the proposed monitoring approach based on SFA. A future step is to develop related fault diagnosis methods, for example, faulty sensor identification[40,41], and quality-relevant monitoring solutions[10] for the SFA monitoring framework.

Acknowledgments

This work is supported by the National Basic Research Program of China (2012CB720505),

the National Natural Science Foundation of China (61433001) and Tsinghua University Initiative Scientific Research Program. Xiaolin Huang and Johan A.K. Suykens acknowledge support from ERC AdG A-DATADRIVE-B, IUAP-DYSCO, GOAMANET, OPTEC, and BIL with Tsinghua University. We would also like to thank our anonymous reviewers who have provided helpful comments.

Literature Cited

1. Wise BM, Gallapher NB. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*. 1996;6:329-348.
2. Qin SJ. Statistical process monitoring: Basics and beyond. *Journal of Chemometrics*. 2003;17:480-502.
3. Qin SJ. Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*. 2012;36:220-234.
4. Macgregor JF, Cinar A. Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. *Computers & Chemical Engineering*. 2012;47:111-120.
5. Jackson JE, Mudholkar GS. Control procedures for residuals associated with principal component analysis. *Technometrics*. 1979;21:341-349.
6. Jackson JE. Principal components and factor analysis: part I-principal components. *Journal of Quality Technology.* 1980;12:201-213.
7. Kano M, Tanaka S, Hasebe S, Hashimoto I, Ohno H. Monitoring independent components for fault detection. *AIChE Journal*. 2003;49:969-976.
8. Lee JM, Yoo CK, Lee IB. Statistical process monitoring with independent component analysis. *Journal of Process Control*. 2004;14:467-485.
9. Lee JM, Qin SJ, Lee IB. Fault detection and diagnosis based on modified independent component analysis. *AIChE Journal*. 2006;52:3501-3514.
10. Qin SJ, Zheng Y. Quality-relevant and process relevant fault monitoring with concurrent projection to latent structures. *AIChE Journal*. 2013;59:496-504.
11. Ku W, Storer RH, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*. 1995;30:179-196.
12. Chen J, Liu KC. On-line batch process monitoring using dynamic PCA and dynamic PLS models. *Chemical Engineering Science*. 2002;57:63-75.

13. Shang C, Huang X, Suykens JAK, Huang D. Enhancing dynamic soft sensors based on DPLS: A temporal smoothness regularization approach. *Journal of Process Control*. 2015;28:17-26.

14. Lee JM, Yoo CK, Lee IB. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science*. 2004;59:2995–3006.

15. Negiz A, Cinar A. PLS, balanced, and canonical variate realization techniques for identifying VARMA models in state space. *Chemometrics and Intelligent Laboratory Systems*. 1997;38:209-221.

16. Dunia R, Qin SJ. Subspace approach to multidimensional fault identification and reconstruction. *AIChE Journal*. 1998;44:1813-1831.

17. Negiz A, Cinar A. Statistical monitoring of multivariable dynamic processes with state-space models. *AIChE Journal*. 1997;43:2002-2020.

18. Russell EL, Chiang LH, Braatz RD. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems.* 2000;51:81-93.

19. Wiskott L, Sejnowski TJ. Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*. 2002;14:715-770.

20. Körding KP, Kayser C, Einhäuser W, König P. How are complex cell properties adapted to the statistics of natural stimuli?. *Journal of Neurophysiology*. 2004;91:206-212.

21. Berkes P, Wiskott L. Slow feature analysis yields a rich repertoire of complex cell properties. *Journal of Vision*. 2005;5:579-602.

22. Blaschke T, Zito T, Wiskott L. Independent slow feature analysis and nonlinear blind source separation. *Neural Computation*. 2007;19:994-1021.

23. Sprekeler H, Zito T, Wiskott L. An extension of slow feature analysis for nonlinear blind source separation. *Journal of Machine Learning Research*. 2014;15:921-947.

24. Zhang Z, Tao D. Slow feature analysis for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012;34:436-450.

25. Wu C, Du B, Zhang L. Slow feature analysis for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 2014;52:2858-2874.

26. Zhang L, Wu C, Du B. Automatic radiometric normalization for multitemporal remote sensing imagery with iterative slow feature analysis. *IEEE Transactions on Geoscience and Remote Sensing*. 2014;52:6141-6155.

27. Turner R, Sahani M. A maximum-likelihood interpretation for slow feature analysis. *Neural Computation*.

2007;19:1022-1038.

28. Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. *Neural Networks*. 2000;13:411-430.
29. Blaschke T, Berkes P, Wiskott L. What is the relation between slow feature analysis and independent component analysis?. *Neural computation*. 2006;18: 2495-2508.
30. Casella G, Berger RL. Statistical inference (Vol. 2). Duxbury: Pacific Grove, CA, 2002.
31. MacGregor JF, Jaeckle C, Kiparissides C, Koutoudi, M. Process monitoring and diagnosis by multiblock PLS methods. *AIChE Journal*, 1994;40:826-838.
32. Li G, Qin SJ, Ji Y, Zhou D. Reconstruction based fault prognosis for continuous processes. *Control Engineering Practice*. 2010;18:1211-1219.
33. Valle S, Li W, Qin SJ. Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. *Industrial and Engineering Chemistry Research*. 1999;38:4389-4401.
34. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*. 2001;58:109-130.
35. Chiang LH, Russell EL, Braatz RD. Fault detection and diagnosis in industrial systems. London: Springer, 2001.
36. Downs JJ, Vogel EF. A plant-wide industrial process control problem. *Computers & Chemical Engineering*. 1993;17:245-255.
37. Yin S, Ding SX, Haghani A, Hao H, Zhang P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control*. 2012;22:1567-1581.
38. Jiang B, Huang D, Zhu X, Yang F, Braatz RD. Canonical variate analysis-based contributions for fault identification. *Journal of Process Control*. 2015;26:17-25.
39. Lyman PR, Georgakis C. Plant-wide control of the Tennessee Eastman problem. *Computers & Chemical Engineering*. 1995;19:321-331.
40. Du M, Scott J, Mhaskar P. Actuator and sensor fault isolation of nonlinear process systems. *Chemical Engineering Science*. 2013;104:294-303.
41. Du M, Mhaskar P. Isolation and handling of sensor faults in nonlinear systems. *Automatica*. 2014;50:1066–1074.

Appendix A: The statistical uncorrelation between different $\dot{s}_j$

The covariance matrix of between $\dot{\boldsymbol{s}}$ can be calculated as

$$\left\langle \dot{\boldsymbol{s}}\dot{\boldsymbol{s}}^{\mathrm{T}} \right\rangle_t = \mathbf{P}^{\mathrm{T}} \left\langle \dot{\boldsymbol{z}}\dot{\boldsymbol{z}}^{\mathrm{T}} \right\rangle_t \mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{P}\boldsymbol{\Omega}\mathbf{P}^{\mathrm{T}}\mathbf{P} = \boldsymbol{\Omega} = \mathrm{diag}\{\omega_1, \cdots, \omega_{m(d+1)}\} . \tag{A1}$$

Therefore $\dot{s}_i$ and $\dot{s}_j$ $(i < j)$ are uncorrelated.

Appendix B: Proof of Proposition 1

For the calculation of slowness, x_j should first be normalized to unit variance:

$$x_j^{\mathrm{norm}} \doteq \frac{x_j}{\sqrt{\left\langle x_j^2 \right\rangle_t}} = \frac{\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{s}}{\| \boldsymbol{r}_j \|_2} . \tag{B1}$$

Then its slowness can be obtained as:

$$\Delta(x_j) \doteq \left\langle (\dot{x}_j^{\mathrm{norm}})^2 \right\rangle_t = \frac{\boldsymbol{r}_j^{\mathrm{T}} \left\langle \dot{\boldsymbol{s}}\dot{\boldsymbol{s}}^{\mathrm{T}} \right\rangle_t \boldsymbol{r}_j}{\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{r}_j} = \frac{\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{\Omega}\boldsymbol{r}_j}{\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{r}_j} \tag{B2}$$

where $\boldsymbol{\Omega} = \mathrm{diag}\{\omega_1, \cdots, \omega_{m(d+1)}\}$, and the last equality holds because $\left\langle \dot{\boldsymbol{s}}\dot{\boldsymbol{s}}^{\mathrm{T}} \right\rangle_t = \boldsymbol{\Omega}$ proved in Appendix A. By replacing ω_j with $\Delta(s_j)$, (B2) can further be decomposed as:

$$\Delta(x_j) = \frac{\sum_i r_{ji}^2 \Delta(s_i)}{\sum_i r_{ji}^2} = \sum_i \frac{r_{ji}^2}{\sum_i r_{ji}^2} \Delta(s_i) \doteq \sum_i \alpha_i \Delta(s_i) \tag{B3}$$

where $\alpha_i = r_{ji}^2 / \sum_i r_{ji}^2$ and $\sum_i \alpha_i = 1$. This completes the proof.

Appendix C: Proof of Proposition 2

For a given index *k* satisfying $\Delta(s_k) > \Delta(x_j)$, we show that

$$
\begin{aligned}
&\Delta(s_k) > \frac{\sum_i r_{ji}^2 \Delta(s_i)}{\sum_i r_{ji}^2} \\
&\Rightarrow \sum_i r_{ji}^2 \Delta(s_k) > \sum_i r_{ji}^2 \Delta(s_i) \\
&\Rightarrow \sum_{i \neq k} r_{ji}^2 \Delta(s_k) + r_{jk}^2 \Delta(s_k) > \sum_{i \neq k} r_{ji}^2 \Delta(s_i) + r_{jk}^2 \Delta(s_k) \\
&\Rightarrow \Delta(s_k) \sum_{i \neq k} r_{ji}^2 > \sum_{i \neq k} r_{ji}^2 \Delta(s_i) \\
&\Rightarrow r_{jk}^2 \Delta(s_k) \sum_{i \neq k} r_{ji}^2 > r_{jk}^2 \sum_{i \neq k} r_{ji}^2 \Delta(s_i) \\
&\Rightarrow \frac{r_{jk}^2 \Delta(s_k)}{\sum_{i \neq k} r_{ji}^2 \Delta(s_i)} > \frac{r_{jk}^2}{\sum_{i \neq k} r_{ji}^2} \\
&\Rightarrow 1 + \frac{r_{jk}^2 \Delta(s_k)}{\sum_{i \neq k} r_{ji}^2 \Delta(s_i)} > 1 + \frac{r_{jk}^2}{\sum_{i \neq k} r_{ji}^2} \\
&\Rightarrow \frac{\sum_i r_{ji}^2 \Delta(s_i)}{\sum_{i \neq k} r_{ji}^2 \Delta(s_i)} > \frac{\sum_i r_{ji}^2}{\sum_{i \neq k} r_{ji}^2} \\
&\Rightarrow \frac{\sum_i r_{ji}^2 \Delta(s_i)}{\sum_i r_{ji}^2} > \frac{\sum_{i \neq k} r_{ji}^2 \Delta(s_i)}{\sum_{i \neq k} r_{ji}^2} \\
&\Rightarrow \Delta(x_j) > \Delta(x_j^{\mathrm{rec}})
\end{aligned}
\tag{C1}
$$

This completes the proof.