Neural Comput & Applic (2010) 19:677–689
DOI 10.1007/s00521-009-0324-5

ORIGINAL ARTICLE

Process disturbance identification through integration of spatiotemporal ICA and CART approach

**Chih-Chou Chiu · Shin-Ying Hwang ·
Deborah F. Cook · Yuan-Ping Luh**

Received: 29 October 2008 / Accepted: 9 December 2009 / Published online: 29 December 2009

Abstract Many studies have been conducted over the past several years evaluating the integrated use of statistical process control (SPC) and engineering process control (EPC). The majority of these studies reported that combining SPC with EPC outperforms the use of only SPC or EPC. Basically, the former aims to rapidly detect assignable causes and time points for abnormalities that take place during process; and the latter is a method in which input variables are adjusted against process outputs through a feedback control mechanism. Although combining SPC with EPC can effectively detect time points when abnormalities occur during process, their combination can also cause an increased occurrence of false alarms when autocorrelation is present in the process. In this study, to increase the accuracy of process disturbance identification, we propose the integration of spatiotemporal independent component analysis (stICA) with the classification and regression tree (CART) approach to improve our capability to identify process disturbances and recognize shifts in the correlated process parameters. The integration of the CART methodology results in the development of decision rules that can provide valuable information related to the impact of variation in process variable values. These decision rules can provide an increased understanding of process behavior and useful information for process control. For comparison, the integration of traditional principle component analysis (PCA) with CART (called PCA-CART), ICA with CART (called ICA-CART) and cumulative sum chart approaches were applied to evaluate the identification capability of the proposed approach. As the results reveal, the proposed approach is more effective for monitoring correlated process.

Keywords Process disturbance · Statistical process control · Engineering process control · Spatiotemporal independent component analysis · Classification and regression tree

C.-C. Chiu (✉)
Institute of Commerce Automation and Management,
National Taipei University of Technology, Taipei, Taiwan
e-mail: chih3c@ntut.edu.tw

S.-Y. Hwang
Department of Management Information System,
National Chengchi University, Taipei, Taiwan
e-mail: 96356511@nccu.edu.tw

D. F. Cook
Department of Business Information Technology,
Virginia Polytechnic Institute and State University,
Blacksburg, VA, USA
e-mail: dcook@vt.edu

Y.-P. Luh
Department of Mechanical Engineering,
National Taipei University of Technology,
Taipei, Taiwan
e-mail: yuan@ntut.edu.tw

1 Introduction

In today's highly competitive global market, continuous enhancement of product quality is key for businesses in their efforts to acquire and maintain a competitive advantage. Production processes are typically monitored in an effort to maintain and control the efficient production of high-quality products. Process monitoring refers to the procedure in which all data are recorded using various techniques without altering a product's quality characteristics, thereby allowing the identification of process disturbances. In the field of quality control, statistical process control (SPC) and engineering process control (EPC) are two tools that are often applied to process monitoring and control. The former aims to rapidly detect assignable causes in a process at their time of occurrence; and the latter is a method in which input variables are adjusted against process outputs through a feedback control mechanism, thereby narrowing the gap between process output and target. Although combining SPC with EPC can provide a means to effectively detect time points when abnormalities occur during process, their combination can also cause an increased occurrence of false alarm signals when auto-correlation exists in the process data. As a result, the development of an on-line system to detect and control abnormalities for correlated processes has become a critical research issue in the quality control field.

Independent component analysis (ICA) is a signal linear decomposition technique that can extract underlying factors from multivariate statistical data [10–13, 16–18, 20–22, 24–45]. Given a set of multivariate data, ICA aims at finding a linear decomposition where statistical independence is maximized over space (sICA) or over time (tICA). Basically, sICA was mainly applied to the field of image signal analysis where mutually independent source signal images and a corresponding dual set of unconstrained time courses are interested. Contrastingly, tICA was used in the context of blind source separation, since a set of temporally independent time courses is sought for in such applications. Spatiotemporal ICA (stICA) is a method that permits a trade-off between the mutual independence of spatial underlying variables and the mutual independence of their corresponding time courses [42]. Observations of data from a correlated production process can be viewed as mixture data that contains noise, process disturbance terms and auto-correlation, and consequently can be approached as a set of multivariate data. Since simply using SPC and EPC to undertake process control cannot obtain the ideal results in a correlated process, we propose the combination of SPC/EPC with stICA to develop an improved technique for online process control. The combined SPC/EPC/stICA algorithm will convert process data into separate independent signals using stICA in the first stage. In the second stage, the separated independent signals are used to construct, through the classification and regression tree (CART) method, a powerful on-line model that can identify process disturbance levels. Various types of pattern identification tools have already been successfully developed, including conventional statistical methods, nonparametric methods, and artificial intelligence methods. The CART algorithm developed by Breiman et al. [4] has been widely discussed and successfully applied and is the method we selected for use in this research study. Basically, it uses a binary segmentation process to analyze a huge data set. Through a recursive procedure, a number, N training samples are segmented into a number, M known pattern types according to prediction variables and related division conditions. A set of rules is then produced and used to identify process disturbances in new samples. In this study, we propose the application of the CART technique to the independent signals developed using stICA to establish an improved process control model and illustrate the advantages of this method with respect to increased accuracy of disturbance level identification. For comparison, the integration of traditional principle component analysis (PCA) with CART (called PCA-CART), ICA with CART (called ICA-CART) and cumulative sum (CUSUM) chart approaches were applied to evaluate the identification capability of the proposed approach. As the results reveal, the proposed approach is more effective for monitoring correlated process.

The rest of this paper is organized as follows. We will briefly review the literature of SPC/EPC, stICA and CART techniques in modeling process control problems in Sect. 2. Section 3 gives a brief outline of stICA and CART. The development as well as the analytic results of process control models using stICA with CART are presented in Sect. 4. Finally, Sect. 5 addresses the conclusions.

2 Literature review

2.1 SPC/EPC in process control

Traditional SPC techniques developed for manufacturing operations are only applicable when the manufacturing process data does not violate the basic assumption of statistical independence. When the value of a particular parameter is dependent on its previous value, autocorrelation is present in the data. In such a case, the strength of traditional control charts to identify the presence of assignable causes is significantly weakened. Much research has been directed at solving this problem through the use of time series modeling techniques [1, 6, 37, 44]. Alwan and Roberts [3] fitted a set of correlated data with an adequate time series model and applied a common-cause control (CCC) chart or a special-cause control (SCC) chart to the stream of residuals from the time series model. Montgomery and Mastrangelo [35] evaluated the properties of the EWMA chart of residuals for process data from IMA(1,1) time series model. Their results showed that under very specific conditions, an EWMA chart of residuals from an IMA(1,1) model performs reasonably well. Lu and Reynolds [30] investigated four different charts (an EWMA chart of the logs of the squared residuals, a Shewhart chart of the squared residuals, a moving range chart, and an EWMA chart of residuals) for their ability to detect a change in the variance of a correlated process. They found that none of the charts performed well for all parameter combinations.

Several authors [5, 8, 14, 31, 36, 38] have recommended the integrated use of SPC and EPC. Most of these studies showed that the assignable causes of process disturbance can be effectively identified by SPC techniques [37, 40]. However, these techniques have seen very limited application in practice because they are time-consuming, apply only in very specific conditions, or are difficult to implement in practice.

2.2 Spatiotemporal ICA in process control

The stICA currently remains an active field of research. Few applications of stICA are reported in literatures, such as image de-noising and feature extraction [9]. However, to date, the stICA model has not been used in process control. Only few applications of ICA in process monitoring have been reported in recent literature. A spectral ICA approach was developed by Xia [47] and Xia and Howell [46] to transform the process measurements from the time domain into the frequency domain and to identify major oscillations. Kano et al. [23] have successfully demonstrated the idea of process monitoring according to the observation of ICs instead of the original measurements. In his work, a set of devised SPC charts have been developed effectively for each IC. Lee et al. [26, 27] investigated the utilization of kernel density estimation to define the control limits of the ICs that do not satisfy Guassian distribution. To monitor batch processes that combine ICA and kernel estimation, Lee et al. [25] extended their original method to multi-way ICA. Their simulation results clearly demonstrate the power and advantages of ICA monitoring in comparison with PCA monitoring. Yao et al. [48] used ICs to represent the state space and subsequently applied the method to process monitoring. In addition to the ICA technique, there are some dynamic methods that have been developed recently and been used in process control. For more details, reader can refer to [2, 15, 28, 29, 49].

3 Research methodology

In this study, we used stICA to convert observed process data into separate independent signals. These separated signals were in turn integrated into the CART approach in the second stage of our technique to construct an on-line model that could identify process disturbance levels. A research framework for this study is presented in Fig. 1, and the stICA method and CART technique applied in this study are described as follows.

3.1 Spatiotemporal ICA

Spatiotemporal ICA is a linear decomposition technique that maximizes the degree of independence over space as

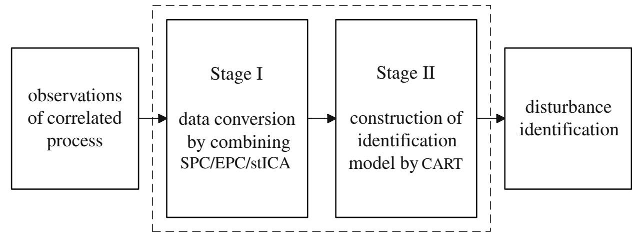

Fig. 1 Research structure

well as over time, without necessarily producing independence in either space or time. In other words, stICA permits a trade-off between the mutual independence of signals and the mutual independence of their corresponding time courses [43].

Let $\mathbf{X} = [X_1, X_2, \ldots, X_m]^T$ be an input matrix of size $m \times n$, $m \leq n$, consisting of observed mixture signals X_i of size $1 \times n$, $i = 1, 2, \ldots, m$. Suppose that the singular value decomposition (SVD) of X is given by $X = UDV^T$, where $U \in R^{m\times n}$ corresponds to eigenarrays, $V \in R^{n\times N}$ is associated with eigengenes, and D is a diagonal matrix containing singular values. Then, following the notations in [43], we can define $\tilde{\mathbf{X}} = UDV^T = \tilde{U}\tilde{V}^T$as $\tilde{U} = UD^{1/2}$ and $\tilde{V} = VD^{1/2}$.

Given $\tilde{\mathbf{X}} = \tilde{U}\tilde{V}^T$, stICA is trying to find the decomposition, $\tilde{\mathbf{X}} = S\Lambda P^T$, where S is a $m \times n$ matrix with a set of n independent m dimensional arrays, P is an $N \times n$ matrix of mutually independent sequences, and Λ is a diagonal scaling matrix. Under the condition of $\tilde{\mathbf{X}} = \tilde{U}\tilde{V}^T$, there exist two $n \times n$ un-mixing matrices, W_S and W_P, such that $S = \tilde{U}W_S$ and $P = \tilde{V}W_P$. Then, if $W_S\Lambda W_P^T = I$, the following relation holds.

$$\tilde{\mathbf{X}} = S\Lambda P^T = \tilde{U}W_S\Lambda(\tilde{V}W_P)^T = \tilde{U}\tilde{V}^T \quad (1)$$

We can estimate the W_S and W_P by maximizing an objective function associated with spatial and temporal entropies at the same time. That is, the objective function for stICA has the following form

$$\text{Entropy (stICA)} = \alpha \times \text{Entropy (sICA)} + (1-\alpha) \times \text{Entropy (tICA)} \quad (2)$$

where α defines the relative weighting for spatial entropy and temporal entropy. For simplifying the setting of α, 0.5 is used in this study. More details on stICA can be found in [42].

3.2 Classification and regression tree

Classification and regression tree, a non-parametric statistical procedure introduced by Breiman et al. [4], is primarily used as a classification tool, where the objective is to classify an object into two or more populations. As the name suggests, CART is a single procedure that can be used to analyze either categorical or continuous data using the same technology. The methodology outlined in Breiman et al. [4] can be summarized into three stages. The first stage involves growing the tree using a recursive partitioning technique to select variables and split points using a splitting criterion. Several criteria are available for determining the splits, including gini, twoing and ordered twoing. A more detailed description of these criteria can be found in Breiman et al. [4]. In addition to selecting the primary variables, surrogate variables that are closely related to the original splits and that may be useful in classifying observations that have missing values for the primary variables can also be identified and selected.

After a large tree is identified, the second stage of the CART methodology uses a pruning procedure that incorporates a minimal cost complexity measure. The result of the pruning procedure is a nested subset of trees starting from the largest tree grown and continuing the process until only one node of the tree remains. Cross-validation or a testing sample will be used to provide estimates of future classification errors for each sub tree. Cross-validation is used when only small numbers of data points are available in building the CART models.

The last stage of the methodology is to select the optimal tree, which corresponds to a tree yielding the lowest cross-validated or testing set error rate. Although the optimal tree has the smallest testing error, it could be identified as unstable because of its large size. To avoid this instability, trees with smaller sizes but comparable in accuracy (i.e., within one standard error) will be chosen as an alternative. This process is referred to as the one standard error rule and can be tuned to obtain trees of varying sizes and complexity. A measure of variable importance can be achieved by observing the drop in the error rate when another variable is used instead of the primary split. Basically, the more frequent a variable appears as a primary or surrogate split, the higher the importance score assigned. Please refer to Breiman et al. [4] and Steinberg and Colla [41] for more details regarding the model building process of CART.

4 Integration of CART with stICA for disturbance detection

In order to demonstrate the effectiveness of our proposed approach, we use two commonly encountered disturbances [34, 38–41], the step-change disturbance and the linear disturbance, as our illustrated examples.

4.1 Step-change disturbance

Typically, a step-change disturbance can be described as:

$$D_t = \begin{cases} L, & t \geq t_s \\ 0, & t < t_s \end{cases} \tag{3}$$

where D_t is the disturbance at time t, L is the level of the step-change disturbance and t_s is the time of the introduction of the disturbance into the process. By assuming B as a backward shift operator (e.g., $B^2 D_t = D_{t-2}$) and β_t as a random variable with value zero most of the time except when the disturbance has occurred, we can transform the (3) as follows.

$$D_t = \frac{\beta_t}{1-B} \qquad \beta_t = \begin{cases} L, & t = t_d \\ 0, & t \neq t_d \end{cases} \tag{4}$$

According to MacGregor [32], applying the first-order model to represent the manufacturing process is applicable. Consequently, we assume that the underlying process is modeled as a first-order process. Moreover, we adopt the typical first-order integrated moving average (i.e., IMA(1,1)) process to represent random noise. The detailed description and definition of first-order processes and IMA(1,1) can be found in [7, 8, 31, 33, 39].

Consequently, a first order process with IMA(1,1) noise used in this paper can be described by

$$y_{t+1} = \frac{q}{1-pB} X_t + d_{t+1} + D_{t+1} \tag{5}$$

$$d_{t+1} = \frac{1-\theta B}{1-B} \varepsilon_{t+1} \tag{6}$$

where y_{t+1} is the output deviation from the target at time $t+1$, p and q are fixed parameters for the first order system, with $p + q = 1$, X_t is the control variable's deviation from the nominal value at time t, d_{t+1} is the noise at time $t+1$ as it follows the IMA(1,1) process, θ is a parameter estimated from the available observations, and ε_{t+1} is white noise, assumed to be normally distributed with mean zero and variance of σ_ϵ^2.

After setting parameters p, q and θ to 0.8, 0.2 and 0.9 (typical values for a manufacturing environment), a step-change disturbance (SD) dataset is first generated to evaluate the performance of the proposed stICA-CART approach for identifying process disturbance level. In the dataset, four subsets (SD1–SD4) are simulated from (5, 6). In each SD subset, 400 non-disturbed data vectors and 400 step-change disturbance data vectors are included. The 400 step-change disturbance data vectors in Subset SD1 are generated from (4–6) at $L = 1$. Likewise, the data vectors in Subsets SD2–SD4 are produced at $L = 5$, 10 and 15, respectively. It is noted that we assume the disturbance took place at 201 $t_d = 201$, when 400 process series was generated. The dataset were partitioned into training and testing parts, each training part representing roughly 70%

of the dataset. The first part (training part) was used to estimate stICA-CART process disturbance identification model (PDIM) parameters, and the second part was used to evaluate the accurate performance of the model. Figure 2 shows an example of process monitoring subsets with $L = 10$ and 15. Every subset was put into a specific category according its reference values such as 0, 1, 2, 3 and 4.

To tie in with monitoring techniques used in practice, in this study we assume that when data correspond to the first-order model and their noise term has the properties of IMA(1,1), the proportional integral (PI) controller (as in 7) will be used to assume the role of EPC, that is, to undertake adjustments to the process. The PI controller in this study is a controller built with the minimum mean squared error (MMSE) for an objective.

$$X_t = -\frac{(1-\theta)p}{q}\left(y_t + \frac{1-p}{p}\sum_{i=-\infty}^{t} y_i\right) \quad (7)$$

If (7) is substituted into (5) (i.e., adjust process data with a PI or MMSE controller), the original process data will be converted into a sum of process data y_t at the present time point and weighted historical process data $(1-\theta)\sum_{i=1}^{t-1} y_i$, as shown in (8).

$$d_t + D_t = y_t + (1-\theta)\sum_{i=1}^{t-1} y_i \quad (8)$$

Besides, if (4) is substituted into (8), process data converted by applying SPC and EPC can then be obtained, as shown in the following:

$$y_{t_d+i} = L\theta^{i-1} + \varepsilon_{t_d+i} \quad (9)$$

Figure 3 shows the output deviations from target with use of MMSE control for the disturbances shown in Fig. 2.

Basically, the disturbance of the process data stream is identified according to what data pattern is present. Hence, if we can extract and filter the noise signal that is mixed in process data, the disturbance level would be clearer. In this study, signal extraction is conducted using stICA. stICA is applied to two step-change disturbance data sets $\mathbf{Y}$, that has been converted using SPC and EPC, by first using SVD to obtain $\tilde{Y} = \tilde{U}\tilde{V}^T$, where columns of $\tilde{U}$ and $\tilde{V}$ contains the $k = 2$ eigensignals and eigensequences of $\mathbf{Y}$, respectively, with eigenvalues greater than unity. A conjugate gradient method then was used to find independent components. The initial condition for the un-mixing matrix $\mathbf{W}_{S(SD)}$ was randomly chosen. The entropies for spatial and temporal, entropy (sICA) and entropy (tICA) are defined following the notations in [43],

$$\text{Entropy (sICA)} = \log|W_s| + \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{k}\log\sigma_i'\left(\tilde{U}W_S\right)^{ij}$$

$$\text{Entropy (tICA)} = \log|W_p| + \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}\log\tau_i'\left(\tilde{V}W_p\right)^{ij}$$

where the functions σ_i and τ_i are assumed to be the cdfs of the individual spatial and temporal signals, respectively, and their derivatives σ_i' and τ_i' are their corresponding pdfs.

Here, we used the two step-change disturbance data sets to estimate matrix, $\mathbf{W}_{S(SD)}$, because this study finds that step-change disturbance process data could contain two different signal sources: disturbance data per se and noise signal. The results of temporal IC estimated through converted data are shown in Fig. 4a. Then, using the matrix $\mathbf{W}_{S(SD)}$, calculations of temporal ICs were conducted by pairing a set of new simulation data with a set of noise signals (at this stage, only a set of simulation was used because this study assumes monitoring is conducted only for one quality characteristic). Relevant results are provided in Fig. 4b. In Fig. 4, we may find that temporal ICs (IC_1 and IC_2) extracted through stICA could indeed successfully separate noise signals from mixture information.

After data conversion through EPC and stICA, noise signals can be successfully separated from process data. As every batch of converted data was assigned to a particular category according to the parameter value ($L = 0, 1, 5, 10, 15$) used in the original simulation process, we proceeded to build up process disturbance identification models for converted dataset. CART was utilized to construct these models. The rules created when a decision tree is established can facilitate improved process understanding by quality control staff and can serve as an important reference for subsequent process adjustments.

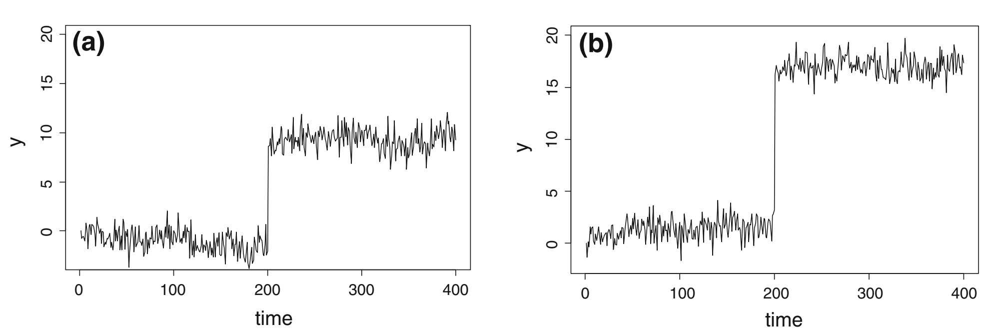

Fig. 2 **a** A step-change disturbance with $L = 10$ starts at $t = 201$; **b** A step-change disturbance with $L = 15$ starts at $t = 201$

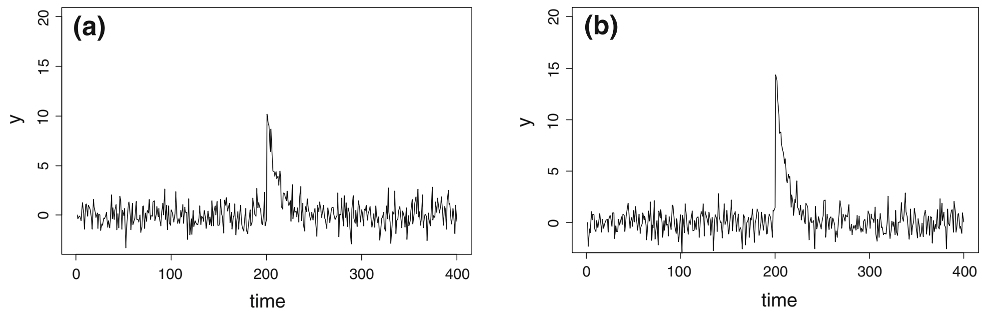

Fig. 3 Output deviations from target with use of MMSE control for a step-change disturbance with **a** $L = 10$ and **b** $L = 15$

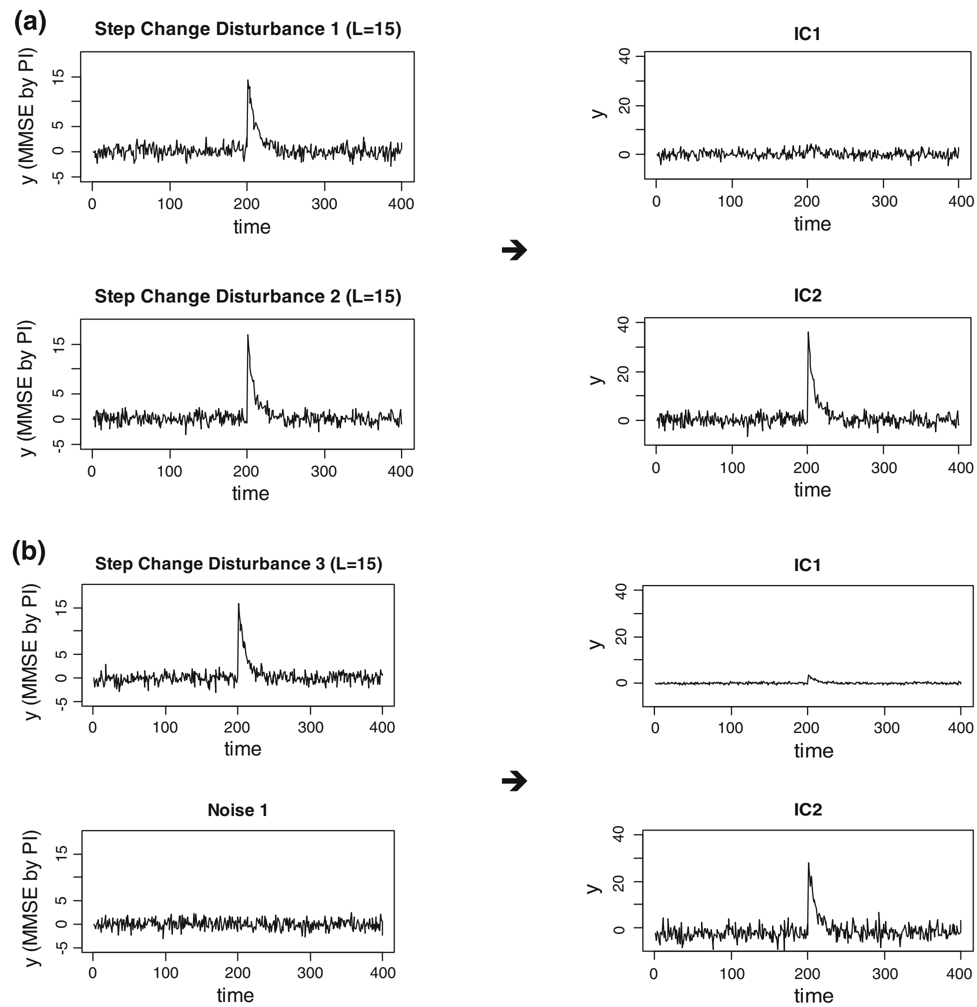

Fig. 4 Estimated temporal IC through converted data using $\tilde{\mathbf{X}} = \tilde{U} W_{S(SD)} \Lambda \left(\tilde{V} W_{P(SD)} \right)^T$. **a** Estimated temporal ICs through two sets of step-change disturbance with $L = 15$, **b** Calculations of temporal ICs by pairing a new set of step-change disturbance data with $L = 15$

In respect of choice of input variables, as there was only one non-noise-term IC separated through stICA on Stage I, this study screened for possible data according to the trend of the non-noise-term IC. Having viewed its pattern, six variables (X_1: average of first five periods; X_2: variance of first ten periods; X_3: slope of $t - 1$ period; X_4: slope of $t - 2$ period; X_5: single point paired with average of five periods; X_6: average of first three periods) were adopted to identify different process disturbance types and sizes. Moreover, to test the accuracy of the disturbance identification model constructed using CART, 2240 batches (70%) of the converted IC data were selected for training, and the remaining 960 batches (30%) served as test data.

The variables' importance and deduction rules (partial) generated using CART for SD dataset are summarized in Tables 1 and 2, respectively. According to Table 1, the six variables are listed by their importance as follows: X_5, X_1, X_6, X_2, X_3, and X_4. This suggests that single point paired with average of five periods could successfully identify the changing trend of SD data. Besides, X_2 and X_6, which represented change and shifting in SD data in a particular period respectively, helped identify the occurrence of a step-change disturbance. In Table 2, two step-change disturbance processes at $L = 10$ and 15 could be successfully identified through the particular upper and lower limits of the three variables X_1, X_3 and X_5. The production manager may look at the three particular values—average of first five periods $= -45.3934$, slope of $t-1$ period $= -0.679057$ and single point paired with average of five periods $= -5.72091$—to improve control of the manufacturing process.

Table 3 provides the identification success rates using PDIM with different correlation values (θ). According to Table 3, correlation values (θ) in data had slightly appreciable influence when the PDIM constructed with CART was used to identify disturbance levels. Furthermore, the success rate of identification for the disturbance identification model rose with increases in L value.

To verify the validity of the integrated monitoring method proposed here, this study applied the PCA-CART, ICA-CART and one commonly used SPC tool, CUSUM chart, as well to identify process disturbance levels. In applying PCA, we use the principle component with largest eigenvalue to represent the underlying process. And then, we applied the CART to construct the disturbance identification model. It should be noted that the customary T^2 chart is not used herein to monitoring of process data because we like to have the identification success rates with different disturbance levels and correlation values (θ). Relevant results are provided in Table 4. A comparison of data in Tables 3 and 4 finds that the identification results made by the proposed integrated approach (stICA-CART) are better than those generated using PCA with CART.

Table 1 Variables' importance (%) for the SD dataset ($\theta = 0.9$)

	X_1	X_2	X_3	X_4	X_5	X_6
Importance	92.14	57.72	54.78	32.36	100.00	76.63

In applying ICA, we adopt the FastICA algorithm proposed by Hyvärinen [19] to estimate de-mixing matrixes and ICs in process data. The testing results of the ICA-CART model with combinations of different parameter sets are summarized in Table 5. From Tables 3, 4, and 5, it can be found that the stICA-CART model gives the best identification result, and the identification results made by the ICA-CART model are slightly better than those generated by PCA-CART. They are consistent with the conclusion made by Lee [25]. The ICA solution extracts the original source signal to a much greater extent than the PCA solution if the latent variables follow non-Gaussian distribution. Moreover, the identification rate made by the ICA-CART method is influenced by the data correlation. But when $L = 15$, both the ICA-CART approach and the PCA-CART method have almost same identification success rates. This is caused by the fact that the step-change disturbance pattern is obvious when $L = 15$.

In applying CUSUM chart, the values of $k = 0.5$ and $h = 5$ were used, since they are effective across a broad

Table 3 Correct classification rate for testing data using stICA-CART (step-change disturbance)

θ	L				
	0 (%)	1 (%)	5 (%)	10 (%)	15 (%)
0.9	98.75	55.02	100.00	100.00	100.00
0.8	99.38	60.03	100.00	95.08	100.00
0.7	97.50	70.06	95.04	100.00	100.00
0.6	98.75	75.03	90.02	100.00	100.00
0.5	99.38	100.00	100.00	100.00	100.00
0.4	99.38	100.00	100.00	100.00	100.00
0.3	100.00	95.03	100.00	100.00	100.00
0.2	99.38	95.05	100.00	99.07	100.00
0.1	98.13	45.06	95.00	90.03	100.00
0.0	96.88	65.05	95.00	100.00	100.00
−0.1	99.38	90.00	100.00	85.05	100.00
−0.2	99.38	55.07	95.03	98.05	100.00
−0.3	96.25	65.02	100.00	100.00	100.00
−0.4	98.75	75.00	100.00	100.00	99.06
−0.5	98.75	90.01	100.00	100.00	100.00
−0.6	98.13	80.00	100.00	100.00	100.00
−0.7	98.13	80.04	100.00	98.09	100.00
−0.8	99.38	80.06	95.01	95.08	100.00
−0.9	99.38	80.00	100.00	100.00	100.00

Table 2 Deduction rule (partial) for SD dataset ($\theta = 0.9$)

Terminate node	Rule	Disturbance level
1	If ($X_5 \leq -5.72091$) and ($X_1 \leq -45.3934$) and ($X_3 \leq -0.679057$)	Class = 4 ($L = 15$)
2	If ($X_5 \leq -5.72091$) and ($X_1 \leq -45.3934$) and ($X_3 > -0.679057$)	Class = 3 ($L = 10$)

Table 4 Correct classification rate for testing data using PCA-CART (step-change disturbance)

θ	L				
	0 (%)	1 (%)	5 (%)	10 (%)	15 (%)
0.9	99.4	59.6	99.8	100.0	99.6
0.8	98.1	45.3	80.1	100.0	99.9
0.7	98.8	20.5	100.0	99.7	99.7
0.6	98.1	45.1	84.7	100.0	100.0
0.5	99.4	40.0	80.0	100.0	99.8
0.4	98.1	30.1	69.6	100.0	100.0
0.3	99.4	35.0	74.6	95.0	99.6
0.2	96.3	34.5	69.9	90.3	99.6
0.1	98.8	39.9	70.5	84.9	100.0
0.0	97.5	29.9	70.4	95.4	94.5
−0.1	97.5	40.3	85.5	95.0	100.0
−0.2	98.1	9.8	44.5	90.4	89.8
−0.3	96.3	45.1	95.1	85.2	99.8
−0.4	97.5	45.3	44.6	85.3	95.4
−0.5	95.6	15.1	70.0	85.1	95.3
−0.6	99.4	30.3	59.6	95.4	94.8
−0.7	96.3	44.8	90.3	95.3	100.0
−0.8	98.8	14.7	89.9	95.2	100.0
−0.9	91.3	45.1	95.2	99.7	99.5

Table 5 Correct classification rate for testing data using ICA-CART (step-change disturbance)

θ	L				
	0 (%)	1 (%)	5 (%)	10 (%)	15 (%)
0.9	98.77	55.02	95.03	100.00	100.00
0.8	100.00	55.09	55.04	100.00	100.00
0.7	98.81	65.08	80.05	100.00	100.00
0.6	98.13	65.03	75.06	90.08	100.00
0.5	96.98	45.03	95.06	100.00	100.00
0.4	99.43	40.01	70.08	95.09	95.05
0.3	98.83	60.06	70.02	100.00	100.00
0.2	97.56	55.04	60.00	100.00	100.00
0.1	99.47	65.08	70.08	100.00	95.07
0.0	100.00	60.04	75.03	95.06	100.00
−0.1	95.61	35.08	60.51	95.09	95.02
−0.2	100.00	60.03	95.09	100.00	100.00
−0.3	98.84	35.03	80.09	95.04	100.00
−0.4	100.00	40.09	65.09	90.02	100.00
−0.5	95.69	20.16	45.07	95.08	95.05
−0.6	96.34	30.08	100.00	100.00	100.00
−0.7	98.17	50.09	95.03	100.00	100.00
−0.8	100.00	25.05	100.00	100.00	100.00
−0.9	98.15	35.02	100.00	100.00	100.00

range of process shifts [36]. It is noted that the CUSUM chart can only be used for process monitoring but not for the identification of disturbance level. Table 6 shows the classification results for step-change disturbances at different levels by using the CUSUM control chart. Based on Table 6, three observations can be made. First, as the magnitude of the level of the step-change disturbance increases, there is also a corresponding increase of the correct classification percentages. However, the correct classification percentages are almost the same for the cases of $L = 10$ and 15. This may imply that there is no difference for the correct classification percentages once L is greater than a certain value. The second observation shows very low correct classification percentages for smaller levels of a disturbance (i.e., $L = 1$). This is due to the MMSE that provides virtually complete compensation for negligible disturbance. Thus, despite a disturbance of trend magnitude $L = 1$, the process acts as if there is no disturbance. When L is small, it is very difficult to detect the rise of the step-change disturbance. Consequently, it is very difficult to correctly identify if there is any disturbance occurrences. Finally, when L is moderate (i.e., $L = 5$), Table 6 indicates that the smaller (e.g., $\theta = -0.9$) is associated with the lower correct classification percentage rate. The reason may be due to the "accumulation" feature of the CUSUM control chart. The output deviation from the target, y_t, could be negative because of the negative θ, and

Table 6 Correct classification rate for testing data using CUSUM control chart (step-change disturbance)

θ	L				
	0 (%)	1 (%)	5 (%)	10 (%)	15 (%)
0.9	98.1	31.2	89.3	95.3	98.7
0.8	98.3	26.8	98.3	97.3	98.9
0.7	98.2	27.6	93.1	97.4	98.4
0.6	98.5	20.9	98.6	97.6	99.4
0.5	98.0	30.1	92.6	98.0	97.0
0.4	98.0	24.4	93.0	99.3	98.9
0.3	98.1	25.5	97.3	96.6	96.3
0.2	98.3	20.1	95.9	95.5	97.2
0.1	97.9	30.7	99.3	96.7	96.7
0.0	98.1	18.3	91.0	99.0	99.7
−0.1	98.3	17.6	91.4	98.8	97.9
−0.2	98.1	30.9	95.0	98.6	98.8
−0.3	98.3	21.2	93.9	98.6	99.9
−0.4	98.3	26.8	96.3	95.5	99.7
−0.5	98.4	20.4	94.3	97.8	99.7
−0.6	98.1	23.4	91.3	99.1	99.3
−0.7	98.2	27.9	98.7	99.3	96.6
−0.8	98.3	21.7	98.1	98.1	99.2
−0.9	98.3	26.9	84.3	93.3	98.5

this may cause the CUSUM statistics to be smaller. Consequently, this may delay the signal and result in a lower correct identification percentage rate.

4.2 Linear disturbance

A linear disturbance can be represented as

$$D_t = \begin{cases} S(t - t_s), & t \geq t_s \\ 0, & t < t_s \end{cases} \quad (10)$$

where S is the slope (or trend rate) of the linear disturbance. And it can be transformed as follows in terms of the B (a backward shift operator) defined in Sect. 4.1.

$$D_t = \frac{\beta_t}{(1 - B)^2} \qquad \beta_t = \begin{cases} S, & t = t_d \\ 0, & t \neq t_d \end{cases} \quad (11)$$

Then, a linear disturbance (LD) dataset which also comprises four subsets (LD1–LD4) can be generated from (5–6) and (11) by the same way as the data in SD dataset. An example of process monitoring subsets with $S = 1.0$ and 1.5 is shown in Fig. 5. About 70% of the dataset is used as the training sample, while the remaining 30% of the dataset is used as the testing sample. Then, by substituting (11) into (8), the linear disturbance process data converted by applying SPC and EPC can be obtained as (12).

$$y_{t_d+i} = S\left(\frac{\theta^i - 1}{\theta - 1}\right) + \varepsilon_{t_d+i} \quad (12)$$

Figure 6 shows the output deviations from target with use of MMSE control for our linear disturbances dataset.

Because this study also finds that linear disturbance process data could contain two different signal sources: disturbance data per se and noise signal, we used the two linear disturbance datasets to estimate $\mathbf{W}_{S(LD)}$ de-mixing matrix. Consequently, the ICs could be calculated by using $\mathbf{W}_{S(LD)}$ and a set of new simulation data with a set of noise signals. The relevant results are provided in Fig. 7. In the figure, we observe that the stICA can successfully separate noise signals from mixture information. After data conversion through EPC and stICA, we successfully excluded the noise signals from process data. And then, we used CART technique to construct the PDIM for the converted linear disturbance datasets.

The variables' importance for the six selected input variables and deduction rules (partial) are summarized in Tables 7 and 8, respectively. Similar to the results shown in Table 1, the linear disturbance could be identified in terms of average of first three periods (X_6). In addition, single point paired with average of five periods (X_5) in LD data could help identify the occurrence of a linear disturbance. However, the rank of the importance of X_1 variable is different from the one in SD data. The difference is mainly caused by the utilization of proposed stICA-based signal extraction technique. Because the extracted signals for step-change disturbance and the linear disturbance are different, the important variables are varied. Consequently, the deduction rules for both disturbances are different. In Table 8, we

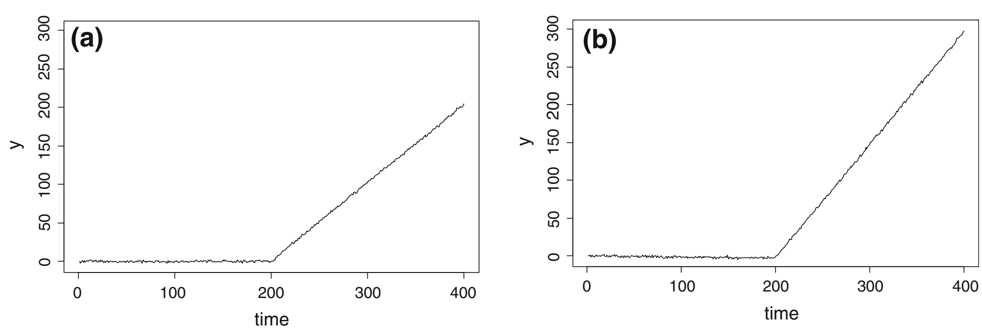

Fig. 5 **a** A linear disturbance with $S = 1.0$ starts at $t = 201$; **b** a linear disturbance with $S = 1.5$ starts at $t = 201$

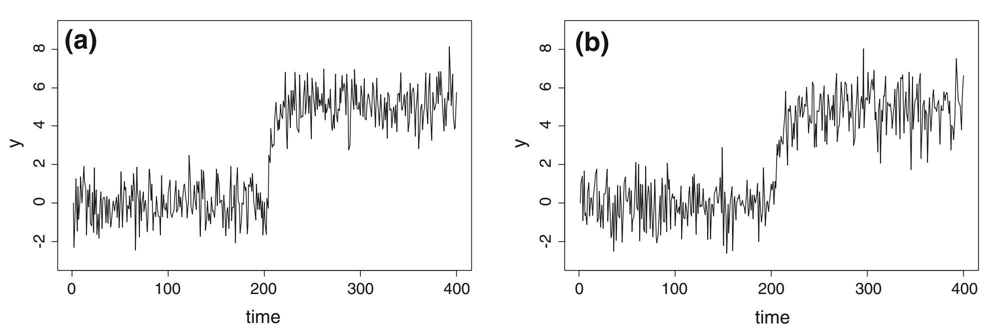

Fig. 6 Output deviations from target with use of MMSE control for a linear disturbance with **a** $S = 1.0$ and **b** $S = 1.5$

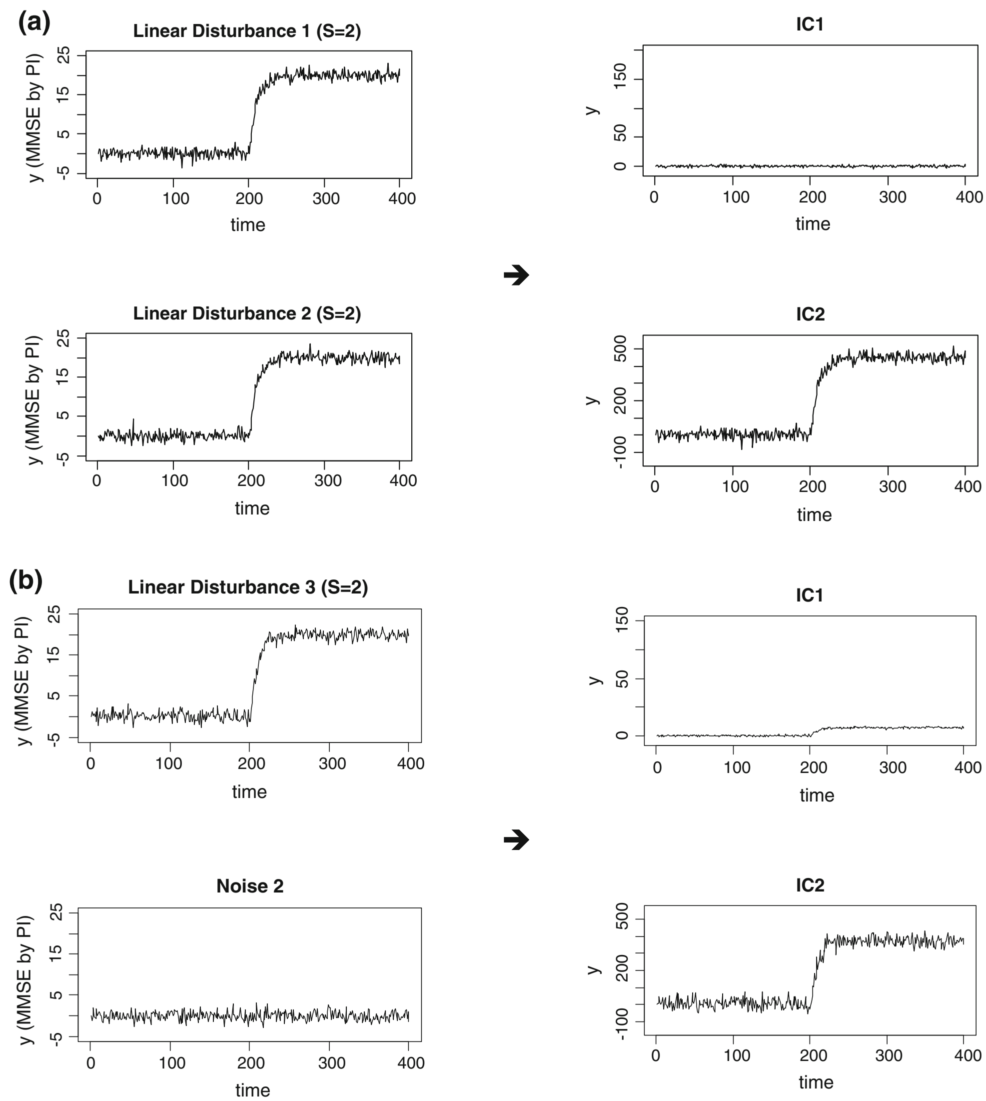

Fig. 7 Calculations of temporal ICs by pairing a set of new simulation data with a set of noise signals. **a** Estimated temporal ICs through two sets of linear disturbance with $S = 2$, **b** Calculations of temporal ICs by pairing a new set of linear disturbance data with $S = 2$

Table 7 Variables' importance (%) for the LD dataset ($\theta = 0.9$)

	X_1	X_2	X_3	X_4	X_5	X_6
Importance	33.29	74.63	33.82	32.85	76.38	100.00

found out that two variables, X_1 and X_5, could be used to identify the linear disturbance process at $S = 0.5$.

Table 9 provides the identification success rates using PDIM with linear disturbance datasets and different correlation values (θ). According to the Table, the identification model had a higher success rate with step-change disturbance data than with linear disturbance data. The cause, this study finds, was related to the converted IC data: step-change disturbance data, after stICA conversion, had higher levels of fluctuation than the linear disturbance data. As a result, when step-change disturbance occurred, the CART identification model could identify the disturbance levels accurately.

For comparison, the PCA-CART, ICA-CART and CUSUM chart approaches were applied to evaluate the identification capability of the proposed stICA-CART

Table 8 Deduction rule (partial) for the LD dataset ($\theta = 0.9$)

Terminate node	Rule	Disturbance level
1	If ($X_1 \leq -8.98377$) and ($X_5 \leq -6.86123$)	Class = 1 ($S = 0.5$)
2	If ($X_5 \leq -0.833547$) and ($-8.98377 < X_1 \leq -7.28314$) and ($1.9506 < X_6 \leq 3.91865$)	Class = 2 ($S = 1.0$)

Table 9 Correct classification rate for testing data using stICA-CART (linear disturbance)

θ	S				
	0 (%)	0.5 (%)	1.0 (%)	1.5 (%)	2.0 (%)
0.9	95.63	50.05	85.07	95.00	100.00
0.8	96.88	100.00	80.02	85.09	90.00
0.7	95.63	50.09	65.01	85.05	100.00
0.6	99.38	45.03	85.07	100.00	100.00
0.5	99.38	95.06	90.00	100.00	100.00
0.4	99.38	90.08	100.00	100.00	100.00
0.3	98.13	85.02	95.08	100.00	90.00
0.2	94.38	60.07	45.01	55.06	85.00
0.1	97.50	80.01	55.00	95.00	100.00
0.0	96.25	85.00	75.02	70.00	95.00
−0.1	99.38	80.02	100.00	95.06	100.00
−0.2	98.75	75.02	90.05	80.07	90.03
−0.3	96.25	75.00	100.00	100.00	100.00
−0.4	95.00	50.00	45.00	80.00	90.04
−0.5	98.75	75.05	95.02	100.00	100.00
−0.6	98.75	80.00	80.09	75.00	100.00
−0.7	97.50	60.09	85.00	95.00	100.00
−0.8	98.75	95.00	95.03	95.00	100.00
−0.9	98.75	85.00	95.00	100.00	100.00

Table 10 Correct classification rate for testing data using PCA-CART (linear disturbance)

θ	S				
	0 (%)	0.5 (%)	1.0 (%)	1.5 (%)	2.0 (%)
0.9	98.8	25.4	74.9	50.1	85.4
0.8	98.7	39.6	29.9	69.8	79.6
0.7	99.2	30.3	54.9	55.2	79.8
0.6	98.2	44.5	34.7	75.4	74.9
0.5	99.0	30.0	20.0	39.9	60.2
0.4	97.5	29.9	35.2	29.6	75.5
0.3	99.1	40.4	20.2	35.0	64.5
0.2	96.4	20.1	19.6	19.9	60.0
0.1	99.1	10.3	20.2	29.9	74.6
0.0	97.2	14.7	14.9	54.6	65.0
−0.1	97.4	10.5	24.9	25.4	59.9
−0.2	97.3	15.4	10.1	30.2	44.8
−0.3	95.9	30.5	39.8	15.2	40.3
−0.4	98.0	5.2	35.5	15.4	49.8
−0.5	95.0	15.5	15.1	14.6	65.1
−0.6	98.8	14.6	14.9	14.8	35.4
−0.7	95.3	19.5	15.4	30.2	34.5
−0.8	99.7	19.7	10.5	25.1	40.4
−0.9	92.2	9.5	5.3	10.3	35.1

Table 11 Correct classification rate for testing data using ICA-CART (linear disturbance)

θ	S				
	0 (%)	0.5 (%)	1.0 (%)	1.5 (%)	2.0 (%)
0.9	98.76	45.01	60.05	75.04	85.00
0.8	100.00	45.03	40.04	60.02	80.07
0.7	98.87	45.03	65.01	60.02	75.05
0.6	98.15	40.09	30.06	40.07	75.05
0.5	96.96	55.02	30.09	50.04	75.01
0.4	99.41	50.08	30.07	50.08	80.08
0.3	98.89	45.05	20.09	45.04	65.10
0.2	97.53	45.03	35.07	50.04	80.08
0.1	99.43	25.09	25.00	40.03	75.02
0.0	100.00	15.08	30.08	30.07	70.09
−0.1	95.61	30.03	45.04	35.08	80.01
−0.2	100.00	15.04	20.09	55.02	50.04
−0.3	98.85	35.07	25.03	35.06	50.02
−0.4	100.00	10.09	35.07	30.02	45.10
−0.5	95.01	25.04	10.03	20.04	75.07
−0.6	96.31	35.04	15.03	25.04	60.01
−0.7	98.18	25.02	30.05	25.02	50.00
−0.8	100.00	30.08	10.08	20.01	45.01
−0.9	98.14	25.07	35.02	20.03	55.06

Table 12 Correct classification rate for testing data using CUSUM control chart (linear disturbance)

θ	S				
	0 (%)	0.5 (%)	1.0 (%)	1.5 (%)	2.0 (%)
0.9	97.8	17.4	97.3	100.0	100.0
0.8	98.8	11.1	90.5	99.9	100.0
0.7	98.5	10.3	76.3	97.8	100.0
0.6	98.7	8.6	58.5	91.3	99.5
0.5	98.3	7.0	42.9	79.0	96.4
0.4	98.4	7.3	33.5	65.0	87.4
0.3	97.9	5.4	23.5	51.9	76.5
0.2	98.1	5.5	20.6	45.8	64.0
0.1	98.2	3.4	17.1	33.3	55.7
0.0	98.5	4.3	15.3	26.4	45.2
−0.1	98.7	3.7	9.9	20.9	39.2
−0.2	98.4	3.9	9.1	19.7	34.9
−0.3	98.2	2.8	7.8	16.3	31.1
−0.4	98.7	3.6	8.2	16.2	26.1
−0.5	98.1	3.8	9.4	13.1	22.6
−0.6	98.2	3.0	7.6	14.4	22.0
−0.7	98.3	2.6	7.0	13.8	22.9
−0.8	98.1	2.1	8.9	17.1	28.2
−0.9	98.3	2.8	11.1	19.7	32.5

model. Tables 10, 11 and 12 show, respectively, the identification results for the PCA-CART, ICA-CART and CUSUM models. The comparison results showed that the integration of stICA and CART not only takes advantage of the superior capability of stICA to identify process disturbances but also adds the benefit of developing decision trees that can provide critical insight into the process behavior and facilitate improved process control. Moreover, compared to PCA and ICA, stICA can extract more information from the autocorrelated process data, which makes stICA to be the most efficient method among these three approaches for handling data autocorrelation.

5 Conclusions

This study views autocorrelated manufacturing process data as a mixture signal that contains noise and process data per se and based on that view introduces the integrated use of stICA and CART techniques for process monitoring by proposing a framework that incorporates SPC/EPC, stICA and CART. Compared with conventional process monitoring methods, the new integrated method eliminates complicated matching of mathematic functions and can also guide fast and accurate corrections to the variables being monitored to enhance the efficiency of process monitoring.

This study first combined SPC/EPC with stICA for data conversion in an aim to use EPC to make timely and proper corrections to disturbances or errors in process. stICA was then applied to separate noise signals from process data. Lastly, a process disturbance identification model was constructed using CART and was applied to the process signals extracted using stICA. Decision tree analysis was adopted for constructing an identification model principally because the inference rules generated can be clearly interpreted. In addition, quality control personnel can use the rules to further improve the understanding and operation of the manufacturing process.

Acknowledgments This research was partially supported by the National Science Council of the Republic of China under Grant Number NSC 95-2221-E-027-072-MY3.

References

1. Adams BM, Tseng LT (1998) Robustness of forecast-based monitoring schemes. J Qual Technol 30:328–329
2. Albazzaz H, Wang XZ (2007) Introduction of dynamics to an approach for batch process monitoring using independent component analysis. Chem Eng Commun 194:218–233
3. Alwan LC, Roberts HV (1988) Time series modeling for statistical process control. J Bus Econ Stat 6:87–95
4. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth International Group, Belmont
5. Box GEP, Coleman DE, Baxley RV (1997) A comparison of statistical process control and engineering process control. J Qual Technol 29:128–130
6. Box GEP, Jenkins GM, Reinsel GC (1992) Time series analysis, forecasting, and control, 3rd edn. Prentice-Hall, Englewood Cliffs
7. Box GEP, Luceno A (1997) Statistical control by monitoring and feedback adjustment. Wiley, New York
8. Box GEP, Kramer T (1992) Statistical process monitoring and feedback adjustment—a discussion. Technometrics 34:251–285
9. Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001) Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum Brain Mapp 13:43–53
10. Peng C, Qian X, Ye D (2007) Electrogastrogram extraction using independent component analysis with references. Neural Comput Appl 16(6):581–587
11. Cichocki A, Amari SI (2002) Adaptive blind signal and image processing: learning algorithms and applications. Wiley, New York
12. Comon P (1994) Independent component analysis: a new concept? Signal Process 36:287–314
13. David V, Sanchez A (2002) Frontiers of research in BSS/ICA. Neruocomputing 49:7–23
14. Faltin FW, Hahn GJ, Tucker WT, Vander Wiei SA (1993) Algorithmic statistical process control: some practical observations. Int Stat Rev 61:67–80
15. Ge ZQ, Song ZH (2007) Process monitoring based on independent component analysis-principal component analysis (ICA-PCA) and similarity factors. Ind Eng Chem Res 46:2054–2063
16. Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. Neural Netw 13:411–430
17. Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 10:626–634
18. Hyvarinen A (1998) Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood. Neurocomputing 22:49–67
19. Hyvarinen A (1999) The fixed-point algorithm and maximum likelihood estimation for independent component analysis. Neural Process Lett 10:1–5
20. Hyvarinen A, Karhunen J, Oja E (2001) Independent component analysis. Wiley, New York
21. Hyvarinen A, Oja E (1997) A fast fixed-point algorithm for independent component analysis. Neural Comput 9:1483–1492
22. Kano M, Hasebe S, Hashimoto I (2003) Evolution of multivariate statistical process control: application of independent component analysis and external analysis. In: Proceedings of the foundations of computer aided process operations conference (FOCAPO 2003), Coral Spring, USA, pp 385–388
23. Kano M, Tanaka S, Hasebe S, Hashimoto I, Ohno H (2003) Monitoring independent components for fault detection. AIChE J 49:969–976
24. Lee TW (1998) Independent component analysis: theory and application. Kluwer, Boston
25. Lee JM, Yoo C, Lee IB (2004) Statistical process monitoring with independent component analysis. J Process Control 14(5): 467–485
26. Lee JM, Yoo C, Lee IB (2003) On-line batch process monitoring using different unfolding method and independent component analysis. J Chem Eng Jpn 36:1384–1396
27. Lee JM, Yoo CK, Lee IB (2003) New monitoring technique with an ICA algorithm in the wastewater treatment process. Water Sci Technol 47(12):49–56
28. Lee JM, Yoo CK, Lee IB (2004) Statistical monitoring of dynamic processes based on dynamic independent component analysis. Chem Eng Sci 59:2995–3006

29. Lee JM, Qin SJ, Lee IB (2006) Fault detection and diagnosis based on modified independent component analysis. AIChE J 52:3501–3514
30. Lu CW, Reynolds MR Jr (1999) EWMA control charts for monitoring the mean of autocorrelated processes. J Qual Technol 31:166–188
31. Luceno A (1995) Choosing the EWMA parameter in engineering process control. J Qual Technol 27:162–168
32. Macgregor JF (1990) A different view of the funnel experiment. J Qual Technol 22:255–259
33. Macgregor JF (1988) On-line statistical process control. Chem Eng Prog 84:21–31
34. Macgregor JF, Harris TJ, Wright JD (1984) Duality between the control of processes subject to randomly occurring deterministic disturbances and ARIMA stochastic disturbances. Technometrics 26:389–397
35. Montgomery DC, Mastrangelo CM (1991) Some statistical process control for autocorrelation data. J Qual Technol 23(3): 179–193
36. Montgomery DC, Keats JB, Runger GC, Messina WS (1994) Integrating statistical process control and engineering process control. J Qual Technol 26(2):79–87
37. Montgomery DC (2001) Introduction to statistical quality control, 4th edn. Wiley, New York
38. Shao YE (1998) Integrated application of the cumulative score control chart and engineering process control. Stat Sinica 8:239–252
39. Shao YE, Chiu CC (1999) Developing identification techniques with the integrated use of SPC/EPC and neural networks. Qual Reliab Eng Int 15:287–294
40. Shao YE, Runger GC, Haddock J, Wallace WA (1999) Adaptive controllers to integrate SPC and EPC. Commun Stat Simul Comput 28:13–36
41. Steinberg D, Colla P (1997) CART—classification and regression trees. Salford Systems, San Diego
42. Stone JV (2004) Independent component analysis: a tutorial introduction. MIT, Cambridge
43. Stone JV, Porrill J, Porter NR, Wilkinson ID (2002) Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. NeuroImage 15(2):407–421
44. He T, Clifford G, Tarassenko L (2006) Application of independent component analysis in removing artifacts from the electrocardiogram. Neural Comput Appl 15(2):105–116
45. Wieringa JE (1999) Statistical process control for serially correlated data. PhD dissertation. University of Groningen, Netherlands
46. Xia C, Howell J (2003) Isolating multiple sources of plant-wide oscillations via independent component analysis. Control Eng Pract 13(8):1027–1035
47. Xia C (2003) Control loop measurement based isolation of faults and disturbances in process plants. PhD thesis. University of Glasgow, UK
48. Yao ZX, Qian Y, Li XX, Jiang YB (2003) A description of chemical processes based on state space. In: Proceedings of the 8th international symposium on process systems engineering. Kunming, China, pp 1112–1117
49. Zhao CH, Wang FL, Mao ZY, Lu NY, Jia MX (2008) Adaptive monitoring based on independent component analysis for multiphase batch processes with limited modeling data. Ind Eng Chem Res 47:3104–3113