Shi-jin Ren, Yin Liang, Xiang-jun Zhao, Mao-yun Yang, 2015. A novel multimode process monitoring method integrating LDRSKM with Bayesian inference. *Frontiers of Information Technology & Electronic Engineering*, **16**(8):617-633. [doi:10.1631/FITEE.1400263]

A novel multimode process monitoring method integrating LDRSKM with Bayesian inference

Key words: Multimode process monitoring, Local discriminant regularized soft *k*-means clustering, Kernel support vector data description, Bayesian inference, Tennessee Eastman process

Contact: Shi-jin Ren
E-mail: sjren_phd@163.com
ORCID: http://orcid.org/0000-0002-8321-1879

Introduction

- The monitoring performance of traditional multivariate statistical process control (MSPC) methods may be degraded for multimode process.

- Clustering and mixture model approaches are typical of monitoring methods used for multimode process. They work well under the assumption that each operation mode follows a Gaussian distribution. However, in many applications, the process data do not exactly follow a Gaussian distribution.

- The separation of common information, specific mode information, and transition information can enhance nonlinear multimode monitoring performance and understanding of multiple mode behaviors.

- It is necessary to simultaneously partition the data into multiple subspaces and find the low-dimensional underlying subspace to fit with instances associated with one cluster.

- A local discriminant regularized soft *k*-means (LDRSKM) approach integrating with Bayesian inference is presented for multimode process monitoring.

A regularized soft *k*-means with locality preservation algorithm

The optimization problem of a regularized soft *k*-means with locality preserving (LPRSKM) proposed in this study is defined as follows:

$$
\begin{aligned}
\arg\min_{u_{ic}} & \sum_{c=1}^{C}\sum_{i=1}^{N} u_{ic} \| \boldsymbol{x}_i - \boldsymbol{m}_c^r \|^2 + \eta \sum_{c=1}^{C}\sum_{i=1}^{N} u_{ic} \log u_{ic} \\
& + \gamma \sum_{i,j}^{N}\sum_{c=1}^{C} \left(u_{ic} \log \frac{u_{ic}}{u_{jc}} + u_{jc} \log \frac{u_{jc}}{u_{ic}} \right) \cdot S_{ij} \\
\text{s.t.} & \sum_{c=1}^{C} u_{ic} = 1 \ \text{ and } \ u_{ic} \ge 0, \quad i = 1, 2, \cdots, N
\end{aligned}
$$

The approach exploits the intrinsic geometry of the probability distribution and locality discriminative information, and the cluster assignments are found via membership degrees.

The centroids, covariance matrices, and membership are iteratively updated as follows:

$$\boldsymbol{m}_c^r = \sum_{i=1}^{N} u_{ic}\boldsymbol{x}_i \Big/ \sum_{i=1}^{N} u_{ic}$$

$$\boldsymbol{\Sigma}_c = \sum_{i=1}^{N} u_{ic}(\boldsymbol{x}_i - \boldsymbol{m}_c^r)(\boldsymbol{x}_i - \boldsymbol{m}_c^r)^{\mathrm{T}} \Big/ \sum_{i=1}^{N} u_{ic}$$

$$u_{ic} = \frac{\exp\left(-\gamma \sum_{j}^{N} l_{ij}^c S_{ij} \Big/ \eta\right)\exp(-\|\boldsymbol{x}_i - \boldsymbol{m}_c^r\|^2/\eta)}{\sum_{c=1}^{C}\exp\left(-\gamma \sum_{j}^{N} l_{ij}^c S_{ij} \Big/ \eta\right)\exp(-\|\boldsymbol{x}_i - \boldsymbol{m}_c^r\|^2/\eta)}$$

where $S_{ij} = \begin{cases} \exp(-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2/t^0), & \\ \qquad \boldsymbol{x}_i \in \mathrm{NN}(\boldsymbol{x}_j) \text{ or } \boldsymbol{x}_j \in \mathrm{NN}(\boldsymbol{x}_i), & \\ 0, \quad \text{otherwise}, & \end{cases}$

and the proof of the convergence of LPRSKM is similar to that described by Miyamoto and Mukaidono (1997) and Jing *et al.* (2007).

Generalized linear discriminant analysis

Just like traditional LDA, GELDA obtains low-dimensional subspaces using the membership degrees from soft clustering in a supervised situation. The between-cluster matrix $\boldsymbol{S}_\mathrm{b}$, within-cluster matrix $\boldsymbol{S}_\mathrm{w}$, and total scatter matrix $\boldsymbol{S}_\mathrm{t}$ are defined by

$$\boldsymbol{S}_\mathrm{b} = \frac{1}{N}\sum_{c=1}^{C} n_c(\boldsymbol{m}_c - \boldsymbol{m})(\boldsymbol{m}_c - \boldsymbol{m})^\mathrm{T}$$

$$\boldsymbol{S}_\mathrm{w} = \sum_{c=1}^{C}\sum_{i=1}^{N} u_{ic}(\boldsymbol{x}_i - \boldsymbol{m}_c)(\boldsymbol{x}_i - \boldsymbol{m}_c)^\mathrm{T}$$

$$\boldsymbol{S}_\mathrm{t} = \sum_{i=1}^{N}(\boldsymbol{x}_i - \boldsymbol{m})(\boldsymbol{x}_i - \boldsymbol{m})^\mathrm{T}$$

An optimal projection matrix is achieved by maximizing the following objective function:

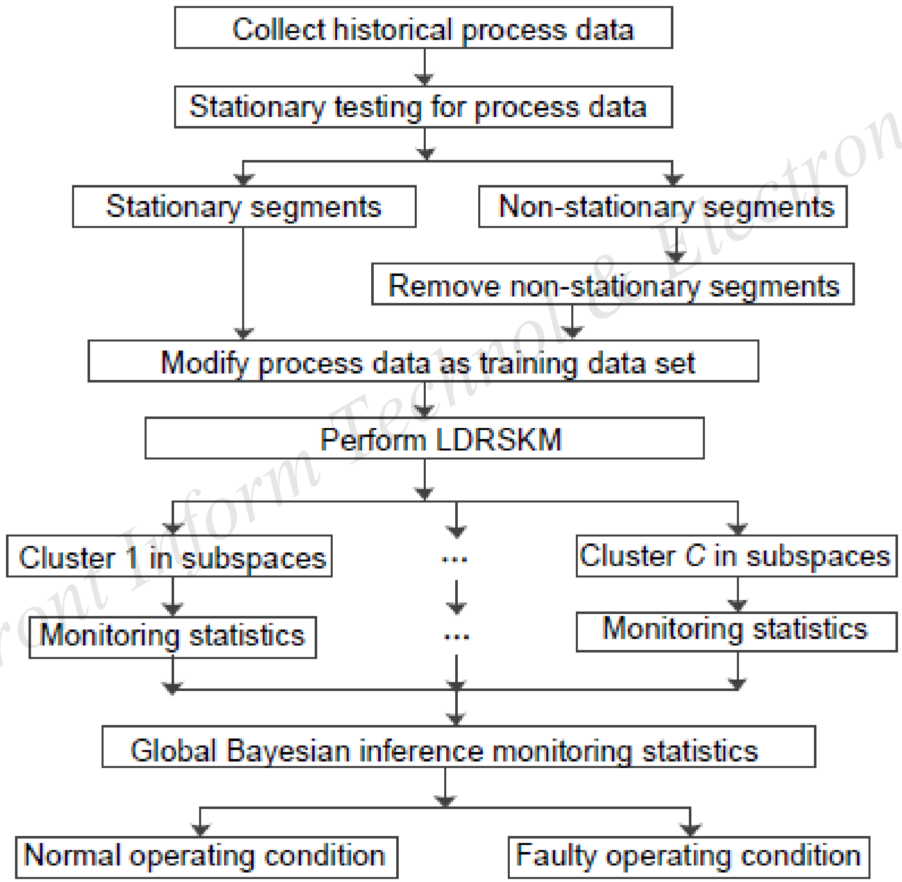
$$\boldsymbol{P}^{*} = \arg\max_{\boldsymbol{P}} \frac{\operatorname{tr}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{b}}\boldsymbol{P})}{\operatorname{tr}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{w}}\boldsymbol{P})},$$

where tr(·) represents the trace operator of a matrix.

The construction of monitoring statistics

- First, multiple kernel SVDDs (KSVDD) are used for the construction of monitoring statistics and control limits for all process modes.
- Then Bayesian inference based monitoring statistics are established by combining monitoring results of all the principal and residual subspaces in a probabilistic manner.

Flow chart of the proposed monitoring method



Tennessee Eastman benchmark simulations

Three operation modes were chosen in this study, 16 variables listed in Table 1 were selected for monitoring purposes, as in Kano *et al.* (2002). The process data were collected by running the simulation for 60 h under normal conditions. Each normal mode lasts 20 h.

Simulation results

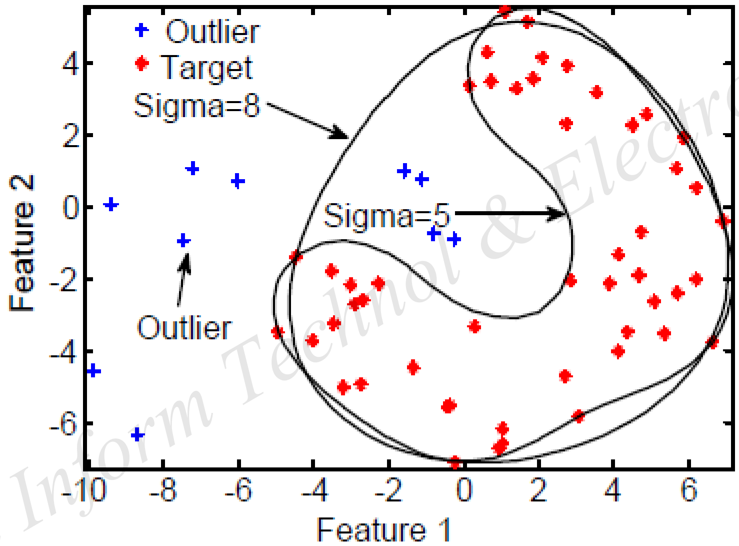


Fig. 3 KSVDD with different kernel parameter (Sigma) values on the Banana set (target class 1)

Simulation results (Con’d)

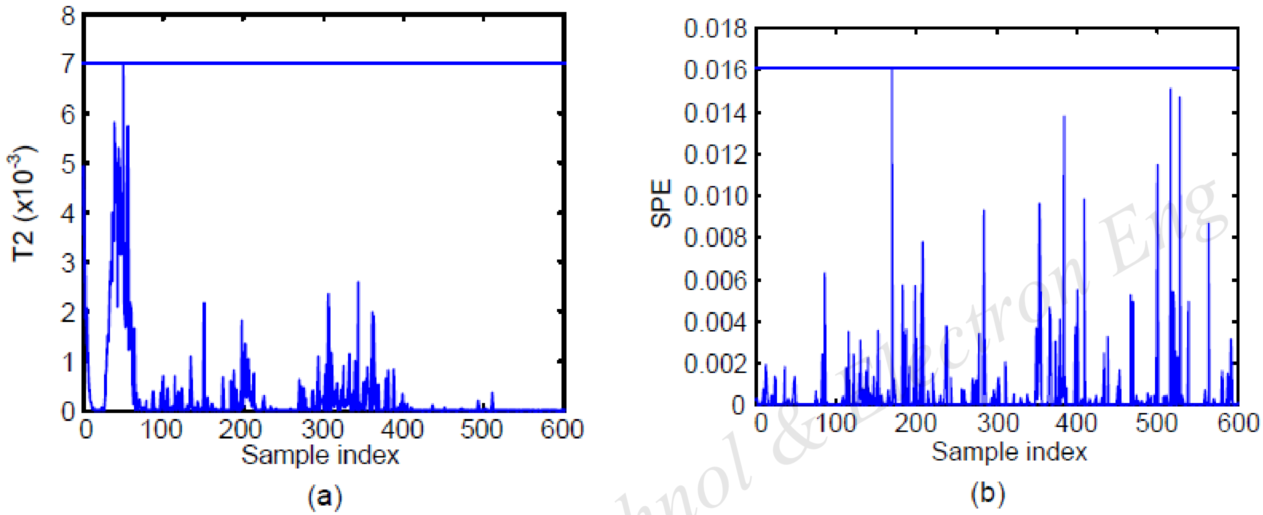


Fig. 4 LPP-FCM monitoring results for the three normal operation modes: (a) T2; (b) SPE

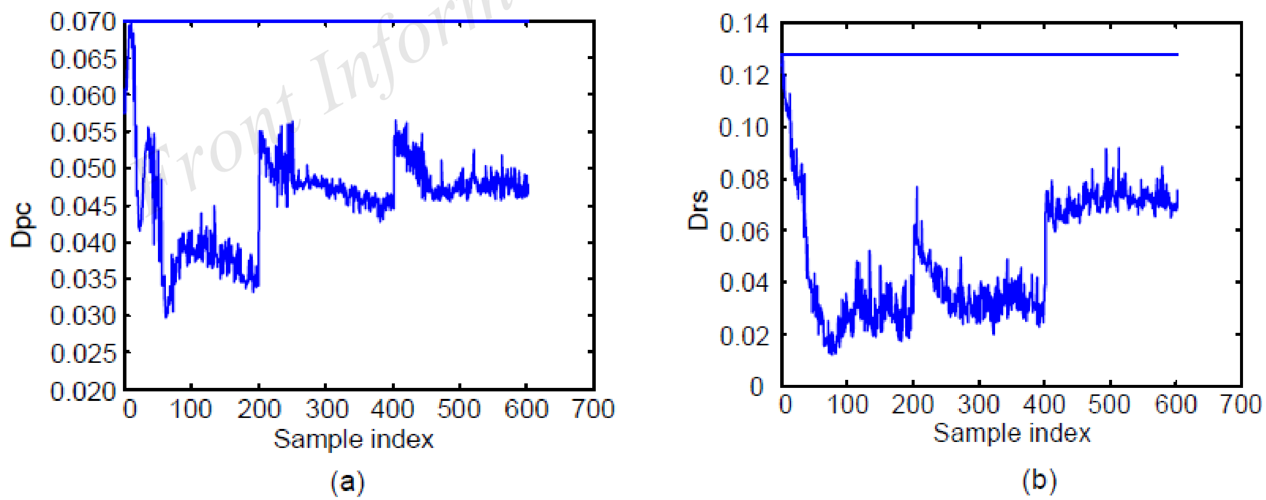


Fig. 5 LDRSKM monitoring results for the three normal operation modes: (a) Dpc; (b) Drs

Simulation results (Con'd)

Table 3 Four test cases with different kinds of process faults

Case	Description
1	Fault 2: 100 faulty samples under mode 1
	Fault 3: 290 faulty samples under each mode
2	Fault 4: 100 faulty samples under mode 1
	Fault 5: 290 faulty samples under each mode
3	Fault 7: 100 faulty samples under mode 1
	Fault 8: 290 faulty samples under each mode
4	Fault 10: 100 faulty samples under mode 1
	Fault 9: 290 faulty samples under each mode

Simulation results (Con’d)

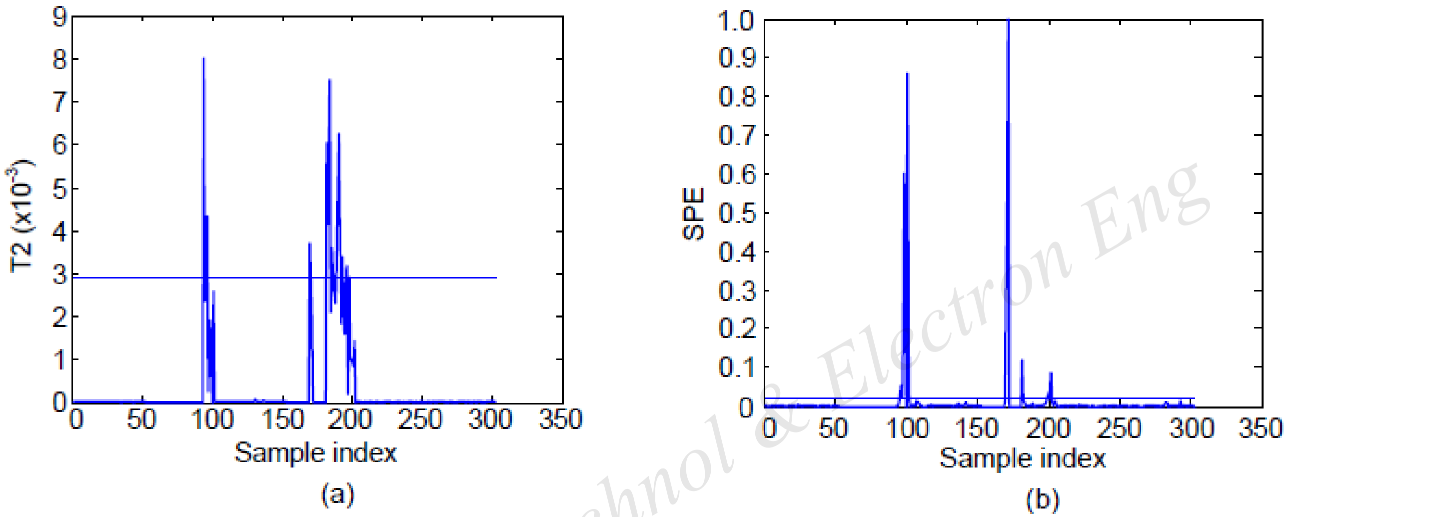


Fig. 6 LPP-FCM monitoring results for transition process data in principal and residual subspaces: (a) T2; (b) SPE

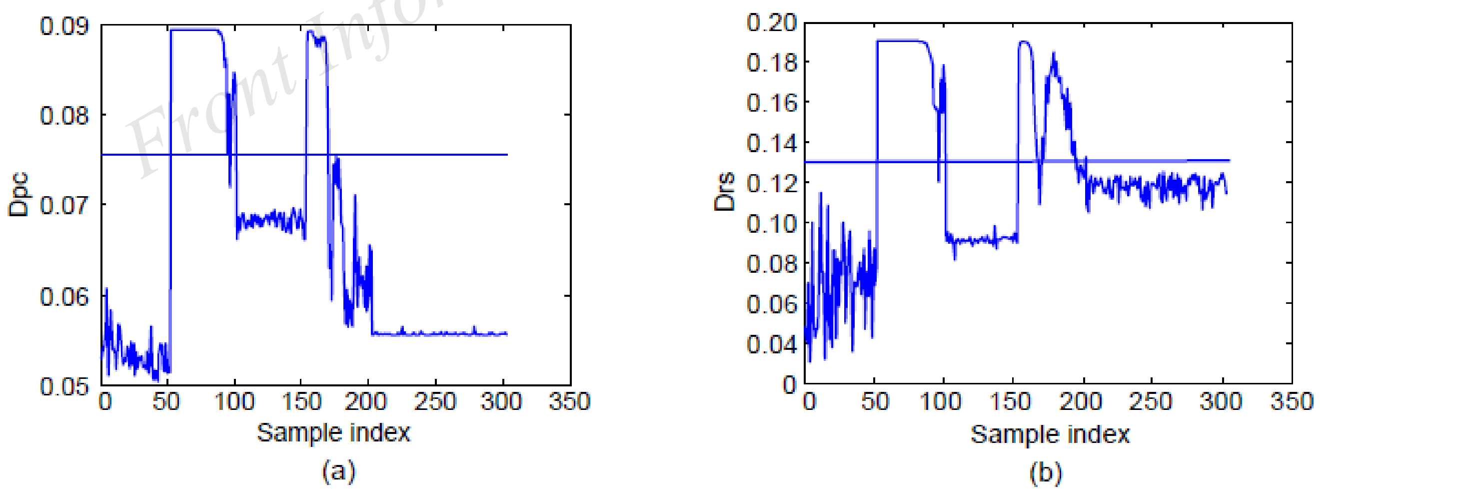


Fig. 7 LDRSKM monitoring results for transition process data in principal and residual subspaces: (a) Dpc; (b) Drs

Simulation results (Con’d)

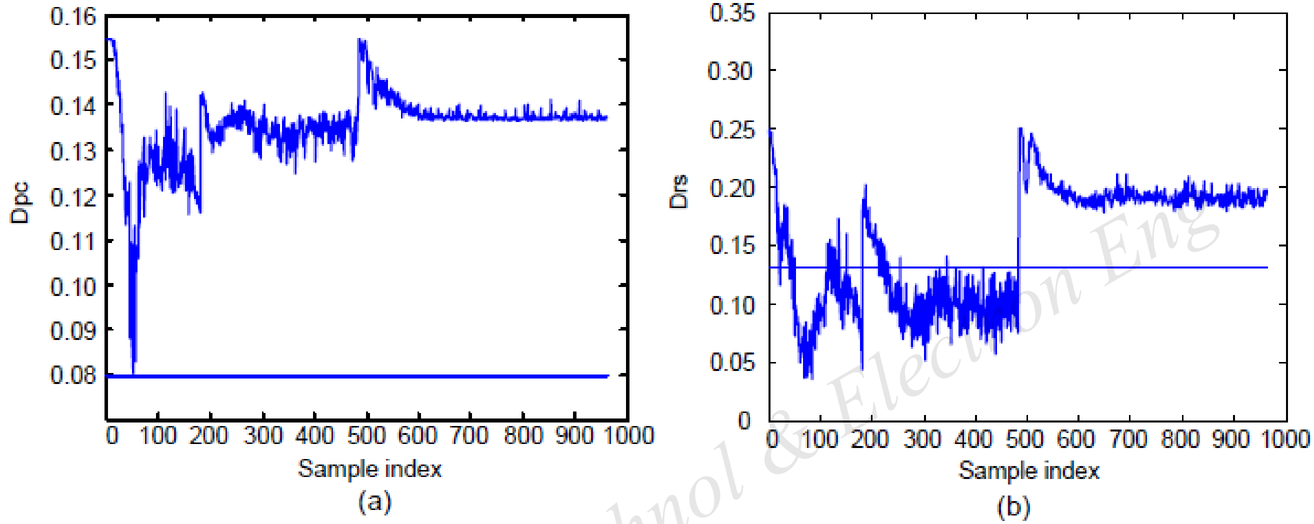


Fig. 8 LDRSKM monitoring results for fault 2 and fault 3: (a) Dpc; (b) Drs

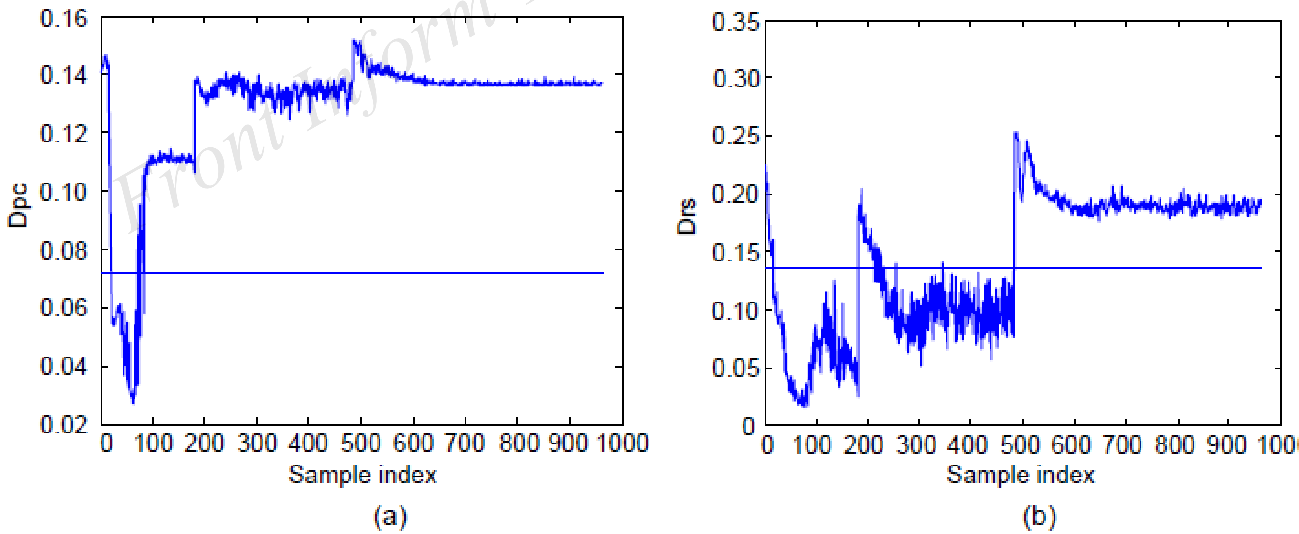


Fig. 9 LDRSKM monitoring results for fault 4 and fault 5: (a) Dpc; (b) Drs

Simulation results (Con'd)

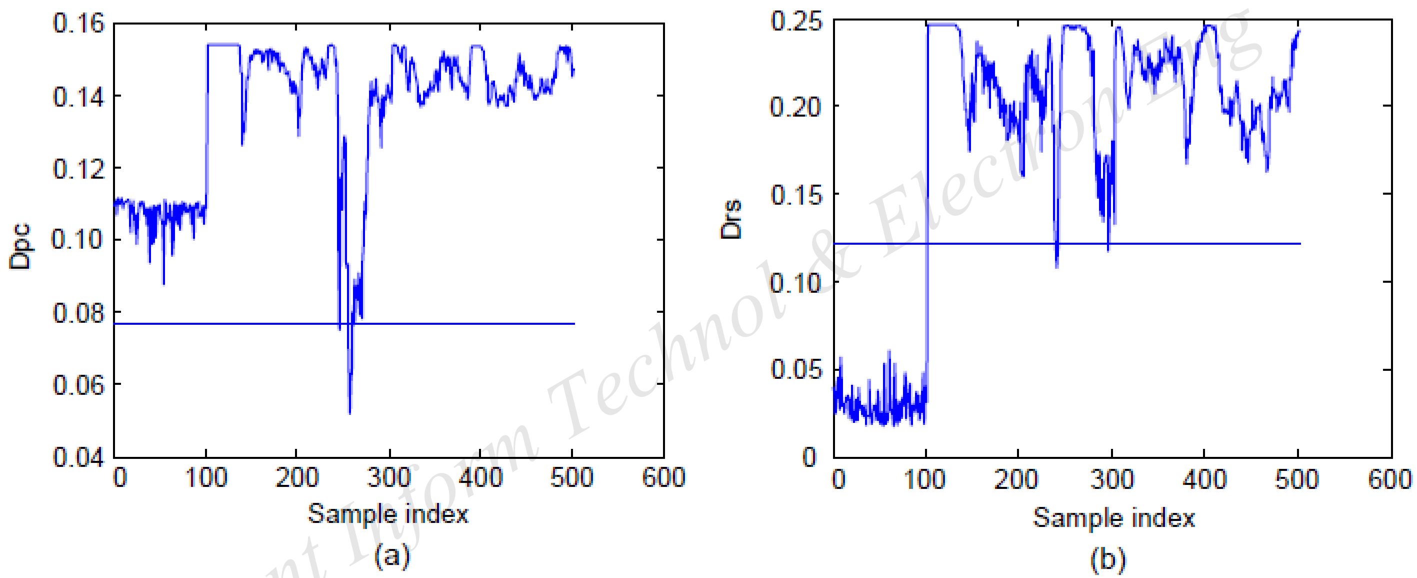


Fig. 10 LDRSKM monitoring results for fault 7 and fault 8: (a) Dpc; (b) Drs

Conclusions

- The proposed approach can partition the overlapped multimodal operating data into optimal discriminant low-dimensional subspaces by iteratively performing soft clustering and discriminant dimensionality reduction, enhancing the separability of data residing in different subspaces.
- Compared to the state-of-the-art multimode process monitoring approaches, our method can detect faults effectively and reliably and show satisfactory monitoring performance.