Krist V. Gernaey, Jakob K. Huusom and Rafiqul Gani (Eds.), 12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering.
31 May – 4 June 2015, Copenhagen, Denmark 

# Process Monitoring and Fault Detection in Non-Linear Chemical Process Based On Multi-Scale Kernel Fisher Discriminant Analysis

Norazwan Md Nor, [a] Mohd Azlan Hussain, [a] Che Rosmani Che Hassan[a]

[a]*Department of Chemical Engineering, University of Malaya, 50630 Kuala Lumpur, Malaysia*

## Abstract

This paper presents a multi-scale kernel Fisher discriminant analysis (MSKFDA) algorithm combining Fisher discriminant analysis (FDA) and its nonlinear kernel variation with the wavelet analysis. This approach is proposed for investigating the potential integration of wavelets and multi-scale methods with discriminant analysis in nonlinear chemical process monitoring and fault detection system. In this paper, a discrete wavelet transform (DWT) is applied to extract the dynamics of the process at different scales. The wavelet coefficients obtained during the analysis are used as input for the algorithm. By decomposing the process data into multiple scales, MSKFDA analyse the dynamical data at different scales and then restructure scales that contained important information by inverse discrete wavelet transform (IDWT). A monitoring statistic based on Hoteling's $T^2$ statistics is used in process monitoring and fault detection. The Tennessee Eastman benchmark process is used to demonstrate the performance of the proposed approach in comparison with conventional statistical monitoring and fault detection methods. A comparison in terms of false alarm rate, missed alarm rate and detection delay, indicate that the proposed approach outperform the others and enhanced the capabilities of this approach for the diagnosis of industrial applications.

**Keywords**: multi-scale, Fisher discriminant analysis, wavelet analysis, process monitoring, fault detection

## 1. Introduction

Various modelling methodologies have been developed for achieving efficient monitoring and control system for chemical processes. These methodologies are broadly divided into three types: quantitative model-based methods; qualitative model-based methods; and process history based methods (Venkatasubramanian *et al*., 2003). The process history based or data-driven method concerns with the transformation of large amounts of data into a particular form of knowledge representation that will enable proper detection and diagnosis of faults. Availability of vast amounts of plant data has encouraged researchers to develop and improve the data-driven-based and multivariable statistical process monitoring based methods, which use statistic projection technique to extract key process information from theses massive process data. These methods including such as Principal Component Analysis (PCA), Partial Least Square (PLS), Independent Component Analysis (ICA) and Fisher Discriminant Analysis (FDA) have been widely applied to on-line continuous and batch process monitoring, fault detection and diagnosis.

In some context, fault diagnosis problems can be considered as classification problems when lots of historical data are obtained from various faulty conditions, and then feature extraction and pattern recognition or classification methods can be used for fault diagnosis. Linear supervised classification methods such as PCA, discriminant PLS (DPLS) and FDA can be used. PCA represents high-dimensional process data in a reduced dimension, which brings convenience for process monitoring. However, PCA aims at reconstruction but not classification, which degrades the performance in classification problems while FDA provides an optimal lower dimensional representation in terms of discriminating among classes, which is very useful for fault diagnosis. Though it has been proven that FDA outperforms PCA in the classification problems, it is still linear in nature, which degrades the performance of FDA in monitoring the nonlinear system.

Furthermore, a chemical process is often characterized by large scale and non-linear behaviour. When linear FDA is used for fault diagnosis in non-linear system, a lot of incorrect diagnosis results will occur. As solution to deal with the nonlinear system, and to improve the classification ability, kernel-based FDA, called kernel FDA (KFDA), is introduced. The basic idea of the KFDA is to map the input sample data into a kernel feature space by a nonlinear kernel function and then perform linear FDA in the nonlinearly mapped feature space to find the discriminant feature vectors for classification. KFDA has turned out to be effective in many real-world applications due to its power of extracting the most discriminatory nonlinear. However, application of KFDA in fault diagnosis of chemical and biological processes is still limited.

There were also limited application on multi-scale FDA and KFDA that integrates FDA and wavelet packet analysis (Vana *et al.*, 2011). Discrete wavelet analysis decomposes the high-frequency part further, which wavelet analysis not does, and adaptively selects relative frequency bond based on character of signal to be analysed. To further improve de-noising character of multi-scale KFDA, the paper describes a discrete wavelet transform KFDA, which combines the ability of KFDA to de-correlate the variables by discriminating a nonlinear relationship with that of discrete wavelet transform to extract auto-correlated measurements. Then, a novel multi-scale kernel FDA is proposed by combining FDA and its nonlinear kernel variation with the wavelet analysis. Finally, an individuals control charts (XmR charts) and Hoteling's $T^2$ statistics are used to monitor the fault data in process monitoring and fault detection. The proposed method is evaluated and compared with the $T^2$ statistical methods based on the PCA in terms of the average run length (ARL). The paper is organized as follows. In Section 2 the background of FDA, KFDA and the discrete wavelet transform (DWT) is introduced. Section 3 introduces the proposed fault diagnosis approach with kernel FDA and integrates it with the wavelet transform. The case study illustrates an application to Tennessee Eastman process is provided in Section 4, and Section 5 concludes the paper.

## 2. Background

### *2.1. Fisher Discriminant Analysis*

Fisher discriminant analysis is a linear dimensionality reduction technique to find a direction for which data classes are optimally separated. The optimal discriminant direction is determined by maximizing the scatter within the classes (Wang and Romagnoli, 2005). Let the training data for all faulty classes be stacked into a $n$ by $m$ matrix $X \in R^{n \times m}$, where $n$ is the observation number and $m$ is the variable number. The within-class-scatter matrices $S_W$ and the between-class-scatter matrix $S_B$ contain all the basic information about the relationship within the groups and between them as in Eq.(1) and Eq.(2) respectively.

$$S_W = \sum_{k=1}^{K} \sum_{n \epsilon C_k} (x_n - m_k)(x_n - m_k)^T \tag{1}$$

$$S_B = \sum_{k=1}^{K} N_k (x_n - m_k)(m_k - m)^T \tag{2}$$

where $N_k$ and $m_k$ are the number and mean vector of the points assigned to class k, respectively; $m$ is the total mean of all the samples, and $K$ is the number of classes and $x_n$ is the data sample. Maximizing Eq.(3) with $w$ is the solution is equivalent to maximizing the between-class scatter $S_B$,and minimizing the within-class scatter $S_W$.

$$J(w) = \arg\max \frac{|w^T S_B w|}{|w^T S_W w|}. \tag{3}$$

*2.2. Kernel Fisher Discriminant Analysis*

The idea of KFDA is to solve the problem of FDA in the feature space $F$. However, since any solution $w \in F$ must lie in the span of all the samples in F, there exists coefficients $\alpha = \{ \alpha ,i = 1, 2, \ldots, n\}$ and the mapping of sample class, $\phi_i$ such that

$$w = \sum_{i=1}^{n} \alpha_i \phi_i. \tag{4}$$

$K_B$ and $K_W$ are the between-class kernel matrix and within-class kernel matrix, as in Eq.(5) and Eq.(6) respectively.

$$w^T S_B w = \alpha^T K_B \alpha \tag{5}$$

$$w^T S_W w = \alpha^T K_W \alpha \tag{6}$$

So the solution can be achieved by maximizing Eq.(7) following Fisher criterion

$$J(\alpha) = \arg\max \frac{|\alpha^T K_B \alpha|}{|\alpha^T K_W \alpha|}. \tag{7}$$

*2.3. Discrete Wavelet Transform*

Wavelets are basis functions that are localized in both time and frequency. Generally, the dyadically discretized form of wavelets is used and it can be represented as

$$\psi_{mk}(t) = 2^{-m/2} \psi(2^{-m} t - k), \tag{8}$$

where, $\psi(t)$ is the mother wavelet, and $m$ and $k$ are dilation and translation parameters, respectively. The translation parameter determines the location of the wavelet in the time domain, while the dilation parameter determines its scale and location in the frequency domain. By projecting a signal on the wavelet basis function, its contributions in different regions of the time-frequency space can be obtained. The scaling and wavelet coefficients represented in terms of original measured data vector, $x$, take the form of

$$a_s = H_s x,\ d_s = G_s x \tag{8}$$

where, $H_s$ represents projection $m$ times on the scaling function, and $G_s$ represents projection $(s-1)$ times on the scaling function and once on the wavelet. The sequences $H$ and $G$ are low-pass and high-pass filters derived from the corresponding basis function, respectively. The scaling coefficients represent the lower frequency approximation of the signal while the wavelet coefficients represent the higher frequency components of the signal (Maulud *et al.*, 2005).

## 3. Methodology

The detailed in multi-scale kernel FDA application procedure of fault diagnosis were discussed briefly in this section. First, each of the $m$ variables is first decomposed individually by applying discrete wavelet transformation (DWT). Then, the Kernel FDA is performed on the wavelet coefficients for each selected scale. Appropriate numbers of component loading vectors are retained and the wavelet coefficients are reconstructed at each selected scale. In this work four scales (s=4) are used for discrete wavelet transformation (DWT) of the original signal. After that, the wavelet coefficients larger than a selected threshold corresponding to a significant event are retained. The variables consisting of deterministic components are reconstructed from the retained wavelet coefficients through inverse discrete wavelet transformation (IDWT) and the loadings of the extracted deterministic components are computed.

The new observations are projected into lower dimensional subspace. This subspaces measures the systematic or state variations occurring in the process. Meanwhile, the lower dimensional subspace corresponding to the $(m-a)$ smaller singular value describes the random variations of the process such as that associated with measurement noise are term as residual space. KFDA is used to search the optimal one-dimensional discriminant direction between the fault data and the normal data. Thus, individuals control charts, also known as XmR charts, are used to monitor the fault data on the optimal discriminant direction with contribution plot based on the optimal discriminant direction from KFDA is also used to improve its performance (Pei *et al.*, 2006). The proposed method is applied to the Tennessee Eastman process. Finally, the proposed method is evaluated and compared with the Hotelling's $T^2$ statistical methods.

## 4. Case study

Tennessee Eastman process (TEP) as described by Downs and Vogel (1993) in Figure 1was used as a case study. The process includes a total of 52 variables with 20 different faults were simulated (Yélamos *et al.*, 2006). The data set for the process and the details can be obtained from the Multi-scale Systems Research Laboratory (MIT, 2013).

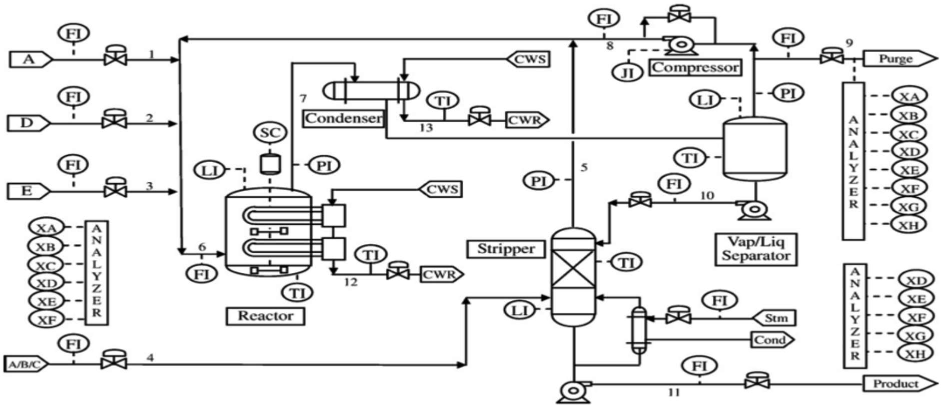


Figure 1 Tennessee Eastman process diagram

Table 1 Selected faults for four combination cases

| | *ID* | *Fault description* | *Type* |
|---|---|---|---|
| Case 1 | Fault 3 | D feed temperature | Step change |
| | Fault 4 | Reactor cooling water inlet temperature | Step change |
| | Fault 11 | Reactor cooling water inlet temperature | Random variation |
| Case 2 | Fault 8 | A, B, C feed composition | Random variation |
| | Fault 12 | Condenser cooling water inlet temperature | Random variation |
| | Fault 15 | Condenser cooling water valve | Sticking |
| Case 3 | Fault 2 | B composition, A/C ration constant | Step change |
| | Fault 9 | D feed temperature | Random variation |
| | Fault 13 | Reactor kinetics | Slow drift |
| Case 4 | Fault 9 | D feed temperature | Random variation |
| | Fault 14 | Reactor cooling water valve | Sticking |
| | Fault 17 | Unknown | Unknown |

These case studies and their respective faults were tabulated in Table 1, covering all fault types in the TEP simulation. Each case study combination includes three different types of faults and each of fault type sampling intervals is set to be 3 min with every sample data contains 52 process variables.

## 5. Results and Discussion

Table 2 Comparison among diagnosis success rates using different approaches

| | | *Diagnosis success rate (%)* | | |
|---|---|---|---|---|
| | | FDA | KFDA | MSKFDA |
| Case 1 | Training data set | 67.5 | 100 | 98.0 |
| | Testing data set | 41.7 | 84.7 | 92.7 |
| Case 2 | Training data set | 89.1 | 100 | 100 |
| | Testing data set | 33.7 | 72.7 | 91.3 |
| Case 3 | Training data set | 100 | 100 | 100 |
| | Testing data set | 86.7 | 100 | 100 |
| Case 4 | Training data set | 93.7 | 100 | 100 |
| | Testing data set | 62.7 | 95.3 | 98.3 |

A summary of the classification results for MSKFDA, KFDA and FDA is listed in Table 2. Compared with the classification performance of KFDA, MSKFDA has a significant improvement. After fault data are decomposed by DWT wavelet analysis, KFDA is performed on these multi-scaled fault data, which offers important supplemental classification information to KFDA. Without proper variable weighting, all variables are used in a same level and the data sets are masked with irrelevant information. The integration of DWT with KFDA improved the extraction of features that are relevant to the abnormal operation in both time and frequency domain and lead to better classification. In addition, the high misclassifications rate for FDA shows the advantage of nonlinear technique when the fault data are highly overlapped.

## 6. Conclusions

In this paper, MSKFDA-based fault diagnosis for the Tennessee Eastman process is presented. The data discrimination and fault detection based on MSKFDA methodology enhanced the diagnosis proficiency by taking into consideration the multi-scale information compared to other methods that considered only single scale nature. Moreover, it can provide a better separation of the deterministic and stochastic features and improve the extraction of features that are relevant to a faulty situation from both time and frequency domain aspects. The application of the proposed MS-KFDA shows better fault diagnosis performance than KFDA and FDA. The high misclassification rate for FDA also shows the advantage of nonlinear technique when the fault data are highly overlapped.

## References

Downs, J.J., Vogel, E.F., 1993. A plant-wide industrial process control problem. Computers and Chemical Engineering, 17, 245–255

Massachusetts Institute of Technology, 2013, The Simulation Code of Tennessee Eastman, Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, U.S.A.

Maulud, A. H. S., Wang, D., & Romagnoli, J. A., 2005, Wavelet-Based Nonlinear Multivariate Statistical Process Control. In L. Puigjaner & A. Espuña (Eds.), European Symposium on Computer Arded Process Engineering – 15.

Pei, X., Yamashita, Y., Yoshida, M., Matsumoto, S., & Pei, X., 2006, Discriminant analysis and control chart for the fault detection and identification. In W. Marquardt & C. Pantelides (Eds.), 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering, pp. 1281–1286.

Vana, Z., Privara, S., Cigler, J., & Preisig, H. A., 2011, System identification using wavelet analysis. In E. N. Pistikopoulos, M. C. Georgiadis, & A. C. Kokossis (Eds.), 21st European Symposium on Computer Aided Process Engineering – ESCAPE 21, Vol. 29.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., & Yin, K., 2003, A review of process fault detection and diagnosis Part III: Process history based methods, Computers and Chemical Engineering, 27, 327–346.

Wang, D., & Romagnoli, J. A., 2005, A Robust Discriminate Analysis Method for Process Fault Diagnosis. In L. Puigjaner & A. Espuña (Eds.), European Symposium on Computer Aided Process Engineering – 15.

Yélamos, I., Escudero, G., & Graells, M., 2006, Fault diagnosis based on support vector machines and systematic comparison to existing approaches. In W. Marquardt & C. Pantelides (Eds.), 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering, pp. 1209–1214.