European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

Wavelet-Based Nonlinear Multivariate Statistical Process Control

Abd Halim S. Maulud*, Dawei Wang and Jose A. Romagnoli

Process System Engineering Laboratory, Department of Chemical Engineering, University of Sydney, NSW 2006, Australia

Abstract

In this paper, an approach of wavelet-based nonlinear PCA for statistical process monitoring is presented. The strategy utilizes the optimal wavelet decomposition in such a way that only approximation and the highest detail functions are used thus simplifying the overall structure and making the interpretation at each scale more meaningful. An orthogonal nonlinear PCA procedure is incorporated to capture the nonlinearity characteristic with minimum number of principal components. The proposed nonlinear strategy also eliminates the requirement of nonlinear functions relating the nonlinear principal scores to process measurements for *Q*-statistics as in other nonlinear PCA process monitoring approaches.

Keywords: Fault detection, Orthogonal Nonlinear PCA, Optimal multivariate wavelet decomposition

1. Introduction

Principal component analysis (PCA) has been widely applied in multivariate statistical process monitoring due to its capability to extract information in multivariate environment. In practice, many processes do exhibit significant nonlinearity correlation and in these cases a linear PCA mapping results in substantial loss of information or large numbers of linear components are required to obtain the required accuracy. Several nonlinear PCA have been proposed in the literature to improve the data extraction when the nonlinear correlations among the variables exist. Wavelets have also become a promising tool in many applications due to its property of time-frequency localization. However, up to date few combinations of PCA and wavelets have been reported for process monitoring. A multi-scale-PCA strategy (Bakshi 1998; Misra et al. 2002) has been proposed to capture information in both time and frequency domains and in addition, some wavelet nonlinear PCA combination have been reported (Shao et al. 1999; Palazoglu et al. 2001)

Despite these developments there are still some issues remaining on the development and implementation of wavelet based nonlinear PCA approaches. Orthogonality is a problem when dealing with nonlinear extensions of PCA strategies. In a conventional nonlinear PCA (NLPCA), the data information tends to be evenly distributed among the

* Author/s to whom correspondence should be addressed: halim@chem.eng.usyd.edu.au

principal components or bottleneck layer (Chessari et al. 1995). Thus, loosing the orthogonally characteristics inherent of linear PCA. On the other hand, when wavelet decomposition is applied in process measurements, the signals are projected onto the approximation and detail coefficients. As the signals are decomposed by using discrete wavelet transform to multi-level decompositions, more information is being transferred from approximation function to detail functions. If the signal is overly decomposed, it will eventually smoothing out the underlying features in the approximation function.

In this paper, an approach of wavelet-based nonlinear PCA for statistical process monitoring is presented. A significant contribution of the proposed strategy is the identification/definition of the optimal level of decomposition making possible to use only the approximation reconstruction and the highest level detail (the lowest frequency) reconstruction for process monitoring. This strategy improves the fault interpretation by separating the deterministic features and localized events. A modified Kramer's NLPCA (Kramer 1991) called orthogonal non-linear PCA (O-NLPCA) is incorporated into the strategy to improve the explained variance by maximizing the data variability in the first few principal. In addition, the number of principal components specified can be relaxed while in NLPCA the number of principal components must be optimally specified in advance. The proposed nonlinear strategy structure also eliminates a requirement of nonlinear functions relating the nonlinear principal scores to process measurements for *Q*-statistics monitoring as in other nonlinear PCA approaches.

2. Orthogonal Nonlinear PCA

Let X be a data matrix with n number of observations and m number of dimensions. In PCA, the X matrix can be decomposed into two matrices as follows:

$$X = TP^T = \sum_{i=1}^{m} t_i p_i^T \tag{1}$$

where T and P are called scores and loadings matrix, respectively. If the variables in X are collinear, the first f principal components can sufficiently explain the variability in data X. Thus, the data X can be written in term of a residual, E, as;

$$X = T_f {P_f}^T + E = \sum_{i=1}^{f} t_i p_i^T + E \tag{2}$$

In order to improve the orthogonality property in NLPCA, the Gram-Schmidt training algorithm is applied in NLPCA in such that the nonlinear score produced are orthogonal at the end of training session (Chessari et al. 1995). A major disadvantage of this scheme is it may suffer a constraint of trade-off between the main objective (overall convergence) and the secondary objective (orthogonal principal components). In addition, the network training is quite complex as it involves iterative procedure. In this paper, a simpler alternative approach to orthogonal NLPCA is proposed. This approach utilizes the Hammerstein model concept by incorporating linear PCA into the NLPCA. In Hammerstein model, the nonlinear and linear parts are separated into two blocks as shown in figure 1. In this case, the bottleneck layer nodes are called non-orthogonal nonlinear principal components and their outputs are called non-orthogonal scores.

When linear PCA is applied on non-orthogonal score, it will produce orthogonal nonlinear principal components as shown in figure 2.

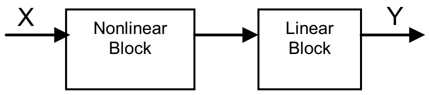

Figure 1: The Hammerstein Model

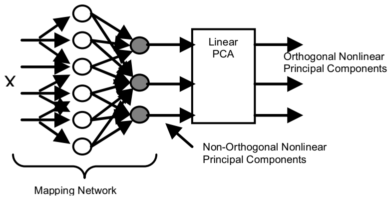

Figure 2 : Orthogonal Nonlinear PCA

Let T is the non-orthogonal nonlinear scores matrix generated at the output of bottleneck layer. Thus, the non-orthogonal matrix T can be transformed to orthogonal matrix U;

$$T = UP^T \tag{3}$$

where P is the eigenvector matrix. Another advantage of this approach compared to conventional NLPCA is a number of bottleneck layer neurons can be relaxed as long as the number is reasonably selected.

3. Optimal Wavelet Decomposition

In multi-resolution analysis theory (Mallat 1989), any signal can be approximated by successively projecting it down onto scaling function and wavelet function. The projections onto the scaling functions are known as scaling or approximation coefficients. The projection onto the wavelet functions are known as the wavelet or detail coefficients which capture the details of the signal lost when moving from an approximation at one scale to the next coarser scale.

Let us define the optimal decomposition level as the highest decomposition level in which the approximation function is adequately representing the actual deterministic signal with a minimum noise for given wavelet type. In a multivariate case, each variable may have a different optimal decomposition level. In most practical application, only a single decomposition level will be applied on all variables for computational simplification. Thus, the decomposition level selected must be appropriate or optimal to ensure that the underlying features of each variable are adequately preserved in approximation function with minimum noises.

In this paper, a graphical method based on PCA concept is presented to determine the optimal decomposition level for multivariate system. It is assumed that the deterministic features of the process data are continuous smooth functions. Let a PCA is being applied on the approximation reconstruction function with single-level wavelet decomposition. Recall equation (2), the E matrix consist mainly of noises. As the wavelet decomposition is recursively applied, the magnitude of matrix E is reduced as more noises have been captured by the detail functions. However, the compositions of the f principal components remain more or a less constant as the underlying features are still adequately preserved. As the decomposition level increases, at the l level some of the underlying features of the signals in the approximation function start to be lost to the detail function and the composition of the f principal components start to change significantly. This change can be detected by plotting the explained variance of the first principal component of the approximation reconstruction function for different decomposition level. An example plot is shown in figure 4. The explained variance is calculated based on the eigenvalue of the approximation reconstruction covariance matrix. The optimal decomposition level is given as (l-1) which is can be determined from the plot.

4. Wavelet-Based Nonlinear Statistical Process Control

Wavelet-based nonlinear statistical process control (WNSPC) combines the optimal wavelet decomposition ability to extract the deterministic features of the process variables in the approximation reconstruction with the ability of orthogonal nonlinear PCA to extract the correlation among the variables. The proposed monitoring strategy is shown in figure 3.

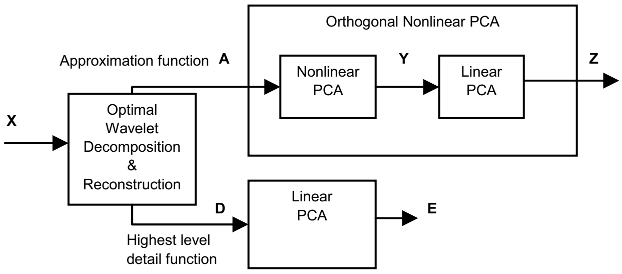

Figure 3. Wavelet Based Nonlinear Statistical Process Control

The measurement data matrix, X, is decomposed and reconstructed by using orthonormal wavelet at optimal decomposition level. Most of the noises will be filtered in this process. Only the approximation function and the lowest frequency detail function (the highest level) are retained for process monitoring. The advantage of this strategy is that the interpretation of a fault detected is more meaningful and simpler since only bi-scale is utilized. The approximation function represents the cleaned process deterministic features while the detail function represents the localized activities or stochastic features. In the multi-scale strategy, the process deterministic features are decomposed onto multi-scale which makes it difficult for interpretation. In addition, cleaned measurement data in the approximation function improves the data analysis. The orthogonal nonlinear PCA is applied onto the approximation reconstruction function as the nonlinearity characteristics are exhibited in the deterministic features. For the detail reconstruction function, a linear PCA is utilized as mostly noise is expected to be present. It is assumed that Y adequately represents the measurement data X in a compact form. Thus, the multivariate statistical process monitoring can be performed in Y instead of X. As a result, this strategy reduces the overall monitoring system to a conventional linear PCA system. The Q-statistic can be computed as in a conventional linear PCA manner between Y and Z. For the detail function, it is not necessary to apply Q-statistic because it mainly contains noises.

5. Simulation Case Study

An irreversible exothermic reaction of A→ B is conducted in an ideal CSTR. There are 9 variables monitored and 1000 data points are generated from the simulation and are corrupted with random noise with zero mean. The Daubechies-4 wavelet is utilized in this application. Figure 4 shows the explained variance for the first principal component of approximation reconstruction at different decomposition level. Level 8 is selected as the optimal decomposition level based on normal operating data. Two types of fault have been introduced. The first fault is a step mean shift in the inlet concentration of component A, between samples 200 and 399. The second fault is a drift deterioration of the overall heat transfer coefficient of the cooling system starting from a sample 601 onward. 95% and 99% significance points are considered as warning and action limits, which are indicated by dotted and solid lines, respectively. The WNSPC utilizes three control charts (T^2-approximation, Q-approximation and T^2-detail) compare to two control charts in conventional PCA (T^2 and Q).

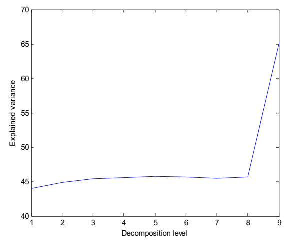

Figure 4: E.V. of 1st PC

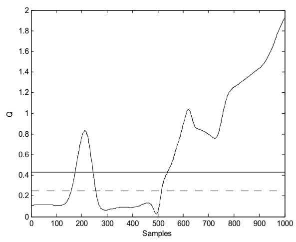

Figure 5b: Q-approx.

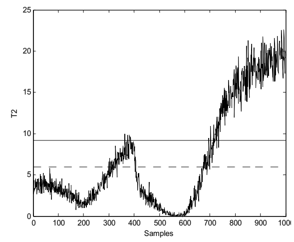

Figure 6a: T^2-PCA

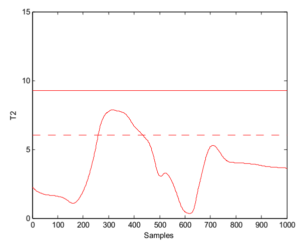

Figure 5a: T^2 approx.

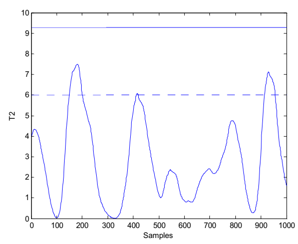

Figure 5c: T^2-detail

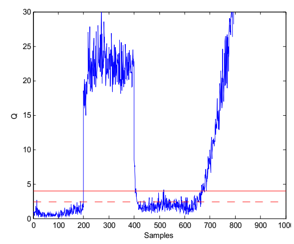

Figure 6b: Q-PCA

From figure 5 through 6, the mean shift can be easily detected by both WNSPC and PCA methods. This can be seen in Q-statistic plots which indicate a sharp increase that goes beyond the 99% limit at sample 200. For the T^2-plots, both methods cross the 95% limit, but the WNSPC response is faster. A sudden change in mean is also warned by the T^2-detail at sample 200 and 400. An advantage of O-NLPCA is being applied in the WNSPC approach can be clearly seen in the second fault. The fault is quickly detected by Q-approximation as it occurred by crossing the 95% and 99% limits at sample 515 and 539, respectively. While Q-PCA is only crossing the 95% and 99% limits at 628 and 668 respectively. This indicates that WNSPC approach is much faster compared to PCA in detecting a gradual process of characteristic change.

A utilization of T^2-detail plot gives some advantages. In addition its ability to warn any sudden or sharp changes (as illustrated in figure 5c), it has a potential to warn some localized events. A simulation of sensor stalled of outlet A concentration is performed from point 201 to 400 as shown in figure 7. Dotted line indicates the actual deterministic path. Both T^2-approximation and Q-approximation (figure 8a-b) do not indicate any warning regarding this fault as it does not involve the mean shift or process characteristic change. However, the T^2-detail (figure 8c) quickly warns that there is some localized event taking place. For PCA, the *Q*-PCA does indicate some warning since some points are crossing 95% limit from points 201 to 400. However, an interpretation of fault type can not be quickly addressed. This illustrates an additional advantage of this strategy as it is able to provide an initial guess of the fault types.

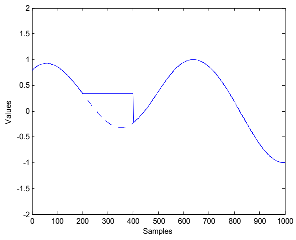

Figure 7: Outlet A concentration

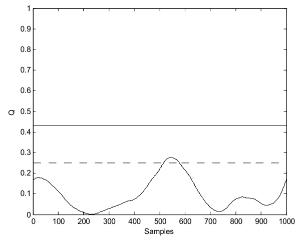

Figure 8b: Q-approximation

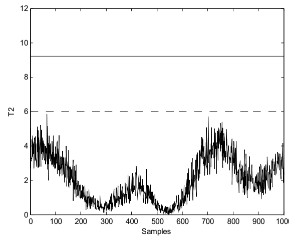

Figure 9a: T^2-PCA

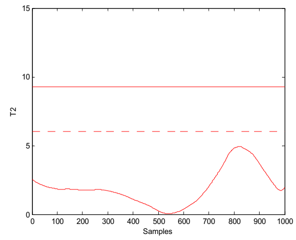

Figure 8a: T^2-approximation

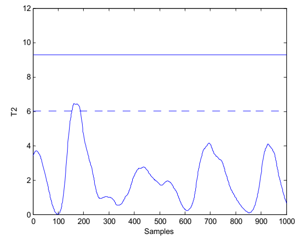

Figure 8c: T^2-detail

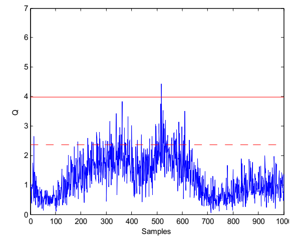

Figure 9b: Q-PCA

6. Conclusion

Wavelet-based Nonlinear Statistical Process Control strategy has been presented which utilizes optimal wavelet decomposition and orthogonal NLPCA. The strategy provides an alternative approach to uncover the nonlinear characteristic and simultaneously provide an initial fault interpretation platform.

References

Bakshi, B. R., 1998, AIChE Journal 44(7), 1596-1610.

Chessari, C. J., G. W. Barton and P. Watson, 1995, IEEE International Conference on Neural Networks 1, 183-188.

Kramer, M. A., 1991, AIChE Journal 37(2), 233-243.

Mallat, S. G., 1989, IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7), 674-693.

Misra, M., H. H. Yue, S. J. Qin and C. Ling, 2002, Computers & Chemical Engineering 26(9), 1281-1293.

Palazoglu, A., F. Doymaz, W. Sun, A. Bakhtazad and J. Romagnoli, 2001, ESCAPE-11 supplementary, 115-120.

Shao, R., F. Jia, E.B. Martin and A.J. Morris, 1999, Control Engineering Practice 7(7), 865-879.