Accepted Manuscript

Title: Improved nonlinear fault detection strategy based on the Hellinger distance metric: Plug flow reactor monitoring

Author: Fouzi Harrou Muddu Madakyaru Ying Sun

PII: S0378-7788(16)30852-0

DOI: <http://dx.doi.org/doi:10.1016/j.enbuild.2017.03.033>

Reference: ENB 7458

To appear in: *ENB*

Received date: 17-9-2016

Accepted date: 13-3-2017

Please cite this article as: Fouzi Harrou, Muddu Madakyaru, Ying Sun, Improved nonlinear fault detection strategy based on the Hellinger distance metric: Plug flow reactor monitoring, *Energy & Buildings* (2017), <http://dx.doi.org/10.1016/j.enbuild.2017.03.033>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Improved nonlinear fault detection strategy based on the Hellinger distance metric: plug flow reactor monitoring

Fouzi Harrou^a, Muddu Madakyaru^b, Ying Sun^a

^a*King Abdullah University of Science and Technology (KAUST)*

Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, Thuwal 23955-6900, Saudi Arabia

Tel: +966.12808.0602, Fax: +966.12802.1386, E-mail: fouzi.harrou@kaust.edu.sa

^b*Department of Chemical Engineering, Manipal Institute of Technology, Manipal University, Manipal, India*

Abstract

Fault detection has a vital role in the process industry to enhance productivity, efficiency, and safety, and to avoid expensive maintenance. This paper proposes an innovative multivariate fault detection method that can be used for monitoring nonlinear processes. The proposed method merges advantages of nonlinear projection to latent structures (NLPLS) modeling and those of Hellinger distance (HD) metric to identify abnormal changes in highly correlated multivariate data. Specifically, the HD is used to quantify the dissimilarity between current NLPLS-based residual and reference probability distributions obtained using fault-free data. Furthermore, to enhance further the robustness of these methods to measurement noise, and reduce the false alarms due to modeling errors, wavelet-based multiscale filtering of residuals is used before the application of the HD-based monitoring scheme. The performances of the developed NLPLS-HD fault detection technique is illustrated using simulated plug flow reactor data. The results show that the proposed method provides favorable performance for detection of faults compared to the conventional NLPLS method. *Keywords:* Anomaly detection, Hellinger distance, Nonlinear PLS, Nonlinear processes, Multiscale filtering.

1. Introduction

Process monitoring is required by various process industries for improving the quality of products and enhancing process safety [1, 2]. Among the most important applications of process safety are those related to environmental and industrial processes. A critical anomaly in a chemical or a petrochemical process not only may cause a degradation in the process performance or lower its product quality; it can result in catastrophes that may lead to fatal accidents and substantial economic losses [3]. Therefore, detecting anomalies in chemical processes is vital for their safe and proper operations. Moreover, proper operation of complex engineering and chemical processes, such as those in the oil and gas industries, requires careful monitoring

of certain key process variables to enhance the productivity of these processes and more importantly to avoid disasters in the cases of failure [3]. Generally, anomalies in modern automatic processes are difficult to avoid and may generate catastrophic process degradations. Possible anomalies could be a result of malfunctioning sensor/s (called sensor anomalies) or to abnormal changes in the process. Sensor anomalies are usually quantified by sudden (or quick) changes in a small number of process variables. Process anomalies, on the other hand, are abnormal changes caused by deviations in the process itself. These anomalies are usually quantified by slow drifts across several process variables. Thus, anomaly detection and diagnosis are two vital components of process monitoring, during which anomalies are first identified and then isolated to ensure that they can be appropriately handled. Of course, as a result of accurate and prompt anomaly detection, downtime is minimized, safety of process operations is improved, and manufacturing costs are reduced [4, 5]. Such monitoring is an important task not only to maintain high quality products and to enhance operational process efficiency and profitability, but also for people's safety.

To improve the reliability, safety and efficiency of advanced supervision methods, anomaly detection and diagnosis have become more and more important for numerous technical processes. Indeed, modeling and monitoring chemical and petrochemical processes is a challenging task because of the complexity and sometimes the lack of understanding about these processes. In these cases, data-based monitoring techniques are more commonly used [6]. Moreover, data-based techniques provide efficient tools for extracting useful feature for design of monitoring schemes based on empirical models derived from the available process data [6, 7, 8, 9]. Such methods require a minimal a priori knowledge about process physics, but depends on the availability of quality input data. A detailed overview of data-based monitoring techniques is presented in [10, 7]. Principal component analysis (PCA) and projection to latent structures (PLS) are two basic methods of multivariate analysis and reputed as powerful tools for monitoring multivariate processes with highly correlated process data [5, 11]. Furthermore, they are among the data-driven methods, which have gained a remarkable acceptance in industry for statistical monitoring and control of multivariate processes [12]. Unlike PCA, PLS finds an optimum pair of latent variables both in predictor (input) and predicted (output) variables such that these transformed variables have the largest covariance [13, 11]. Despite their simplicity, these linear latent variable regression (LVR) methods are based on assumption that process variables are linear and may not be appropriate for nonlinear process. However, most engineering and chemical processes are nonlinear. For highly nonlinear data the limitations of such linear methods are obvious. The use of linear LVR models (PLS and PCA) to approximate the nonlinearity may not be adequate; therefore in this situation the use of a nonlinear model is required to bypass such deficiency.

For nonlinear input-output systems, nonlinear PLS (NLPLS) has been extensively used to extract the linear and nonlinear relations between the process variables. Numerous variants of nonlinear PLS were developed for process monitoring, including the quadratic PLS (QPLS) [14, 13, 15], spline PLS[16], neural network PLS[17, 18], kernel PLS [19] and fuzzy PLS [20, 21]. Furthermore, the NLPLS methods have been found useful in many applications, mainly in industrial manufacturing and production processes, including chemicals, power industries and microelectronics. Usually, for anomaly detection purposes, NLPLS is first performed on the normal operating data (fault-free data) enabling us to obtain a reference model that will be used next for anomaly detection. Abnormal events are detected if the measurements deviate from the region of normal operation in latent space or in the residual space. The early and accurate detection of incipient anomalies represent an important step ahead that helps to design safe and reliable systems. However, conventional NLPLS-based monitoring indices such as T^2 and Q charts lose the ability to detect small changes in the mean of process data. Alternatively, one may use a NLPLS model with a distributional divergence-based monitoring metric, the Hellinger Distance (HD) to remedy the drawbacks described above and enhances the sensitivity for incipient change detection.

In this paper, an efficient anomaly detection scheme to identify incipient anomalies in highly correlated multivariate data is presented. We propose a statistical strategy that combines the flexibility of a NLPLS modeling approach, the greater sensitivity of HD metric to quantify dissimilarity between two distributions, and that may be efficiently enhance the sensitivity for incipient change detection. In such framework, NLPLS model is used to capture nonlinear relationships in multivariate input-output data to build predictive models. Specifically, in such NLPLS model, an optimal polynomial function in each inner relation via cross validation method is used to describe the inner relation of PLS model [22]. In order to monitor a nonlinear multivariate process and to achieve greater sensitivity to incipient anomalies, the HD-based detection index is used to identify deviation between actual NLPLS-based residuals and reference probability distributions. The HD metric has been used because it is very useful to quantify the similarity between two probability distributions. However, the presence of measurement errors (noise) in the data and model uncertainties degrade the quality of fault detection techniques. The second objective of this paper is to utilize wavelet-based multiscale representation of data [23] to further enhance the effectiveness of the proposed method. Specifically, multiscale data pre-filtering will be used before applying the HD-based fault detection method to enhance the robustness of this method to measurement noise. We apply the proposed model to monitor a plug flow reactor and compare its performance with convectional NLPLS-based methods. Obtained results show the ability of the new NLPS-HD index in detecting incipient anomalies, pointing out the promising

application of these statistical tool in the supervision of nonlinear multivariate processes.

The rest of the paper is structured as follows. In Section 2, a brief description of linear PLS and detailed discussion on NLPLS and fault detection methods are discussed. In Section 3, some backgrounds of the Hellinger distance and its use for anomaly detection purpose, are briefly described. Then, in Section 4, a brief introduction to multi-scale filtering is presented. Next, in Section 5, we assess the proposed scheme and present some simulation results. Finally, some conclusions are given in Section 6.

2. Preliminary materials

In this paper, our recent results towards aiming at improving fault detection will be presented. Specifically, to achieve greater sensitivity to small changes in nonlinear processes a new NLPLS-based HD anomaly detection strategy will be described. The developed method utilizes NLPLS as a modeling framework and uses HD-based index for fault detection. The performance of the developed fault detection strategy will be illustrated and compared to that of the conventional NLPLS method. One applications will be presented, using simulated plug flow reactor (PFR) data. However, since NLPLS is utilized in this developed anomaly detection method, a brief introduction to NLPLS and how it can be used in anomaly detection, will be presented next.

2.1. Linear PLS model

For input-output models, PLS has been widely used to extract linear relationships between two sets of variables, inputs and outputs. PLS is usually used to predict the dependency of certain key variables that are challenging or expensive to measure (called outputs) from other process variables that are easier to measure (called inputs). Examples where PLS models can be used include predicting the ozone concentration from measurements of various meteorological conditions (such as temperature, wind direction, etc) or predicting the compositions in a distillation column using temperature and pressure measurements at different trays of the column.

Consider a pair of datasets $\mathbf{X} \in \mathbb{R}^{N \times M}$ and $\mathbf{Y} \in \mathbb{R}^{N \times 1}$, where $\mathbf{X}$, $\mathbf{Y}$ are the input and output variables, respectively. After the data standardization by first subtracting the sample mean of the training data and then dividing by the sample standard deviation of the training data, PLS projects $\mathbf{X}$ and $\mathbf{Y}$ on to a lower

dimension subspace defined by number of latent variable $[z_1, z_2, \dots, z_l]$ as follows:

$$\begin{cases} \mathbf{X} = \mathbf{Z}\mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} = \mathbf{Z}\mathbf{Q}^T + \mathbf{F} \end{cases} \quad (1)$$

Where $\mathbf{Z} \in \mathbb{R}^{N \times l}$ (l is the number of latent variable) is the score matrix represents the projection of the variables on the subspace, $\mathbf{P} \in \mathbb{R}^{M \times l}$ represents the loading matrix for $\mathbf{X}$ and $\mathbf{Q} \in \mathbb{R}^{1 \times l}$ represents the loading matrix for $\mathbf{Y}$. $\mathbf{E}$ and $\mathbf{F}$ represents the model residue of input and output respectively. PLS calculates the input loading vectors, $\mathbf{P}_i$, so that the covariance between the estimated latent variable $\hat{\mathbf{Z}}_i$ and model output, $\mathbf{Y}$, i.e., [24]:

$$\hat{\mathbf{P}}_i = \arg \max_{\mathbf{P}_i} \text{cov}(\mathbf{Z}_i, \mathbf{Y}) \quad (2)$$

can be maximized with constraint

$$s.t. \quad \mathbf{P}_i^T \mathbf{P}_i = 1; \quad \mathbf{Z}_i = \mathbf{X}\mathbf{P}_i$$

where, $i = 1, \dots, l$, $l \leq m$. Note that PLS utilizes an iterative algorithm [24] to estimate the latent variables, where one latent variable or principal component is added iteratively to the model. After the inclusion of a latent variable, the input and output residuals are computed and the process is repeated using the residual data until a cross validation error criterion is minimized [25].

2.2. Nonlinear PLS model

PLS has been extensively used as data analysis tool because of its well developed statistical foundation in the literature. Though it is used but it is limited to modeling linear relationship and most of the plant data are non linear in nature. For nonlinear processes, however, the linearity assumption made in the above PLS model formulation may not provide satisfactory model predictions. In these cases, nonlinear PLS models would be needed. In the proposed nonlinear PLS modeling, outer transformation of linear PLS retained to get the robust properties of linear PLS. The inner relation of the linear PLS is modeled by an optimal polynomial function, $f(\cdot)$. The advantage of doing is that dimension reduction method to eliminate collinearity among the data is achieved by the outer transformation of the linear PLS and nonlinear relationship is captured in the latent variable space. Specifically, the latent variables ($\mathbf{Zc}_k$) obtained by the outer transformation of the linear PLS is used map the output variable $\mathbf{Y}$ using optimal polynomial function. which is described as,

$$\mathbf{Y}_k = f(\mathbf{Zc}_k) + \mathbf{R}_k. \quad (3)$$

where $\mathbf{R}_k$ represents the model error and $f(\cdot)$ is a polynomial function. The parameters of the optimal polynomial function are selected by minimizing $\mathbf{R}_k$. The major advantage of the proposed NLPLS is that the outer layer of the linear PLS decomposition properties retained so that variables are projected into the directions maximizing the covariance. The schematic representation of the NLPLS algorithm is presented in Figure 1 and summarized in Table 1.

Table 1: Nonlinear PLS modeling algorithm.

-
- (1) Given the raw input $\mathbf{X}$ output $\mathbf{Y}$ data in to zero mean and unit variance.
 - (2) At $k=1$, $\mathbf{E}_{\mathbf{X}_k}=\mathbf{X}$, $\mathbf{E}_{\mathbf{Y}_k}=\mathbf{Y}$
 - (3) Determine the loading vector $\mathbf{P}_k$ and score vector (latent variable) $\mathbf{z}_k$.
 - (4) Continue the step A-3 till it converges to the required threshold. This step uses the nonlinear iterative least square (NIPALS) algorithm [26].
 - (5) Determine the optimal polynomial structure for inner relation $f(\cdot)$, which predicts the output $\mathbf{E}_{\mathbf{y}_k}$ with the input score vector (latent variable) $'Z'_k$ has the

$$E_{Y_k} = \xi_0 + \xi_1 Z c_k + \xi_2 Z c_k^2 + \xi_3 Z c_k^3 + \cdots + \mathbf{R}_k. \quad (4)$$

The model is estimated by minimizing the regression vector $\mathbf{R}_k$ such that it does parsimonious in parameters. This part is computed as follows:

- i) Select the data (i.e., latent variable $\mathbf{Z}c_k$ and output $\mathbf{E}_{\mathbf{Y}_k}$
 - ii) Choose a order of the polynomial model structure
 - iii) Split the univariate data into Training and Testing data
 - iv) Estimate the parameters of the polynomial model
 - v) Determine the mean square error using testing data set
 - vi) Return to step (ii) with higher order model structure
 - vii) Repeat the procedure iii to v.
 - viii) Decide the model structure based on the minimum MSE criteria.
- (6) Determine the input and output residuals i.e., $\mathbf{E}_{X_k} = \mathbf{X} - \hat{\mathbf{X}}$; $\mathbf{E}_{Y_k} = \mathbf{Y} - \hat{\mathbf{Y}}$
 - (7) Let $k=k+1$, then return to (3) until all score vectors (latent variables) are calculated. The number of latent variables are optimized and compared using cross validation, by minimizing the output prediction mean squared error (MSE) using unseen data set.
-

Figure 1: Nonlinear PLS modeling algorithm

2.3. NLPLS-based monitoring

After a model is obtained using NLPLS method, various methods for fault detection can be applied. Two monitoring statistics, the T^2 and Q statistics, are usually utilized for fault detection purpose [27] (see Figure 2). First, Hotelling T^2 statistics indicates the variation within the process model in LVs space. The other is the Q statistic, also known as the squared prediction error (SPE), which monitors how well the data conforms to the model.

Figure 2: A flowchart of a NLPLS-based fault detection scheme.

The T^2 statistic is used in detecting anomalies corresponded at atypical variations within a the latent variable subspace (scores space). T^2 can be computed as [27]:

$$T^2 = \sum_{i=1}^l \frac{t_i^2}{\sigma_i^2}, \quad (5)$$

where, σ_i^2 represents the estimated variance of the corresponding latent variable t_i . When the monitored

process is conform with the desired performances at the $i - th$ time point, it is obvious that $T^2 < T_{l,n,a}^2$, where $T_{l,n,a}^2$ is a control limit given in [27]. The T^2 chart gives a signal of the presence of an anomaly when $T^2 > T_{l,n,a}^2$. It has been shown in [11], that the T^2 statistic can result in missed detection due to the latent space sometimes being insensitive to changes with small magnitude, which is because each latent variable is a combination of all process variables.

The squared prediction error (SPE) or Q statistic, on the other hand, which is defined as [27]:

$$Q = \mathbf{e}^T \mathbf{e} \quad (6)$$

captures the changes in the residual subspace. Where $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ represents the residuals vector, which is the difference between the new observation, $\mathbf{x}$, and its prediction, $\hat{\mathbf{x}}$, via MSPLS model. Equation 6 provides a direct mean of Q statistic in terms of measured total sum of variation in the residual vector $\mathbf{e}$. Once Q statistic of the new observation exceed the control limits Q_α at significance level α given in [27]. Figure 3 shows an example of a data with one observation has a large Q value while an other with large T^2 .

Figure 3: A set of projected observations (green) and two original three-dimensional observations a and b; T^2 and Q metrics.

The main drawback of such monitoring charts, T^2 and Q , is the lack to detect small faults because treat they each observation individually and don't take into account information from past data [28]. That makes them insensitive to small faults in process variables and causes many missed detections and false alarms [28]. Furthermore, it has been shown in the literature that the T^2 chart is not effective to detect moderate or small changes in multivariate data [29]. Therefore, only the Q -based chart will be used in this study as a benchmark for fault detection with NLPLS. To remedy these drawbacks and achieve greater sensitivity to incipient anomalies, we propose an alternative anomaly detection chart based on a HD metric. More details about the HD metric and how it can be used in anomaly detection are presented next.

3. Hellinger distance-based monitoring scheme

In this paper, a distance based fault detection method that are based on NLPLS models will be developed in order to achieve enhanced detection performance compared to the conventional NLPLS fault detection method. Indeed, the proposed monitoring strategy combines modeling using NLPLS modeling with the Hellinger distance (HD) based monitoring index. HD metric, which is one of the most common measures of distance between probability distributions, is central to the problems of inference and discrimination [30, 31, 32]. Specifically, HD metric, which was initially developed by Ernest Hellinger, is an important statistical measure that can be used to quantify the dissimilarity or closeness between two probability density functions (PDFs) [33, 31, 32]. This measure has been used extensively by scientists and engineers in various disciplines, including pattern recognition [34], image processing [35], classification [36, 37], and anomaly detection [38, 32, 39]. In addition, it has been used effectively within the ecological domain [40], network intrusion detection [39], and fraud detection in insurance applications [41]. Here, we address the problem of anomaly detection following a HD-based approach.

Definition 1 (Hellinger distance). *Let us consider $P_1(x)$ and $P_2(x)$ to be two probability distributions with probability density functions $p_1(x)$ and $p_2(x)$, respectively. The squared HD of $p_1(x)$ relative to $p_2(x)$, which is a measure of the dissimilarity between $p_1(x)$ and $p_2(x)$, is defined by:*

$$HD^2(p_1(x), p_2(x)) = \frac{1}{2} \sum (\sqrt{p_1(x)} - \sqrt{p_2(x)})^2. \quad (7)$$

which can be considered as the Euclidean norm of the difference of the square root vectors:

$$HD^2(p_1, p_2) = \frac{1}{\sqrt{2}} \|\sqrt{p_1} - \sqrt{p_2}\|_2. \quad (8)$$

It is important to note that the HD measure is a symmetric distance or metric in the Euclidean sense (i.e. $HD^2(p_1, p_2) = HD^2(p_2, p_1)$). It is non-negative and bounded (i.e., the HD satisfies $0 \leq HD^2(p_1, p_2) \leq 1$) and null only when the two densities are equal, $p_1 = p_2$.

The closed form expression of HD can easily be computed in the case of normal distributions. For univariate normal distributions, $p_1(x)$ and $p_2(x)$ of a random variable x , where $p_1 \sim \mathcal{N}(\mu_0, \sigma_0)$ and $p_2 \sim \mathcal{N}(\mu_1, \sigma_1)$, where μ_0 and μ_1 are the means and σ_0^2 , σ_1^2 are the variances for p_1 and p_2 , the HD between p_1 and p_2 is given by:

$$HD^2(p_1, p_2) = 1 - \sqrt{\frac{2\sigma_0\sigma_1}{\sigma_0^2 + \sigma_1^2}} \exp\left(-\frac{1}{4} \frac{(\mu_0 - \mu_1)^2}{\sigma_0^2 + \sigma_1^2}\right). \quad (9)$$

When changes affect only on the mean (i.e., $\sigma_1^2 = \sigma_0^2$), Equation (9) can be rewritten as:

$$HD^2(p_1, p_2) = 1 - \exp\left(-\frac{1}{8} \frac{(\mu_0 - \mu_1)^2}{\sigma_0^2}\right). \quad (10)$$

Lemma 1. *The squared Hellinger distance between two multivariate normal distributions $P_1 \sim \mathcal{N}(\mu_1, \Sigma_1)$ and $P_2 \sim \mathcal{N}(\mu_2, \Sigma_2)$ yields also a closed form expression [42]:*

$$H^2(P_1, P_2) = 1 - \frac{\det(\Sigma_1)^{1/4} \det(\Sigma_2)^{1/4}}{\det\left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{1/2}} \exp\left\{-\frac{1}{8}(\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1} (\mu_1 - \mu_2)\right\} \quad (11)$$

Note that the HD measure in the case of non-Gaussian variables requires that non-Gaussian distributions to be used instead. The HD metric can be computed in closed form for exponential distributions, Weibull distributions, Poisson distributions and Beta distributions. In many application P_1 and P_2 are generally not available and must be estimated from the data sets. In such a case, the HD measure has no closed form and the function given by Equation (7) should be numerically approximated [43].

More importantly, two similar distributions will have a small HD close to zero due to measurement noise and errors, while very different distributions would have a larger HD. In other words, the smaller the HD measure, the more similar the distributions, and vice versa. It is this comparison operation that makes it a useful indicator of anomaly detection. Since HD is a useful tool for measuring the dissimilarity between two distributions, it can be a suitable statistic for anomaly detection. More specifically, the HD can be used as an anomaly indicator by comparing the statistical similarity between residual distributions before and after an anomaly. Therefore, it seems meaningful to adopt a NLPLS-based HD scheme for statistical process monitoring instead.

The main objective of this work is to exploit the advantages of the HD metric and those of NLPLS modeling to achieve enhanced detection performance compared to conventional NLPLS-based methods. Subsequently, the HD metric will be integrated with NLPLS to extend its anomaly detection abilities for detecting incipient anomalies and to make further amelioration and broaden the practical employment of such method.

3.1. NLPLS-HD based fault detection scheme

In this subsection, a brief description of the NLPS-HD monitoring chart is introduced. The difference between the observed value of the output variable, y , and the predicted value, $\hat{y}$, obtained from NLPLS model represent the residual of the output variable, $\mathbf{F} = [f_1, \dots, f_t, \dots, f_n]$ which can be used as an indicator to

detect a possible anomaly. The residual $\mathbf{F}$ obtained from PLS model is assumed to be Gaussian. It is assumed that the anomaly affect the mean parameter of residual distributions and the variance is supposed unchanged after the anomaly occurrence. In this approach, the HD has been used to measure the distance between the probability distribution of current residual $p_0(f) \sim \mathcal{N}(\mu_0, \sigma_0^2)$ against a reference one $p_1(f) \sim \mathcal{N}(\mu_1, \sigma_0^2)$, where μ_0 and μ_1 are the means and $\sigma_0^2 > 0$ is the variance for $p_0(f)$ and $p_1(f)$. Then, the HD distance based on the residuals distributions of the response variables from the NLPLS model can be computed as follows: $HD^2(p_0(f), p_1(f)) = 1 - \exp(-\frac{1}{8} \frac{(\mu_0 - \mu_1)^2}{\sigma_0^2})$. Where μ_0 and σ_0 are the mean and the standard deviation of NLPLS-based residuals obtained with fault-free data. The normal operating conditions is guaranteed by a zero HD when both PDFs are equal. Although, in real situation should deviate from zero only due to measurement errors or modeling uncertainties. The HD-based test makes decision between the null hypothesis $\mathcal{H}_0$ (absence of anomalies) and alternative hypothesis $\mathcal{H}_1$ (presence of anomalies) by comparing between the decision statistic $HD^2(p_0(f), p_1(f))$ and a given value of the threshold h .

$$HD^2(p_0(f), p_1(f)) = \frac{(\mu_1 - \mu_0)^2}{\sigma_0^2} \begin{matrix} \geq \mathcal{H}_0 \\ \leq \mathcal{H}_1 \end{matrix} h. \quad (12)$$

For setting the detection threshold h , a simple approach based on the three-sigma rule was used.

$$h = \mu_0^{HD} + L\sigma_0^{HD} \quad (13)$$

where μ_0^{HD} and σ_0^{HD} are the mean and the standard deviation for the nominal behavior of the anomaly indicator HD, and L is the width of the control limits which determines the confidence limits, usually specified in practice as 3 for a false alarm rate of 0.27%. If the decision function $HD^2(p_0(f), p_1(f))$ is larger than the threshold h , the HD-based test decides for $\mathcal{H}_1$, otherwise $\mathcal{H}_0$ is assumed to be true.

The overall concept of the proposed method is illustrated in Figure 4. The developed anomaly-detection schemes based on the HD metric consist of two stages: training and testing. In the first stage, a reference NLPLS model is constructed in the case of healthy fault-free data. During the testing or monitoring stage, the NLPLS reference models previously obtained are compared with the new data and an anomaly statistical index is calculated and compared with the threshold previously computed using fault-free data.

Figure 4: Flowchart of the proposed NLPLS-based the HD distance metric.

4. Brief review of Multiscale filtering

Data observed from engineering processes are usually noisy and correlated in time, which makes fault detection more difficult because the presence of noise degrades detection quality and most methods are developed for independent observations. Wavelet-based multiscale modeling of data is an efficient tool for extracting feature that is suited well to denoising and decorrelating series time data [23]. In this direction, the multiscale filtering will be used to reduce measurement noise. More specifically, wavelet-based multiscale filtering of data will be utilized before the application of the HD-based monitoring scheme to ameliorate the robustness of this method to measurement noise, and reduce the false alarms due to modeling

errors.

4.1. Wavelet-based multiscale representation

Multiresolution time-series decomposition was initially applied by Mallat, who used orthogonal wavelet bases during data compression for image decoding [44]. Wavelets represent a family of basis functions that can be expressed as the following localized in both time and frequency [45]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (14)$$

where a is the dilation parameter, b is the translation parameter [46] and $\psi(t)$ is the mother wavelet. Both these parameters are commonly discretized dyadically as $a = 2^m$, $b = 2^m k$, $(m, k) \in \mathbb{Z}^2$, and the family of wavelets can be represented as $\psi_{mn}(t) = 2^{-\frac{m}{2}} \psi(2^{-m}t - m)$. Here, $\psi(t)$ is the mother wavelet and m and k are the respective dilation and translation parameters. Different families of basis functions are created based on their convolution with different filters, such as the Haar scaling function and the Daubechies filters [46, 47]. Parameters that are discretized dyadically force downsampling, reducing the number of parameters dyadically with every decomposition; However, dyadically discretized wavelets force samples at non-dyadic locations to become decomposed only after a certain time delay.

Using a discrete wavelet transform, an original signal space, S , can be decomposed into two sub-spaces: an approximation subspace, S_a , and detailed subspaces, S_d . The scale function $\phi_{j,k}(t) = \sqrt{2^{-j}} \phi(2^{-j}t - k)$, $k \in \mathbb{Z}$ and wavelet functions $\psi_{j,k}(t) = \sqrt{2^{-j}} \psi(2^{-j}t - k)$, $j = 1, \dots, J$, $k \in \mathbb{Z}$, where the coarsest scale J normally termed the decomposition level, span the approximation and detailed subspaces, respectively. Any signal can be represented by a summation of all scaled and detailed signals as follows [46]:

$$x(t) = \underbrace{\sum_{k=1}^{n2^{-J}} a_{Jk} \phi_{Jk}(t)}_{A_J(t)} + \sum_{j=1}^J \underbrace{\sum_{k=1}^{n2^{-j}} d_{jk} \psi_{jk}(t)}_{D_j(t)}. \quad (15)$$

where j , k , J and n represent the dilation parameter, translation parameter, number of scales, and number observations in the original signal, respectively [47, 44]. d_{jk} and a_{Jk} represent the scaling and the wavelet coefficients, respectively, and $A_J(t)$ and $D_j(t)$, ($j = 1, 2, \dots, J$) are the approximated signal the detail signals, respectively.

In other words, the detailed signal $D_j(t)$ at scale j can be obtained by passing the original and scaled signals through a high-pass filter (g), and the scaled signals are generated by passing the original and scaled

signals through a low-pass filter (h) [48]. A signal can be described at multiple resolutions by decomposing it on a family of wavelets and scaling functions. For example, consider the series time measurements of the feature indicator shown in Figure 5. The signals in Figures 5(b, d and f) are at increasingly coarser scales compared to the original signal in Figure 5(a).

Figure 5: Illustration of data representation at multiple scales of a heavy-sine signal.

4.2. Multiscale data filtering algorithm

Multiscale de-noising via wavelets is rested on the observation that random errors in a signal are present over all wavelet coefficients while deterministic changes get captured in a small number of relatively large coefficients [49, 50, 51]. A wavelet-based de-nosing algorithm comprises the following three main steps [50]:

- (1) Decompose the original signal at multiple scales via wavelet transform to obtain wavelet coefficient series in different level;
- (2) Select thresholds for each level and remove the wavelet coefficients that are below a threshold value;
- (3) Inverse wavelet transform based on the detail coefficients to obtain a de-noised signal.

Several threshold selection criterion have been proposed including Fixed Threshold [52], Rigorous Sure threshold, Heursure threshold and Min-max threshold. The simplest wavelet threshold method, which is proposed in [52], uses the same threshold to deal with coefficients in the expansion. It determines the threshold as following form:

$$t_j = \sigma_j \sqrt{2 \log n} \quad (16)$$

where, n denotes the length of the analyzed signal and σ_j is the standard deviation of the errors at scale j , which can be estimated from the wavelet coefficients at that scale by Equation 17,

$$\sigma_j = \frac{\text{median}(|d_{jk}|)}{0.6745}. \quad (17)$$

One of the biggest advantages of multiscale representation is its capacity to distinguish measurement noise from useful data features [53, 54], by applying low and high pass filters to the data during multiscale decomposition. This allows the separation of features at different resolutions or frequencies, which makes multiscale representation a better tool for filtering or denoising noisy data than traditional linear filters, like the mean filter and the exponentially weighted moving average (EWMA) filter [48]. The ability of multiscale representation to separate noise has been used not only to improve data filtering, but also to improve the prediction accuracy of several empirical modeling methods and the accuracy of state estimators. In the proposed HD-based fault detection chart, we can apply a filter just before plotting the results of monitoring chart to reduce false alarms and, at same time, to enhance the sensitivity to detecting minor faults.

5. Fault detection in plug flow reactor

Plug flow reactors consist of a hollow pipe or tube through which reactants flow [55]. Figure 6 shows a plug flow reactor in the form of a tube wrapped around an acrylic mold which is encased in a tank. Water at a controlled temperature is circulated through the tank to maintain constant reactant temperature. Plug flow reactors have a wide variety of applications in either gas or liquid phase systems. In this example, the ability of the fault detection method using NLPLS is tested on simulated plug flow reactor data. The dynamic model of the PFR in which two first order reaction take place as per the equation 18. The process is modelled by a system of first order PDE's shown in equation from 19 to 22. For the simulation purpose, the model parameters are used as listed in Table 2 [56].

$$\frac{\partial C_A}{\partial t} = -v_l \frac{\partial C_A}{\partial x} - k_{10} e^{-E_1/RT_r} C_A \quad (19)$$

$$\frac{\partial C_B}{\partial t} = -v_l \frac{\partial C_B}{\partial x} + k_{10} e^{-E_1/RT_r} C_A - k_{20} e^{-E_2/RT_r} C_B \quad (20)$$

$$\frac{\partial T_r}{\partial t} = -v_l \frac{\partial T_r}{\partial x} + \frac{\Delta H_{r1}}{\rho_m c_{pm}} k_{10} e^{-E_1/RT_r} C_A + \frac{\Delta H_{r2}}{\rho_m c_{pm}} k_{20} e^{-E_2/RT_r} C_B + \frac{U_w}{\rho_m c_{pm} V_r} (T_j - T_r) \quad (21)$$

$$\frac{\partial T_j}{\partial t} = -u \frac{\partial T_j}{\partial x} + \frac{U_{wj}}{\rho_{mj} c_{pmj} V_j} (T_r - T_j) \quad (22)$$

The above reaction is series in nature, where A is the reactant species, which converts in to desired product B. Which is further decomposed into undesired product C. The reaction system are endothermic, therefore fluid in the jacket of the reactor is circulated to maintain the reactor temperature at desired level. The flow rate of the jacket fluid (heating fluid) is manipulated to get the desired product concentration i.e., B. There are ten temperatures measuring device used to measure the temperature at different location of the reactor as shown in the Figure 6.

Figure 6: Plug Flow Reactor: Schematic diagram of plug flow reactor.

For the simulation purpose the reactor length converted to normalized length of $x \in [0, 1]$. The inlet concentration of the reactant and inlet temperature of the feed are defined as C_{A0} and T_{A0} . The flow rate of heating fluid in the jacket is u , which is used to regulate the temperature of the reacting mixture. The concentrations of reactant A and B are C_A and C_B in the reactor, the temperature inside the reactor is T and in the jacket T_j . The enthalpy of the two reactions as mentioned in the equation 18 are ΔH_{r1} and ΔH_{r2} . The density of fluid inside the reactor and heating fluid in the jacket are ρ_m and ρ_{mj} respectively. The heat capacity of the fluid in the reactor and in the jacket are c_{pm} and c_{pmj} , respectively. The size of the reactor and jacket interns of volume are V_r and V_j , respectively. The overall heater transfer coefficient

of reactor and jacket side are U_w and U_{wj} , respectively. The Arrhenius constants and the activation energy of the reactions 18 are k_{l0} , k_{20} , E_l , E_2 .

5.1. Data generation

The training and testing data are generated by perturbing the input (i.e., flow rate of the reactant feed) around the steady state nominal condition mentioned in the Table 2. The flow rate is perturbed using pseudo random binary signals (PRBS) signal of the frequency range $[0 \ 0.05 \ w_N]$, where $w_N = \pi/T$ represents the Nyquist frequency. In the modeling problem, ten temperatures at different location in the reactor (see Figure 7) and feed flow considered as inputs and product concentration considered as output (i.e., C_B). Measurements of ten temperatures and feed flow were labeled as $x_1, x_2, \dots, x_{11}$, respectively, and the product concentration is labeled as $y = C_B$. The data generated are noise free, which are then added with zero mean gaussian noise. The training and testing test data with Signal-to-Noise Ration (SNR) of 30 are shown in Figure 7.

Table 2: Plug flow reactor - model parameters.

Process Variable	Value	Process Variable	Value
v_l	1 ml/min	C_{pm}	0.231 kcal/(kg K)
L	1.0 m	R	1.987 kcal/(min K)
V_r	10.0 lt	ρ_m	0.09 kg/lt
E_1	20000 kcal/kmol	U_w	0.20 kcal/(min K)
E_1	50000 kcal/kmol	C_{pmj}	0.8 kcal/(min K)
k_{10}	$5.0 \times 10^{12} \text{ min}^{-1}$	V_j	366 K
k_{20}	$5.0 \times 10^2 \text{ min}^{-1}$	ρ_{mj}	0.10 kg/lt
H_{r1}	0.5480 kcal/kmol	C_{A0}	4 mol/lt
H_{r2}	0.9860 kcal/kmol	C_{B0}	0 mol/lt
T_{r0}	320 K	T_{j0}	375 K

To evaluate the performance of the inferential model, four numerical criteria were used: R^2 and the root mean square error (RMSE). These were calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (23)$$

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \text{mean}(Y))^2}, \quad (24)$$

where y_t are the measured values, $\hat{y}_t$ are the corresponding predicted values by the MSPLS model and n is the number of samples.

A NLPLS model is fitted to the fault-free training dataset, and the goodness of fit is shown in Figure 8.

Figure 7: Plug Flow Reactor: Dynamic input-output data used for training and testing the model with noise SNR=30 (solid black line: noise-free data; blue dots: noisy data).

This figure represents the scatter plot of observed testing data versus predicted values of the constructed NLPLS model. Furthermore, the R^2 values of the constructed NLPLS model is above 0.9, which demonstrates that this modeling method can be used for modeling nonlinear processes effectively. The R^2 was 0.95. The constructed NLPLS model provides a good predictive quality, with the $R^2 = 0.95$ and the low RMSE of 0.0192. Thus, it can be concluded that the constructed model describe well the data.

Figure 8: Scatter plots of predicted and observed testing data.

In general, we have to obtain the model first and then perform fault detection procedures accordingly. Before applying the NLPLS-HD technique for fault detection purpose, we need to check whether the residuals of response variables follows Gaussian distributions to ensure that the data are well represented using a linear

NLPLS model. Figure 9(a) shows that the output residuals obtained from NLPLS model is normal, whereas Figure 9(b) depicts the scatter plots of residuals versus predicted values. The latter shows a random pattern (an indication of mild heteroscedasticity) indicating a good fit for the selected NLPLS model. The residuals histogram shown in Figure 9(c) confirms that the normality assumption appears to be a reasonable one. On the other hand, Figure 9(d), which shows plots of the output residuals for the constructed NLPLS, show a random pattern and small variations around the zero mean, indicating a support to the constructed NLPLS model. The autocorrelation function (ACF) of the residuals depicted in Figure 10 show that the residual are approximately uncorrelated. As the residuals are normally distributed with mean zero and uncorrelated, it can be deduced that the model describes the data well.

Figure 9: Residual plots for the constructed NLPLS model: Normal probability plot(a), predicted values versus residuals for constructed NLPLS model (b), histogram showing the normality of the residuals (c) and residuals versus order.

5.2. Detection results

After the NLPLS model is identified, it is used to monitor the abnormal events (faults) in the process that may lead the process to depart from its normal state. To quantify the efficiency of the proposed strategy, we use two metrics. The false detection rate (FAR) (i.e. the number of the normal observation that is wrongly judged as fault (false alarms) over the total number of fault-free data) and the miss detection rate (MDR)

Figure 10: ACF of the residuals

(i.e. the number of fault that is wrongly judged as normal observation (missed detection) over the total number of faults). In this section, the anomaly detection abilities of the developed NLPLS-HD monitoring chart for detecting incipient fault in multivariate data is shown and compared to NLPLS-Q chart. To assess the strength of the HD-based monitoring scheme, three case studies involving different types of faults were performed. In the first case study, it is assumed that a sensor measuring the temperature is assumed to be faulty with abrupt faults (case A). In the second case study, an intermittent fault in temperature sensors is considered (case B). And in third case study, the monitored PFR is contaminated by gradual fault (case C).

5.2.1. Case (A) - Abrupt anomaly detection

In the first case study, the detection of an abrupt anomaly in temperature sensor of a plug flow reactor is investigated. To generate the fault-contaminated data, we generate 500 observations of fault-free data and then introduce a sensor bias to the data. Three examples are given here to show the detection performance of the fault detection techniques to detect an abrupt fault. In the first example, for data with SNR=30 an abrupt sensor failure related to temperature measurement of x_5 , with a magnitude of 10% of the total variation in the temperature measurement, occurs between samples 300 and 350. This could represent a sudden sensor offset or miss-calibration. The monitoring results of NLPLS-Q and NLPLS-HD charts are demonstrated in Figure 11(a) and 11(b), respectively. From the figure, all charts can give signals of anomaly because the bias shift in this case is quite large. The performances of the NLPS-HD chart when it is applied to the filtered residuals are presented in Figure 11(c). These results show a slight improvement of the NLPS-HD chart applied to the filtered residuals over unfiltered residuals, where the number of false alarms is minimally reduced from 3.11% to 2.66%. If the SNR in the data is large and faults are quite large, this enhancement is not significant. In that case, filtering is not required.

Figure 11: The time evolution of the Q statistic (a), HD statistics (b) and the HD when it is applied to the filtered residuals (c) for the PFR system with abrupt sensor failure in X_3 between samples 300-350 (Case (A), first example).

In the second example, the testing data with low SNR (i.e. SNR=5) are generated for the purpose of evaluation of NLPLS-Q, NLPLS-HD and filtered NLPLS-HD monitoring performances. The three monitoring charts are shown in Figure 12(a)-(c). Figure 12(a) shows that the NLPLS-Q is capable of recognizing this fault but with the expense of a missed detections (i.e., MDR=30%). Thus, for low SNR values, it becomes more difficult for NLPLS-Q chart to accurately detect this type of fault. On the other hand, the plot in Figure 12(b) clearly show the capability of this proposed NLPLS-HD monitoring chart in detecting this moderate anomaly with few false alarms and missed detection (i.e., FAR=3.55% and MDR=2%). Furthermore, it is shown through this simulation that significant improvement in sensor fault detection can be obtained by using HD-based approach as compared to the use of Q chart which is conventionally used with NLPLS based technique. By comparing the two charts, for the testing data with SNR=5, the NLPLS-HD chart provides acceptable monitoring performance, and NLPLS-Q chart yields a high missing detection. This finding maybe can attributed to the noisy data that can affect significantly the Q-based chart. The NLPLS-HD chart when it is applied to the filtered residuals is presented in Figure 12(c). This figure shows the accuracy of the detection was slightly better in the case of filtered residual compared to the used of non-filtered residuals.

Figure 12: The time evolution of the Q statistic (a), HD statistic (b) and the HD statistic when it is applied to the filtered residuals (c) for the PFR system with abrupt sensor failure in X_5 between samples 300-350 (Case (A), second example).

The detection of small changes is central in maintaining the normal operation of a system by providing an early warning which helps in avoiding serious damage and subsequent economic loss. Incipient anomalies is characterized by a weak signature that requires detection indicators that have high sensitivity to small changes. In the third example, a small abrupt fault in the temperature measurement T_5 with a magnitude of 2% is introduced at the 300th sample and removed at the 350th sample of the testing data. Monitoring results of the conventional NLPLS-Q chart are shown in Figure 13(a). From this example, it can be seen that the NLPLS-Q chart could not detect the mean shift in this case. This is mainly due to the fact the conventional NLPLS-Q chart only take into account the information provided by the present data samples in the decision making process, which makes this chart insufficiently powerful and not sensitive for detecting small changes. The result of NLPLS-HD scheme for the considered bias anomaly, which are displayed in Figure 13(b), clearly indicate that the proposed strategy can recognize this anomaly but yields a high missed detection rate (i.e., MDR=32% and FAR =2.66%). Figure 13(c) shows the results of the NLPLS-HD chart when it is applied on the filtered residuals with multiscale filtering. From this figure, it can be seen that the use of filtered residuals via multiscale filter significantly reducing the missed detection rate form 32% to MDR=2%. This finding can be attributed to the powerful of multiscale filtering that can provide a good separation between the deterministic and stochastic parts of a measured signal. This case study clearly testify the superiority of the NLPLS-HD over the conventional NLPLS-Q method. On the other hand, NLPLS-HD method can provide good monitoring performance by using the filtered residuals. In conclusion,

this case study demonstrates that the conventional NLPLS-Q chart cannot be effective in detecting both small and large shifts, while the performance of the proposed NLPLS-HD fault detection technique is quite robust to shifts size.

Figure 13: The time evolution of the Q statistic (a), HD statistic (b) and the HD statistic when it is applied to the filtered residuals (c) for the PFR system with abrupt sensor failure in X_5 between samples 300-350 (Case (A), third example).

5.2.2. Case (B) - Intermittent anomaly detection

The aim of the second case study was to assess the potential of the proposed NLPLS-based HD method to detect intermittent faults. Towards this end, a small bias level, which is 10% of the total variation in T_5 , is injected between samples between intervals [200, 250] and a bias of 15% is introduced between sample interval [400,450]. Figure 14(a)-(b) compares the detection results obtained by NLPLS-Q and NLPLS-HD charts. From Figure 14(a), it can be seen that the NLPLS-Q chart is capable of detecting these moderate faults but with some missed detection (i.e., MDR=3%). Indeed, the NLPLS-Q chart is sensitive to atypical observations having abnormal high or moderate magnitude as the case of this example. Figure 14(b) shows that although the NLPLS-HD chart correctly detected this intermittent faults with few false alarm (i.e., FAR=2.25%). The same NLPLS-HD chart when it constructed using filtered residuals is shown in Figure 14(c). All charts can detect this simulated fault because the mean shift in this case is quite large. It can be seen that applying the NLPLS-HD chart to filtered residuals, it is not very helpful in the case of large fault.

In the second example, we introduce into testing data a bias of amplitude 2% of the total variation in testing data between samples 200 and 250, a bias of 3% from sample 300 to 350, and a bias of 4% is introduced between samples 400 and 450. Figure 15(a)-(c) shows the monitoring results of the NLPLS-Q,

Figure 14: The time evolution of the NLPLS-Q statistic (a), NLPLS-HD chart when it is applied to unfiltered residuals (b), and NLPLS-HD chart when it is applied to the filtered residuals (c) in the presence of intermittent anomalies in X_5 (Case (B), first example).

NLPLS-HD, NLPLS-HD based on filtered residuals techniques, respectively. From Figure 15(a), it can be seen that NLPLS-Q chart has no power to detect this small faults due to its decision based only on the last observation of the data and ignoring the information enclosed in a past data. The NLPLS-Q chart is less sensitive to small or incipient faults, but is able to detect moderate to large shifts in the mean of process variables. In Figure 15b we can see that detection performance is much enhanced buy using the NLPLS-HD chart. On the other hand, NLPLS-HD chart based on filtered residuals gives a much higher detection rate relative to that of the NLPLS-HD method (see Figure 15). Therefore, the NLPLS-HD method based on filtered residuals is more suitable to detect incipient faults. Of course, in this case, the NLPLS-Q chart gives poor results because it is designed for detecting relatively large mean shifts while the true mean shift in this case is quite small. In such example, the NLPLS-HD chart resulted in FAR=3.71% and MDR=5.34%, however NLPLS-HD chart with filtered residuals correctly detected this bias faults without false alarm (i.e., FAR=5.14%). Beside the high sensitivity of HD-based chart with respect to small changes (whose amplitude is less than 10% of the signal magnitude), it has also demonstrated its ability to reflect the fault magnitude as illustrated in Figure 15.

5.2.3. Case (C) - Drift anomaly detection

Early and accurate fault detection is one of the most important challenges during the operation of modern day process systems. The aim of the third case study is to assess the potential of the proposed NLPLS-HD anomaly detection methods to detect incipient or gradual anomalies. To do so, a slow increase in the input

Figure 15: The time evolution of the NLPLS-Q statistic (a), NLPLS-HD chart when it is applied to unfiltered residuals (b), and NLPLS-HD chart when it is applied to the filtered residuals (c) in the presence of intermittent anomalies in X_5 (Case (B), second example).

variable T_5 with a slope of 0.01 was added to the test data starting at sample number 300 of the simulated testing data. The NLPLS-Q chart is shown in Figure 16(a), from which it can be seen that the Q statistic overpass the control limit after the occurrence of anomaly with the first signal at the 372nd time point. On the other hand, the NLPLS-HD statistic gradually increased as the anomaly slowly developed and began to violate the threshold value when the size of the anomaly became sufficiently large that it was detected by this model (see Figure 16(b)). From this figure, it can be seen that the first signal of anomaly is given by the NLPLS-HD chart at sample number 328th. From the Figure 16(a) and Figure 16(b), it is clear that the proposed NLPLS-HD chart gives very early signals of fault compared to the conventional NLPLS-Q chart. The same NLPLS-HD chart when it is constructed from the filtered residual obtained from the NLPLS model is shown in Figure 16(c), and the first signal of anomaly is given by this chart at sample number 309th. Therefore, the NLPLS-HD chart with filtered residuals performs the best in this case and the NLPLS-Q chart performs the worst. Furthermore, in such cases, we have shown that NLPLS-HD is more sensitive to small fault than the conventional NLPLS-Q method. It is worth noting that the combined NLPLS-HD chart with multiscale filtering is often more sensitive to the faults than the NLPLS-HD when it is applied to unfiltered residuals. Of course, the proposed HD-based monitoring method is more reliable for detecting incipient faults.

As a summary, the developed NLPLS-HD monitoring chart showed satisfactory performance compared with the conventional NLPLS-Q method through their application to monitor a plug flow reactor. The

Figure 16: The time evolution of the NLPLS-Q statistic (a), NLPLS-HD chart when it is applied to unfiltered residuals (b), and NLPLS-HD chart when it is applied to the filtered residuals (c) on a semi-logarithmic scale in the presence of a drift sensor fault in X_5 (Case (C)).

results of three simulated case studies show an even clearer advantage for NLPLS-HD method in the presence of smaller faults. It is shown through simulation that significant improvement in sensor fault detection can be obtained by using HD-based fault detection approach as compared to the use of Q chart which is conventionally used with NLPLS based technique. These results are encouraging especially when it is of interest to detect incipient faults. Furthermore, this study shows that a better result is obtained when HD is applied to filtered residual than to non-filtered residuals. If the SNR in the data is large, this enhancement is not significant. Multiscale filtering reduces false alarms and helps detecting minor sensor biases. Results indicate that the proposed chart is a very promising HD-based method because HD-based charts are, in practice, designed to detect small shifts in process parameters.

6. Conclusion

Fault detection can effectively help the operators and engineers in control and automation, and therefore significantly reduce the risk of safety problems or losses in profitability. This paper proposes a new approach to process monitoring that uses NLPLS models to capture nonlinear features in multivariate data and Hellinger distance metric as anomaly indicator to achieve enhanced multivariate statistical process monitoring. Motivated by the sensitivity of HD-based detection to small changes in the process mean, we developed an efficient chart for detecting incipient anomalies in nonlinear multivariate processes. The HD-based anomaly detection using NLPLS is conceptually more straightforward and more sensitive for the detection of incip-

ient anomalies. The effectiveness and superiority of NLPLS-HD chart chart was demonstrated with data from a simulated plug-flow reactor. We showed that satisfactory detection results were obtained using the proposed method, especially for detecting small anomalies. The new scheme, NLPLS-HD, was verified to be more powerful than the conventional NLPLS-Q method. This study shows also that a better result is obtained when HD is applied to filtered residual than to non-filtered residuals, in particular for small faults. Of course, results indicate that the proposed chart is a very promising HD-based method because HD-based charts are, in practice, designed to detect small shifts in process parameters.

This work further can be extended to to develop adaptive multiscale fault detection methods that extend the advantages of the HD-based fault detection methods to handle processes with varying conditions.

Acknowledgement

This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No: OSR-2015-CRG4-2582.

References

- [1] S. Ding, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.
- [2] F. Harrou, L. Fillatre, and I. Nikiforov, "Anomaly detection/detectability for a linear model with a bounded nuisance parameter," *Annual Reviews in Control*, vol. 38, no. 1, pp. 32–44, 2014.
- [3] F. Khan and S. Abbasi, "Major accidents in process industries and an analysis of causes and consequences," *Journal of Loss Prevention in the process Industries*, vol. 12, no. 5, pp. 361–378, 1999.
- [4] R. Isermann, "Model-based fault-detection and diagnosis-status and applications," *Annual Reviews in control*, vol. 29, no. 1, pp. 71–85, 2005.
- [5] F. Harrou, M. Madakyaru, Y. Sun, and S. Khadraoui, "Improved detection of incipient anomalies via multivariate memory monitoring charts: Application to an air flow heating system," *Applied Thermal Engineering*, DOI: 10.1016/j.applthermaleng.2016.08.047, vol. 109, pp. 65–74, 2016.
- [6] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.
- [7] S. Yin, S. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 11, pp. 6418–6428, 2014.
- [8] F. Harrou, F. Kadri, S. Khadraoui, and Y. Sun, "Ozone measurements monitoring using data-based approach," *Process Safety and Environmental Protection*, vol. 100, pp. 220–231, 2016.
- [9] F. Harrou, Y. Sun, and S. Khadraoui, "Amalgamation of anomaly-detection indices for enhanced process monitoring," *Journal of Loss Prevention in the Process Industries*, vol. 40, pp. 365–377, 2016.

- [10] Z. Ge, Z. Song, and F. Gao, "Review of recent research on data-based process monitoring," *Industrial & Engineering Chemistry Research*, vol. 52, no. 10, pp. 3543–3562, 2013.
- [11] F. Harrou, M. Nounou, H. Nounou, and M. Madakyaru, "PLS-based EWMA fault detection strategy for process monitoring," *Journal of Loss Prevention in the Process Industries*, vol. 36, pp. 108–119, 2015.
- [12] T. Kourti, "Process analysis and abnormal situation detection: from theory to practice," *IEEE control systems*, vol. 22, no. 5, pp. 10–25, 2002.
- [13] R. Rosipal and L. Trejo, "Kernel partial least squares regression in reproducing kernel hilbert space," *Journal of machine learning research*, vol. 2, no. Dec, pp. 97–123, 2001.
- [14] S. Wold, N. K. Wold, and B. Skagerberg, "Nonlinear pls modeling," *Chemometrics and Intelligent laboratory Systems*, vol. 7, pp. 53–65, 1989.
- [15] S. Wold, "Nonlinear partial least squares modelling II. Spline inner relation," *Chemometrics and Intelligent Laboratory Systems*, vol. 14, no. 1-3, pp. 71–84, 1992.
- [16] —, "Nonlinear partial least squares modeling: li spline inner relation," *Chem. Int. Lab. Sys.*, vol. 14, pp. 71–84, 1992.
- [17] Z. Tian, B. Gu, L. Yang, and Y. Lu, "Hybrid ANN-PLS approach to scroll compressor thermodynamic performance prediction," *Applied Thermal Engineering*, vol. 77, pp. 113–120, 2015.
- [18] D. Lee, M. Lee, S. Woo, Y. Kim, and J. Park, "Nonlinear dynamic partial least squares modeling of a full scale biological wastewater treatment plant," *Process Biochemistry*, vol. 41, pp. 2050–2057, 1992.
- [19] F. Chu, F. Wang, X. Wang, and S. Zhang, "Performance modeling of centrifugal compressor using kernel partial least squares," *Applied Thermal Engineering*, vol. 44, pp. 90–99, 2012.
- [20] I. Araby, Abdel-Rahman, and G. J. Lim, "A nonlinear partial least squares algorithm using quadratic fuzzy system," *J. of Chemometrics*, vol. 23, pp. 530–537, 2009.
- [21] Y. H. Bang, C. K. Yoo, and I.-B. Lee, "Nonlinear pls modeling with fuzzy inference system," *Chemometrics and Intelligent laboratory Systems*, vol. 64, no. 2, pp. 137–155, 2003.
- [22] M. Madakyaru, M. Nounou, and H. Nounou, "Nonlinear latent variable regression," in *IEEE Symposium on Computational Intelligence in Control and Automation (CICA)*. IEEE, 2013, pp. 81–88.
- [23] R. Ganesan, T. Das, and V. Venkataraman, "Wavelet-based multiscale statistical process monitoring: A literature review," *IIE transactions*, vol. 36, no. 9, pp. 787–806, 2004.
- [24] Y. Hiroyuki, Y. B. Hideki, F. C. E. O. Hiromu, and F. Hideki, "Canonical correlation analysis for multivariate regression and its application to metabolic fingerprinting," *Biochemical Engineering Journal*, vol. 40, pp. 199–204, 2008.
- [25] M. Stone and R. Brooks, "Continuum regression: Cross validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal component regression," *Journal of the Royal Statistical Society, Ser. B*, vol. 52, pp. 237–269, 1990.
- [26] S. J. Qin and T. J. McAvoy, "Nonlinear pls modeling using neural networks," *Comput. Chem. Eng.*, vol. 16, no. 4, pp. 379–391, 1992.
- [27] S. Qin, "Statistical process monitoring: Basics and beyond," *Journal of Chemometrics*, vol. 17, no. 8/9, pp. 480–502, 2003.
- [28] D. C. Montgomery, "Introduction to statistical quality control," *John Wiley& Sons, New York*, 2005.
- [29] F. Harrou, M. Nounou, H. Nounou, and M. Madakyaru, "Statistical fault detection using PCA-based GLR hypothesis testing," *Journal of loss prevention in the process industries*, vol. 26, no. 1, pp. 129–139, 2013.
- [30] M. Basseville, "Divergence measures for statistical data processing—An annotated bibliography," *Signal Processing*, vol. 93,

- no. 4, pp. 621–633, 2013.
- [31] G. Ditzler and R. Polikar, “Hellinger distance based drift detection for nonstationary environments,” in *IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE)*. IEEE, 2011, pp. 41–48.
 - [32] C. Li, B. Huang, and F. Qian, “Hellinger distance based probability distribution approach to performance monitoring of nonlinear control systems,” *Chinese Journal of Chemical Engineering*, vol. 23, no. 12, pp. 1945–1950, 2015.
 - [33] I. Csiszár and P. Shields, *Information theory and statistics: A tutorial*. Now Publishers Inc, 2004.
 - [34] D. González and J. Hayet, “Fast human detection in rgb-d images with progressive SVM-classification,” in *Pacific-Rim Symposium on Image and Video Technology*. Springer, 2013, pp. 337–348.
 - [35] A. Korostelev and A. Tsybakov, *Minimax theory of image reconstruction*. Springer Science & Business Media, 2012, vol. 82.
 - [36] D. Cieslak and N. Chawla, “A framework for monitoring classifiers’ performance: when and why failure occurs?” *Knowledge and Information Systems*, vol. 18, no. 1, pp. 83–108, 2009.
 - [37] V. González-Castro, R. Alaiz-Rodríguez, and E. Alegre, “Class distribution estimation based on the hellinger distance,” *Information Sciences*, vol. 218, pp. 146–164, 2013.
 - [38] L. Aggoune, Y. Chetouani, and H. Radjeai, “Change detection in a distillation column using non-linear auto-regressive moving average with exogenous input model and hellinger distance,” *IET Science, Measurement & Technology*, vol. 10, no. 1, pp. 10–17, 2016.
 - [39] J. Tajer, A. Makke, O. Salem, and A. Mehaoua, “A comparison between divergence measures for network anomaly detection,” in *7th International Conference on Network and Service Management*. IEEE, 2011, pp. 1–5.
 - [40] C. R. Rao, “A review of canonical coordinates and an alternative to correspondence analysis using hellinger distance,” vol. 19, pp. 23–63, 1995.
 - [41] K. Yamanishi, J.-I. Takeuchi, G. Williams, and P. Milne, “On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms,” *Data Mining and Knowledge Discovery*, vol. 8, no. 3, pp. 275–300, 2004.
 - [42] L. Pardo, *Statistical inference based on divergence measures*. CRC Press, 2005.
 - [43] D. Romano and M. Kinnaert, “Robust fault detection and isolation based on the kullback divergence,” *IFAC Proceedings Volumes*, vol. 39, no. 13, pp. 426–431, 2006.
 - [44] S. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
 - [45] B. Bakshi, “Multiscale pca with application to multivariate statistical process monitoring,” *AIChE Journal*, vol. 44, pp. 1596–1610, 1998.
 - [46] R. Gao and R. Yan, *Wavelets: Theory and applications for manufacturing*. Springer Science & Business Media, 2010.
 - [47] I. Daubechies, “Orthonormal bases of compactly supported wavelets,” *Communications on pure and applied mathematics*, vol. 41, no. 7, pp. 909–996, 1988.
 - [48] M. Sheriff, F. Harrou, and M. Nounou, “Univariate process monitoring using multiscale Shewhart charts,” in *International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2014, pp. 435–440.
 - [49] D. Donoho, I. Daubechies, and V. Pierre, “Wavelets on the interval and fast wavelet transforms,” *Appli. Comput. Harmonic Analysis*, vol. 1, no. 7, pp. 64–81, 1993.
 - [50] D. L. Donoho, I. M. Johnstone, G. Kerkycharian, and D. Picard, “Wavelet shrinkage: Asymptotia?” *J. Roy. Stat. Soc. B*, vol. 57, no. 301, 1995.

- [51] B. Bakshi, "Multiscale analysis and modeling using wavelets," *Journal of Chemometrics*, vol. 13, no. 3, pp. 415–434, 1999.
- [52] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.
- [53] F. Harrou, M. Nounou, and H. Nounou, "Enhanced monitoring using PCA-based GLR fault detection and multiscale filtering," in *IEEE Symposium on Computational Intelligence in Control and Automation (CICA)*. IEEE, 2013, pp. 1–8.
- [54] M. Madakyaru, M. Nounou, and H. Nounou, "Integrated multiscale latent variable regression and application to distillation columns," *Modelling and Simulation in Engineering*, vol. 2013, p. 3, 2013.
- [55] C. Wang, L. Chen, S. Xia, and F. Sun, "Maximum production rate optimization for sulphuric acid decomposition process in tubular plug-flow reactor," *Energy*, vol. 99, pp. 152–158, 2016.
- [56] H. Shang, J. F. Fraser, and M. Guay, "Characteristics-based model predictive control of distributed parameter system," *Americal Control Conference PRoc., Anchorage, Alaska*, pp. 4383–4388, May 8-10, 2002.