Journal of Process Control 33 (2015) 60–76

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Reconstruction-based contribution approaches for improved fault diagnosis using principal component analysis

Baligh Mnassri*, El Mostafa El Adel, Mustapha Ouladsine

Aix Marseille Université, CNRS, ENSAM, Université de Toulon, LSIS UMR 7296, 13397, Marseille, France

ARTICLE INFO

Article history:
Received 12 October 2014
Received in revised form 17 April 2015
Accepted 3 June 2015
Available online 24 June 2015

Keywords:
Principal component analysis
Fault detection and diagnosis
Reconstruction-based contribution
Continuous stirred tank reactor (CSTR)

ABSTRACT

This paper provides two new proposed data-based fault diagnosis approaches using the principal component analysis (PCA). Since faults are really complex and may be in multidimensional directions, the first proposal is a generalized *RBC* method. The theoretical diagnosability analysis using this one guarantees correct fault diagnosis in the case of complex faults that have large magnitudes. Nevertheless, the common assumption resulting from the use of the contribution methods in general is that variables with largest contributions to the fault detection indices are more likely to be the root causes of the fault occurrence. To handle the more complex faults and remedy the defective of the *RBC* method, the second proposal presents an alternative approach called *RBC* ratio (*RBCR*). A theoretical diagnosability analysis testifies to its strong performance in identifying the detectable faults. Indeed, the isolation of fault is guaranteed once its magnitude has satisfied a sufficient condition for isolability. These proposed approaches have been successfully applied to a numerical example as well as the CSTR process.

1. Introduction

Fault detection and diagnosis (FDD) approaches are necessary to ensure safe and efficient operation of many industrial processes. In the literature, three main review parts of process FDD techniques have been distinguished and discussed such as quantitative model-based methods, qualitative models and search strategies, and process history-based methods [1–3]. By focusing on the latter topic, multivariate statistical process monitoring (MSPM) and its application for fault detection and isolation based on the historical process data have received considerable attention in research area as well as in industrial applications. Indeed, the MSPM includes modelling techniques which can handle and exploit the process data. Its successful use is usually attributed to the high accuracy ability of the developed process models in detecting any deviation from the normal operating conditions. Until a few decades ago, such data exploitation was neither easy nor efficient due to many problems particularly related to the nature of data. These ones are often enormous, highly correlated and non-causal in nature. The information contained in any process variable is often very small due to the low signal-to-noise ratios. The measurements of many process variables are often missing. In this context, empirical modelling and projection methods are the best to deal effectively with all these difficulties and adequately handling the process data. The literature has provided several academia research efforts interesting on developed models by using latent variable methods such as principal component analysis (PCA) and projection to latent structures or partial least squares (PLS). A main benefit of these two approaches is their capacities to handle multivariate systems that consist of large number of correlated variables. They address all the mentioned problems in a straightforward way and provide quite presentable and interpretable analysis tools [4,5].

PCA is a well-used method in multivariate statistical analysis. It has had several extensions, as the Multiway PCA and the robust PCA [5,6], which are widely applied to many industrial process monitoring and analysis applications [4,7]. PCA-based techniques are typically conceived through a partitioning of the data space in a principal component subspace (PCS) and a residual subspace (RS). Indeed, an optimal partition is the key in determining the PCA-model that adequately represents the process. The partition procedure requires the estimation of the number of the more significant principal components (PCs) in a multivariate data matrix. Ideally, this number should be equal to the correct rank of the covariance (or correlation) matrix of the historical normal data. However, in practice and although the wide-use of PCA, its determination is rather subjective and not unique because real data often contain artifacts such as baseline problems, sensor outputs that are

* Corresponding author. Tel.: +33 491 056 041; fax: +33 491 056 033.
E-mail addresses: baligh.mnassri@lsis.org (B. Mnassri), mostafa.eladel@lsis.org (E.M. El Adel), mustapha.ouladsine@lsis.org (M. Ouladsine).

http://dx.doi.org/10.1016/j.jprocont.2015.06.004
0959-1524/

usually disturbed by different types of noise, errors generated by data preprocessing and so on [8]. In this context, the authors of [9] have reviewed and compared with their variance of the reconstruction error (VRE) criterion a plethora of other criteria that are available in the signal processing and chemometrics literature. To enhance the model selection accuracy, some extensions of the VRE criterion have led to the proposition of other new ones [10–12]. The aim is to overcome the disadvantage of the classical VRE criterion face the presence of independent and/or very weakly correlated process variables. The first two proposed criteria [10,11], which both depend on a confidence level, are considered as heuristic criteria. Indeed, this dependence on the confidence level often causes difficulties that make the results and decisions uncertain. We have extended our research in this framework by proposing a new more efficient and consistent criterion [12].

Several fault detection indices which are usually derived from data projection onto the mentioned subspaces have been widely used in PCA-based process monitoring. The Hotelling's $T2$ statistic captures the amount of faulty variations from the PCS, while the squared prediction error (*SPE*) and Hawkins's T_H^2 or *SWE* indices measure those from the RS. For an easier process monitoring, it has been advocated to combine the weighting of *SPE* and $T2$ statistics into one combined index noted φ [13]. All the cited detection indices are analyzed and unified since they are characterized by a unified quadratic form [7]. This characteristic has led to the emergence of several studies that generalize and analyze the use of the unified form of the quadratic indices for fault detection, isolation and diagnosis concepts [14–16]. Process monitoring operation refers to thresholds to ensure if a process is capable of producing normally. The process thresholds, also known as natural process limits, are horizontal lines drawn on a statistical process control chart. They are used to detect signals in process data that indicate if the process is not in control and, therefore, not operating predictably. The abnormal process behaviour usually involves the existence of faults and its detection is ensured by comparing the values of the detection indices to their corresponding thresholds. It is noted that these fault detection indices are used to only detect the abnormal process behaviour without explicitly provide any diagnosis information.

Although many works have focused on the fault detection using data-based PCA-models, unfortunately far fewer fault diagnosis methods are discussed in the literature. The primary and the more popular diagnosis tool is the contribution analysis, which does not require specific fault information to diagnose the fault. Its theoretical development has had several manners providing, therefore, many various contribution methods [4,5,7,14,17–19]. Most of the contribution-based diagnosis approaches was well summarized, analyzed and generalized by Alcala and Qin [15]. Despite their different forms, the common assumption behind the use of all the contribution methods is that faulty variables tend to provide the largest contributions to the fault detection indices. Furthermore, contribution analysis investigates individual variables one by one. It can be ineffective in isolating several faulty variables which jointly contribute to the fault occurrence and can result a misleading diagnosis due to the correlation effect between the process variables.

On the other hand, this correlation has been the backbone of a more efficient fault diagnosis approach called the fault reconstruction. Its principle is founded on the removing of the fault effect onto the detection indices by reconstructing the faulty variables through different PCA-models. The success in using PCA for fault process monitoring and reconstruction-based fault diagnosis was enriched by the development of two fundamental concepts which are the fault detectability and the fault isolability [7,13,20,21]. In PCA framework, fault detectability involves the capability of the PCA-model to detect the presence of a fault based on a given detection index. Moreover, fault isolability makes a fault capable to be distinguished from one another by using the PCA-model and the fault behaviour. More recently, these concepts have been analyzed and generalized considering a common quadratic-form index [16].

The reconstruction-based contribution (*RBC*) is a recent fault diagnosis method that combines both the contribution analysis and the reconstruction-based identification in order to enhance the diagnosis results [14,15]. This method guarantees correct diagnosis of only the unidimensional faults that have large magnitudes. By extending the application of the *RBC* to the case of known fault direction, a generalized *RBC* method which has the ability to use the fault information when available, is proposed to diagnose output-relevant faults using total projection to latent structures (T-PLS) models [22]. To take advantage of the simplicity of contribution plots and the rigor of multivariate fault reconstruction methods, the *RBC* method has been extended to reconstruction-based block contribution as well as reconstruction-based variable contribution in order to diagnose the faulty block and, therefore, the faulty variables, respectively [23]. To diagnose faults in nonlinear processes, the *RBC* was extended to be used with kernel PCA [24]. It works like the standard contribution plots in linear PCA, which do not require a prior knowledge about the fault direction. Indeed, the *RBC* is computed using the estimated fault magnitudes that usually contain cross-term from other fault directions. According to [25], orthogonal directions projected onto lower-rank subspace lose orthogonality and, therefore, introduce smearing caused by cross-direction effects. Due to the fault smearing problem, the classical *RBC* can provide misdiagnosis particularly for the complex fault cases. To overcome the cited problem, the authors of [25] have proposed an improved method called weighted *RBC* (*WRBC*).

In the *RBC* context, the fault smearing effect seems an evident problem when diagnosing multidimensional or complex faults whereas the fault reconstruction applied in the *RBC* is achieved by reconstructing the variables one by one. In this paper, we propose a new contribution approach that addresses the diagnosis of multidimensional faults based on PCA-models. The proposed generalized or multidimensional *RBC* guarantees correct diagnosis results when the actual faults have large magnitudes. In the intention to deal with the more complex faults, it has advocated to compare the *RBC* to a given threshold. Unfortunately, we have proved that such a proposal cannot ensure correct fault diagnosis due to the fault smearing in variables. Our second aim is to find an alternative approach for the *RBC*, that guarantees the identification and isolation of the detectable complex faults. The isolation of these fault categories usually needs valid and efficient thresholds. Therefore, we have proposed a second new approach called *RBC* ratio (*RBCR*). Its diagnosability analysis ensures the isolability of faults by satisfying an isolability sufficient condition.

The following section of this paper exposes the principle of process modelling using PCA as well as a brief review of PCA-based fault detection and detectability. An overview of the reconstruction-based fault isolation and isolability is presented in Section 3. It also reviews the principle of the classical *RBC* approach and its diagnosability analysis. Section 4 provides the limitation proof on the use of this approach to diagnose the complex faults. It presents the new proposed multidimensional *RBC* and analyses its fault diagnosability in the case of complex faults that have large magnitudes. Further, a fundamental study on the use of the *RBC* threshold is discussed. This section also provides a new alternative contribution method that addresses the more complex fault diagnosis, as well as its fault diagnosability analysis. Section 5 presents case studies to illustrate and validate the proposed fundamental issues, followed by concluding remarks in Section 6.