Control Engineering Practice 46 (2016) 51–58

Contents lists available at ScienceDirect

Control Engineering Practice

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

Canonical correlation analysis-based fault detection methods with application to alumina evaporation process ☆

Zhiwen Chen [a,*], Steven X. Ding [a], Kai Zhang [a], Zhebin Li [b], Zhikun Hu [b]

[a] Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Bismarckstrasse 81 BB, 47057 Duisburg, Germany
[b] School of Physics and Electronics, University of Central South, 410083 Changsha, China

ARTICLE INFO

Article history:
Received 26 March 2015
Received in revised form
8 October 2015
Accepted 12 October 2015
Available online 23 October 2015

Keywords:
Canonical correlation analysis
Residual generation
Fault detection
Alumina evaporation process

ABSTRACT

In this paper, canonical correlation analysis (CCA)-based fault detection methods are proposed for both static and dynamic processes. Different from the well-established process monitoring and fault diagnosis systems based on multivariate analysis techniques like principal component analysis and partial least squares, the core of the proposed methods is to build residual signals by means of the CCA technique for the fault detection purpose. The proposed methods are applied to an alumina evaporation process, and the achieved results show that both methods are applicable for fault detection, while the dynamic one delivers better detection performance.

1. Introduction

The proper functioning of industrial processes has a profound impact on production, product quality, and safety. Efficient detection of faults is essential in avoiding performance degradation and damage to equipment. For these purposes, process monitoring and fault diagnosis (PM-FD) methods have been widely studied and implemented in industry. The existing PM-FD methods are subdivided into model-based and data-driven methods. The model-based methods have received considerable attention (Basseville & Nikiforov, 1993; Gertler, 1998; Isermann, 2006; Ding, 2013) and found a large number of successful applications in automotive, aerospace systems, etc., where the first-principle models or mathematical models are available (Clark, 1978; Gertler et al., 1995; Ding, Fennel, & Ding, 2004). For large-scale processes like alumina evaporation process (AEP) (Tang, Yang, & Gui, 2011), accurate physical models are often unavailable. On the other hand, techniques for routine data collecting, storing and processing have been significantly improved in past years. Thus, these have given rise to the extensive development of data-driven methods (Ge & Song, 2008; Peng, Zhang, Li, & Zhou, 2013; Hu, Chen, Gui, & Jiang, 2014; Chen et al., 2014). Common data-driven methods are based on multivariate analysis (MVA) techniques, such as principal component analysis (PCA) and partial least square (PLS), which have attracted more and more attention from academia and industry (Qin, 2012). In industry, there are numerous successful applications of MVA-based methods, e.g. a PCA-based method for semiconductor manufacturing fault detection (Wise & Gallagher, 1996). In the aluminium smelting industry, a PCA-based method and its variants have been applied to the performance analysis of a line of operating cells and to the detection of anode spikes and anode effects (Tessier, Duchesne, Tarcy, Gauthier, & Dufour, 2008; Majid et al., 2011). In Ding, Yin, Peng, Hao, and Shen (2013), modified PLS has been applied to the prediction and diagnosis of key performance variables of an industrial hot strip mill. The reader is referred to (Russell, Chiang, & Braatz, 2000; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003; Kano & Nakagawa, 2008; Qin, 2012; Yin, Ding, Haghani, Hao, & Zhang, 2012; Ding, 2014) for a comprehensive literature study.

MVA-based PM-FD methods consist of two major steps: off-line training and on-line monitoring. The differences between PCA- and PLS-based methods are that the PCA-based methods only consider the process measurements in both steps and monitor changes in the condition of the process, sensors and actuators, while PLS-based methods are applied to process variables and output variables (quality variables) or key performance indicators (KPI), which are on-line unmeasurable or measurable only with a large time delay. In off-line training, quality data are used to guide the decomposition of the process data and to extract latent variables that are mostly relevant to the product quality. In on-line monitoring, only process variables are available and used to detect

☆Part of the work was presented at the 19th IFAC World Congress held in Cape Town, August 2014.

* Corresponding author.
E-mail address: zhiwen.chen@uni-due.de (Z. Chen).

http://dx.doi.org/10.1016/j.conengprac.2015.10.006
0967-0661/

Table 1
Tabular comparison between CCA-, PCA- and PLS-based monitoring methods.

Method	Assumption on data	Variables	Detection purpose
PCA-based	Multivariate normal distribution	Process measurements (sensors)	Changes in sensor and process
PLS-based	Same as PCA, clear input–output relationship	Both input and output, output is on-line unmeasurable	Changes related with quality variable
CCA-based	Same as PCA, clear input–output relationship	Both input and output, both are on-line measurable	Changes in input, output and process

faults that are mostly related to the product quality variables or KPIs (Zhang, Hao, Chen, Ding, & Peng, 2015).

When input–output relationship explicitly exists and the two blocks of input and output data are on-line measurable, canonical correlation analysis (CCA) technique (Hotelling, 1936) is an efficient tool to design a fault detection (FD) system. However, as a representative multivariate analysis technique, CCA has been rarely employed for fault detection. It is the first objective of this work to deal with fault detection issues for linear static and dynamic processes using CCA technique. In the static case, the output is assumed to be affected only by the current measurements. In the dynamic case, process input and output data in a time interval are applied for the detection purpose. *CCA-based FD methods can be viewed as an extension of PCA-based or PLS-based methods for detecting faults in a process with input and output data.* Table 1 presents a tabular comparison between CCA-, PCA-, and PLS-based methods to clarify our objective.

FD based on residual generation is the state of the art in the model-based fault diagnosis framework (Chow & Willsky, 1984). In this paper, a canonical correlation-based residual generation is first realized by the CCA technique in the data-driven fashion. It is then applied to FD in static processes. In order to address FD in dynamic processes, the proposed static method is further extended to the dynamic version, which is similar with the well-established dynamic PCA (DPCA)- and dynamic PLS (DPLS)-based techniques. The last objective of our work is to apply the two proposed FD methods to the monitoring of an alumina evaporation process.

The rest of this paper is organized as follows. In Section 2, we briefly introduce the basic ideas and formulate the essential form of canonical correlation-based residual generation. Subsequently, a CCA-based FD scheme is proposed in Section 3. Section 4 deals with the dynamic CCA (DCCA). Noting that the proposed method for dynamic case is similar with canonical variate analysis (CVA)-based methods for system identification (Larimore, 1983), we discuss, in Section 5, about the differences between them. Furthermore, possible extensions of the proposed methods are stated. In Section 6, the AEP is first described for the case study. It is followed by the application of both CCA- and DCCA-based methods to the monitoring of the AEP.

Notation: The notation used in this paper is standard. $\mathcal{R}^n$ denotes the n-dimensional Euclidean space, $\mathcal{R}^{n\times m}$ is the set of all $n\times m$ real matrices, and $diag(.,\ldots,.)$ is a square diagonal matrix. $rank(\mathbf{A})$ denotes the rank of matrix $\mathbf{A}$. $\mathbf{A}(:,i)$ represents the i-th column of $\mathbf{A}$. $\mathbf{I}_n$ is an $n\times n$ identity matrix. $\mathbf{x}\sim\mathcal{N}(\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x)$ denotes that $\mathbf{x}$ is a normally distributed random vector with mean $\boldsymbol{\mu}_x$ and covariance $\boldsymbol{\Sigma}_x$. $\chi^2(m)$ stands for the chi-square distribution with m degrees of freedom and $\mathcal{F}(m,n)$ stands for F-distribution with m and n degrees of freedom. Let $prob(\chi^2>\chi^2_{1-\alpha}(m))=\alpha$ represents the probability that $\chi^2>\chi^2_{1-\alpha}(m)$ equals α (significance level) and $prob(\mathcal{F}>\mathcal{F}_{1-\alpha}(m,n))=\alpha$ represents the probability that $\mathcal{F}>\mathcal{F}_{1-\alpha}(m,n)$ equals α.

2. Basic ideas and problem formulation

FD methods based on parity relation have been widely studied (Chow & Willsky, 1984; Ding, 2013). In those methods, residual generation is an essential step. In Gertler (1998), the residual generation is described in the following general form:

$$\mathbf{r}(k)=\mathbf{V}(\phi)\mathbf{u}(k)+\mathbf{W}(\phi)\mathbf{y}(k),$$

where $\mathbf{r}(k)$ is the residual signal at time k, $\mathbf{u}$ and $\mathbf{y}$ are input and output vectors, respectively. $\mathbf{V}(\phi)$ and $\mathbf{W}(\phi)$ are transfer function matrices, and ϕ is the shift operator. Let the nominal system model be $\mathbf{y}(k)=\mathbf{G}(\phi)\mathbf{u}(k)$, then,

$$\mathbf{V}(\phi)\mathbf{u}(k)+\mathbf{W}(\phi)\mathbf{G}(\phi)\mathbf{u}(k)=0\Longrightarrow\mathbf{V}(\phi)=-\mathbf{W}(\phi)\mathbf{G}(\phi)$$

has to be satisfied for all $\mathbf{u}(k)$, in order to achieve a successful residual generation. As a result, a general form of residual generators can be written as

$$\mathbf{r}(k)=\mathbf{W}(\phi)\mathbf{y}(k)-\mathbf{W}(\phi)\mathbf{G}(\phi)\mathbf{u}(k).$$

Thus, the key step for generating the residual signal is to identify $\mathbf{W}(\phi)$ and $\mathbf{G}(\phi)$. Note that $\mathbf{G}(\phi)$ is a transfer function, which is given. In this context, the parity relation-based residual generation is conventionally referred as a model-based approach. Motivated by the facts that a process model is often not available or only achievable at high engineering costs, we address the residual generation problem in the data-driven fashion.

For static processes, CCA technique is a powerful tool to analyze the correlation between process input and output variables. Motivated by the parity relation-based residual generation, in this paper, canonical correlation-based residual signal is defined as follows:

$$\mathbf{r}(k)=\mathbf{L}^T\mathbf{y}(k)-\mathbf{M}^T\mathbf{u}(k),\tag{1}$$

where $\mathbf{r}(k)$ is the residual signal at time k. $\mathbf{L}$ and $\mathbf{M}$ are unknown constant matrices. An extension to dynamic processes by means of DCCA-based method will be addressed in Section 4.

3. CCA-based FD method in the case of static processes

This section is devoted to the development of a CCA-based FD scheme for static processes. We assume that the processes under consideration are described by

$$\boldsymbol{\Psi}_y\mathbf{y}(k)=\boldsymbol{\Psi}_u\mathbf{u}(k)+\mathbf{v}(k),\tag{2}$$

where $\boldsymbol{\Psi}_y\in\mathcal{R}^{m\times m}$ and $\boldsymbol{\Psi}_u\in\mathcal{R}^{m\times l}$ are constant but unknown matrices, $\mathbf{u}\in\mathcal{R}^l$ is the input vector and $\mathbf{y}\in\mathcal{R}^m$ is the output vector, $\mathbf{v}\in\mathcal{R}^m$ is a normally distributed vector with zero mean and unknown constant covariance.

Let $\mathbf{u}_{obs}\in\mathcal{R}^l$ and $\mathbf{y}_{obs}\in\mathcal{R}^m$ be the measured process input and output vectors, respectively. Assume that

$$\mathbf{u}_{obs}\sim\mathcal{N}(\boldsymbol{\mu}_u,\boldsymbol{\Sigma}_u),\ \mathbf{y}_{obs}\sim\mathcal{N}(\boldsymbol{\mu}_y,\boldsymbol{\Sigma}_y),$$

where $\boldsymbol{\mu}_u$, $\boldsymbol{\Sigma}_u$, $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_y$ are unknown but constant. Denote the