Journal of Process Control 24 (2014) 1548–1555

Contents lists available at ScienceDirect

Journal of Process Control

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

Online process operating performance assessment and nonoptimal cause identification for industrial processes

Yan Liu [a], Yuqing Chang [a,b,*], Fuli Wang [a,b]

[a] College of Information Science & Engineering, Northeastern University, Shenyang, Liaoning 110819, China
[b] State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang, Liaoning 110819, China

ARTICLE INFO

Article history:
Received 28 November 2013
Received in revised form 4 June 2014
Accepted 1 August 2014
Available online 30 August 2014

Keywords:
Operating performance
Online assessment
Nonoptimal cause identification
Performance grade
T-PLS
Comprehensive economic benefit (CEB)

ABSTRACT

Although industrial processes are usually operated at the optimal point in the early stage of the production, the operating performance may deteriorate with time due to process disturbances. In order to pursue optimal comprehensive economic benefit (CEB), online process operating performance assessment on optimality has become a key issue. However, a little work has been published in this research area. In this paper, a new online operating performance assessment and nonoptimal cause identification method for industrial process are proposed. The contributions of this paper can be summarized as follows: a novel performance-similarity-based online operating performance assessment method is proposed; total projection to latent structures (T-PLS) is applied to the area of process performance assessment for the first time; the online assessment results include not only the deterministic performance grades, but also the performance grade conversions which were not covered in the existing assessment method; when the assessment result is nonoptimal, a novel automatic nonoptimal cause identification strategy is developed based on variable contributions, which is meaningful for guiding the further production adjustment. Finally, the feasibility and efficiency of the proposed method are illustrated with a case of gold hydrometallurgical process.

1. Introduction

In order to maximize the comprehensive economic benefit (CEB), process operating optimization becomes a crucial issue in industrial process and has attracted great attention from both academia [1–3] and practice [4–6] in recent years. However, it is well known that the process operating performance will deviate from the optimal operating point owing to process disturbances and uncertainties. Therefore, to develop the appropriate evaluation strategies is much crucial for online operating performance assessment on optimality.

As we all know, the purpose of process monitoring [7–11] is to decide whether the operating is normal or fault, but operating performance assessment on optimality [12] is to answer "how optimal is the current operating performance under normal operating condition?". That is to say, through operating performance assessment, managers and operators can grasp the process performance grade, such as optimal and suboptimal, and propose reference suggestions for further operating adjustment and performance improvement.

However, to the best of our knowledge, only a little work has been reported on operating optimality assessment for industrial process. Recently, a probabilistic framework of operating assessment for multimode industrial processes has been proposed [12], where an optimality assessment index is defined and used to evaluate the process operating optimality. However, the assessment index is constructed based on the value of optimized objective function, such as plant cost, profits, and product quality, which usually cannot be obtained online in actual production, especially for the time-consuming and large scale complex industrial processes; furthermore, only limited experience-based qualitative analysis is provided for nonoptimal operating performance and without a general method of quantitative analysis, that is, an automatic nonoptimal cause identification method has not been explored so far.

Generally, the process operating performance on optimality is closely associated with the CEB. If the CEB approaches or reaches the history optimal level, we believe that the process operating performance is optimal. However, CEB is often difficult to be calculated online and usually comes with significant sampling delays. Hence, it is not suitable to evaluate the process operating performance

* Corresponding author at: College of Information Science & Engineering, Northeastern University, 3 Lane 11, Wenhua Road, Heping District, Shenyang, Liaoning 110819, China. Tel.: +86 024 83681106; fax: +86 024 23890912.
E-mail address: xiaoyan_105@163.com (Y. Chang).

http://dx.doi.org/10.1016/j.jprocont.2014.08.001
0959-1524/

based on CEB directly. In practice, CEB is affected and determined by raw materials fluctuation, changes of the external environment and process disturbance which are usually reflected in the change of process variations information that are closely related with CEB, i.e. CEB-related process variations. That is to say, the process variations are different from performance grade to performance grade. Thus, we should extract the process variations from the process measurement information and make full use of them for operating performance assessment.

From the above analysis, we know that how to accurately extract the CEB-related process variations is extremely important for online operating performance assessment. Recently, Zhou et al. [13,14] proposed a total projection to latent structures (T-PLS) for process monitoring and fault diagnosis. They analyzed the problem of conventional PLS which only divides the measured variable space into two subspaces and indicated that the output-unrelated variations are also included by PLS scores while PLS residuals do not necessarily cover only small process variations. The proposed T-PLS algorithm further decomposed the PLS systematic subspace and residual subspace, where the output-related part is separated from the output-unrelated part. From the perspective of operating performance assessment, the output-related process variations extracted by T-PLS can more accurately reflect the CEB. Therefore, T-PLS is used for information extracting in this study. Another important issue is how to automatically identify the causes when the process operating performance is nonoptimal which, however, has not been well addressed yet.

In this paper, a new performance-similarity-based online operating performance assessment and variable contributions-based nonoptimal cause identification method are proposed, which consists of offline part and online part. In offline part, the modeling data corresponding to different performance grades can be divided into several data sets according to the level of CEB, and each data set roughly represents one performance grade. However, owing to the fact that the existence of outliers seriously affects the accuracy of the assessment models, the outliers are be further deleted from each data set and only the remained data are used to model by T-PLS. Based on the understanding that different performance grades contain different CEB-related process variations, the assessment models for each performance grade are built by T-PLS.

In online implement, it should be noted that if the operating performance of the online data is one of the performance grades which have been modeled in the offline part, the CEB-related process variations contained in the online data must be consistent with that extracted from the modeling data of the corresponding performance grade. Therefore, we can use the similarities between CEB-related process variations of online data and those of each performance grade modeling data, named as performance-similarity for simplicity, to evaluate the current process operating performance. In addition, a production process usually includes multiple performance grades and performance grade conversions, and the process variations are usually unchanged when the process operates on a deterministic performance grade; however, the conversion between two performance grades is not at one stroke, and the process variations would change gradually from one performance grade to another in the process of performance grade conversion, which reflects in the change of the performance-similarity. Thus, the online assessment results based on the performance-similarity can not only include the deterministic performance grades but also the performance grade conversions. Moreover, when the operating performance is nonoptimal, the responsible variables can be identified by the variable contribution-based nonoptimal cause identification method, which dose not need any help of the expert knowledge. Through combining some production experience, the managers and operators can take appropriate operating adjustment strategy for production improvement.

The contributions of the present paper are summarized as follows: (i) a novel performance-similarity-based online operating performance assessment method is proposed for complex industrial processes, where the online assessment results contain both the deterministic performance grades and the performance grade conversations which were not covered in the existing assessment methods; (ii) a new variable contributions-based nonoptimal cause identification approach is developed, which is committed to identify the cause variables for nonoptimal operating performance without the help of expert knowledge.

The rest of this paper is organized as follows. First, a simple preparatory theoretical support is framed by a revisit to the standard T-PLS algorithm and then the offline modeling procedure is introduced based on T-PLS. Subsequently, the online operating performance assessment strategy and the nonoptimal cause identification method are presented in Section 3. In illustrations section, a case of a gold hydrometallurgical process is studied to demonstrate the feasibility and efficiency of the proposed method. Finally, the paper ends with some conclusions and acknowledgements.

2. Offline modeling for performance grades

2.1. T-PLS algorithm

T-PLS [13] is considered as an improved version of PLS. Through further decomposition, the output-unrelated and output-related part can be separated in PLS systematic subspace, and large process variations are separated from noise in residual subspace, which provides more accurate information for those who are more concerned with certain aspects of the whole information.

Assume that process data $\boldsymbol{X} \in R^{N \times J}$ consists of N samples with J process variables, and output data $\boldsymbol{y} \in R^{N \times 1}$ only contains the CEB, such as plant cost, profits, and product quality, etc. It should be noted that we shall treat only the case with the single output variable in the present paper; for the multiple output case, one just need to treat all of the output variables simultaneously.

Based on the nonlinear iterative partial least-squares (NIPALS) algorithm [15], the normalized ($\boldsymbol{X}$, $\boldsymbol{y}$) are projected into a low-dimensional space, and the decomposed formulation is shown in Eq. (1):

$$\begin{cases} \boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E} \\ \boldsymbol{y} = \boldsymbol{T}\boldsymbol{q} + \boldsymbol{f} \end{cases}, \tag{1}$$

where $\boldsymbol{T} \in R^{N \times A}$ and $\boldsymbol{P} \in R^{J \times A}$ are the score and loading matrices of $\boldsymbol{X}$, respectively. A is the number of PLS component, and $\boldsymbol{q} \in R^{A \times 1}$ is the loading vector of $\boldsymbol{y}$. To obtain $\boldsymbol{T}$ from $\boldsymbol{X}$ directly, the original weight matrix $\boldsymbol{R} = \boldsymbol{W}(\boldsymbol{P}^T\boldsymbol{W})^{-1} \in R^{J \times A}$ [16] is needed, where $\boldsymbol{W}$ is the weight matrix of $\boldsymbol{X}$ and obtained from the PLS decomposition.

Then, based on T-PLS algorithm, $\boldsymbol{X}$ and $\boldsymbol{y}$ can be further modeled as follows:

$$\begin{cases} \boldsymbol{X} = \boldsymbol{t}_y\boldsymbol{p}_y^T + \boldsymbol{T}_o\boldsymbol{P}_o^T + \boldsymbol{T}_r\boldsymbol{P}_r^T + \boldsymbol{E}_r \\ \boldsymbol{y} = \boldsymbol{t}_y + \boldsymbol{f} \end{cases}, \tag{2}$$

where $\boldsymbol{t}_y \in R^{N \times 1}$, $\boldsymbol{T}_o \in R^{N \times (A-1)}$ and $\boldsymbol{T}_r \in R^{N \times A_r}$ are scores of $\boldsymbol{X}$, and $\boldsymbol{p}_y \in R^{J \times 1}$, $\boldsymbol{P}_o \in R^{J \times (A-1)}$ and $\boldsymbol{P}_r \in R^{J \times A_r}$ are the loadings, correspondingly. $\boldsymbol{E}_r \in R^{N \times J}$ is the new residual matrix. A_r is the number of output-unrelated components. The detailed procedures for single output T-PLS algorithm is given in Appendix. In Eq. (2), $\boldsymbol{t}_y$ represents the output-related variations in score matrix $\boldsymbol{T}$ of original PLS model, $\boldsymbol{T}_o$ represents the output-unrelated variations in $\boldsymbol{T}$, $\boldsymbol{T}_r$ is the major part of original $\boldsymbol{E}$, and $\boldsymbol{E}_r$ is the residual part of $\boldsymbol{X}$. More details about T-PLS can be found in [13].