Engineering 3 (2017) 214–219

Contents lists available at ScienceDirect

Engineering

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

Research
Smart Process Manufacturing—Article

Real-Time Assessment and Diagnosis of Process Operating Performance

Shabnam Sedghi, Biao Huang*

Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta T6G 1H9, Canada

ARTICLE INFO

Article history:
Received 30 November 2016
Revised 26 January 2017
Accepted 2 February 2017
Available online 16 March 2017

Keywords:
Optimality assessment
Probabilistic principal component regression
Multi-mode

ABSTRACT

Over time, the performance of processes may deviate from the initial design due to process variations and uncertainties, making it necessary to develop systematic methods for online optimality assessment based on routine operating process data. Some processes have multiple operating modes caused by the set point change of the critical process variables to achieve different product specifications. On the other hand, the operating region in each operating mode can alter, due to uncertainties. In this paper, we will establish an optimality assessment framework for processes that typically have multi-mode, multi-region operations, as well as transitions between different modes. The kernel density approach for mode detection is adopted and improved for operating mode detection. For online mode detection, the model-based clustering discriminant analysis (MclustDA) approach is incorporated with some *a priori* knowledge of the system. In addition, multi-modal behavior of steady-state modes is tackled utilizing the mixture probabilistic principal component regression (MPPCR) method, and dynamic principal component regression (DPCR) is used to investigate transitions between different modes. Moreover, a probabilistic causality detection method based on the sequential forward floating search (SFFS) method is introduced for diagnosing poor or non-optimum behavior. Finally, the proposed method is tested on the Tennessee Eastman (TE) benchmark simulation process in order to evaluate its performance.

1. Introduction

Process operating performance assessment is an important subject in the process industry and has attracted attention in both academia and industry. Since the performance of processes may deteriorate over time and depart from the initial design due to process variations or process condition changes, it is necessary to continuously monitor process performance. This type of analysis is a step forward from traditional control performance assessment and has been named "optimality assessment."

Some studies on optimality assessment [1–4] have recently been conducted. However, these studies do not describe a method that is applicable to general, complicated process operations. In this paper, a systematic framework for optimality assessment is proposed that addresses the main issues associated with the performance assessment of modern industrial processes. First, we consider multiple operating modes due to the process condition and product demand changes. Second, we introduce multiple operating regions in each steady-state mode due to uncertainties and disturbances. Third, we consider transitions between different operating modes.

To solve these problems, a novel method for optimality assessment based on probabilistic principal component regression (PPCR) is proposed. It is first described for the unimodal processes that are common in practice, and is then extended to multiple operating mode processes. For unimodal processes, the developed method consists of two stages: offline training and online assessment. In offline training, the steady-state data, including process variables and the optimality index (OI), are collected. Note that the OI definition depends on the process. For example, depending on the process, OI can refer to operation costs, profit, product quality, environmental index, and so on. To obtain an online estimation of OI, it is necessary to build a predictive model of OI based on the process variables. Since each operating mode usually has multiple operating regions, the mixture probabilistic principal component regression (MPPCR)

* Corresponding author.
E-mail address: biao.huang@ualberta.ca

http://dx.doi.org/10.1016/J.ENG.2017.02.004

model is employed for modeling. The MPPCR model describes the Gaussian distribution of OI in each operating region, based on which the local value of OI in each operating region can be obtained. By comparing the local OI in each operating region, their optimality condition is analyzed. In online assessment, the operating region of a new data point is estimated based on its posterior probability. Based on the constructed model, OI is predicted using Bayesian inference to evaluate the process performance. When the process performance is non-optimum, diagnosing the cause of the problem helps steer the process to a better performance. A probabilistic contribution analysis technique based on the missing variable approach [5] is adopted to address this issue. The sequential floating forward search (SFFS) method is utilized, instead of a branch and bound method, in order to decrease computational time and simplify the solution.

For multiple operating mode processes, it is assumed that the data points are unlabeled with respect to the operating modes. In other words, the number of operating modes and the operating mode of each data point are unknown. To estimate the labels of the dataset, critical process variables that govern the change of operating modes are selected and named as scheduling variables. Based on the selected scheduling variables, a local kernel-density-based approach [6] is adopted and improved to detect the labels of the data points. In order to estimate the operating modes in online assessment, a mixture discriminant analysis (MDA) is built based on the labeled dataset. In addition, to improve the accuracy of online mode detection, the process knowledge is incorporated into the MDA results. Optimality assessment of steady-state modes is the same as for unimodal processes. For transitions between modes, a dynamic principal component regression (DPCR) model was built, and the performance grades are compared based on the DPCR loading matrices [7].

The rest of this paper is arranged as follows: In Section 2, the problem and the proposed solution are discussed. In Section 3, the proposed optimality assessment strategy for steady-state modes is described. In Section 4, the assessment method for transitions is studied. In Section 5, the mode detection method for multiple operating mode processes is described. In Section 6, the proposed approach is tested on a Tennessee Eastman (TE) process. Finally, conclusions are presented.

2. Problem statement and proposed solution

General process operations have multi-modal characteristics with non-Gaussian behavior in each steady-state mode. An overview of these systems is given in Fig. 1. It is considered that the change of operating modes is caused by known governing factors such as product demand. In addition, each steady-state operating mode consists of different operating regions that are caused by uncertain process variations. The optimality level is altered based on the operating position in the system.

In this paper, the goal is to assess online operating process performance based on routine operating process data by characterizing the data—that is, by estimating operating mode and operating region (or transition grade), predicting the OI value, and diagnosing the cause of poor performance. The proposed framework includes offline training and online assessment. An overview of the proposed framework and methods is given in Fig. 2 and Fig. 3, and the details are described in the following sections.

3. Steady-state modes: Definition

Steady-state modes are the main operating conditions of processes during which no essential change occurs in the critical process variables, flowsheet configuration, product demand, and so on. The MPPCR model is utilized to estimate the model of the training dataset. In the next step, based on the detected model, the local OI values of each operating region are obtained. Furthermore, based on process knowledge, some classes for optimality values are defined, and the obtained operating regions are assigned with various corresponding classes.

3.1. Data modeling

Suppose $X = [x(1), x(2), \ldots, x(n)]^{\mathrm{T}} \in R^{n\times p}$ and $Y = [y(1), y(2), \ldots, y(n)]^{\mathrm{T}} \in R^{n\times 1}$ are the available datasets of the process variables and OI, respectively, where n is the number of data points and p is the number of process variables. Since operating modes consist of several operating regions, the MPPCR model is employed to build a predictive model for which the input is X and the output is Y.

3.2. Analysis of optimality index

The Gaussian distribution for the OI in each operating region k is estimated in the modeling part. As a result, the local OI [1] in each operating region is equal to the mean value of the obtained Gaussian distribution for y:

$$\mathrm{OI}_k = E(y) = \int f_k(y) y \, \mathrm{d}y = \mu_{y,k} \tag{1}$$

3.3. Non-optimum cause detection

In order to find causal variables in the presence of non-optimum or poor performance, one can utilize a probabilistic contribution

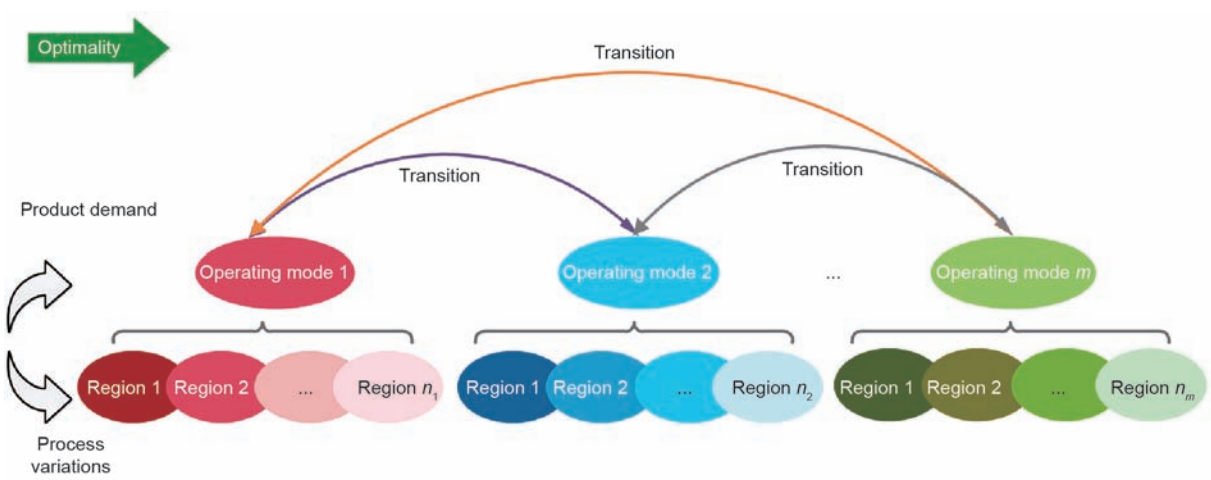

Fig. 1. An overview of general process operations.