Chemical Engineering Science 137 (2015) 140–151

Contents lists available at ScienceDirect

Chemical Engineering Science

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

Ensemble local kernel learning for online prediction of distributed product outputs in chemical processes

Yi Liu [a], Zhengjiang Zhang [b], Junghui Chen [c,*]

[a] Engineering Research Center of Process Equipment and Remanufacturing, Ministry of Education, Zhejiang University of Technology, Hangzhou 310014, People's Republic of China
[b] College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, People's Republic of China
[c] Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taoyuan 320, Taiwan, Republic of China

HIGHLIGHTS

- Kernel learning (KL) models for prediction of the distributed product outputs.
- A just-in-time KL (JKL) method for better description of local distribution shapes.
- A *T*-region-based strategy is proposed to reduce the unnecessary search regions.
- An ensemble JKL modeling approach for getting more reliable prediction performance.
- Research shows online distribution shape prediction outperforms the other methods.

ARTICLE INFO

Article history:
Received 13 October 2014
Received in revised form
11 May 2015
Accepted 3 June 2015
Available online 12 June 2015

Keywords:
Batch
Chemical processes
Distributed product outputs
Kernel learning
Process control
Systems engineering

ABSTRACT

The crystal size distribution in crystallization processes, the molecular weight distribution in polymerization processes, the particle size distribution in powder industries, and the pulp fiber length distribution in paper industries are all distributed product outputs. Reliable online quality prediction of these chemical processes with distributed outputs is important but challenging. In this work, the kernel learning (KL) framework is introduced to model and online predict the distributed product outputs. First, the KL method is proposed to construct a global distributed shape. Then, without resorting to a KL-based global distributed model, a just-in-time KL (JKL) model is presented for better description of local distributed shapes with more accurate and quick prediction performance. Moreover, an ensemble JKL (EJKL) modeling approach is developed to obtain more reliable prediction performance of the distributed outputs. The proposed modeling methods are applied to online prediction of the molecular weight distribution in polymerization processes and of the crystal size distribution in crystallization processes. The prediction results show the proposed method is superior to the traditional counterparts.

1. Introduction

In recent years, there has been increasing demand for product diversification because of the highly competitive market in the chemical industry. Shorter product life cycle and more stringent specifications in the properties are required. Moreover, there are some chemical processes where the output product of interest has the nature of a distribution rather than a single value. For example, as shown in Fig. 1, the molecular weight distribution (MWD) in polymerization processes (Clarke-Pringle and MacGregor, 1998; Crowley and Choi, 1998; Kiparissides, 2006; Richards and Congalidis, 2006; Takamatsu et al., 1988; Vicente et al., 2003), the crystal size distribution (CSD) in crystallization processes (Braatz, 2002; Chung et al., 1999; Fiordalis and Georgakis, 2013; Hsu and Ward, 2013), the particle size distribution (PSD) in powder industries (Crowley et al., 2000; Flores-Cerrillo and MacGregor, 2002; Immanuel and Doyle III,

Abbreviations: CSD, crystal size distribution; CSF, cumulative similarity factor; EJKL, ensemble just-in-time kernel learning; FLOO, fast leave-one-out; GPR, Gaussian process regression; JIT, just-in-time; JKL, just-in-time kernel learning; KL, kernel learning; LOO, leave-one-out; LSSVR, least squares support vector regression; MLP-NN, multilayer-perceptron-neural networks; MWD, molecular weight distribution; NN, neural networks; PSD, particle size distribution; RBF-NN, radial basis function-neural networks; RMSE, root-mean-square error; RVM, relevance vector machine; SF, similarity factor; SVR, support vector regression

* Corresponding author. Tel.: +886 3 2654107; fax: +886 3 2654199.
E-mail address: jason@wavenet.cycu.edu.tw (J. Chen).

http://dx.doi.org/10.1016/j.ces.2015.06.005
0009-2509/

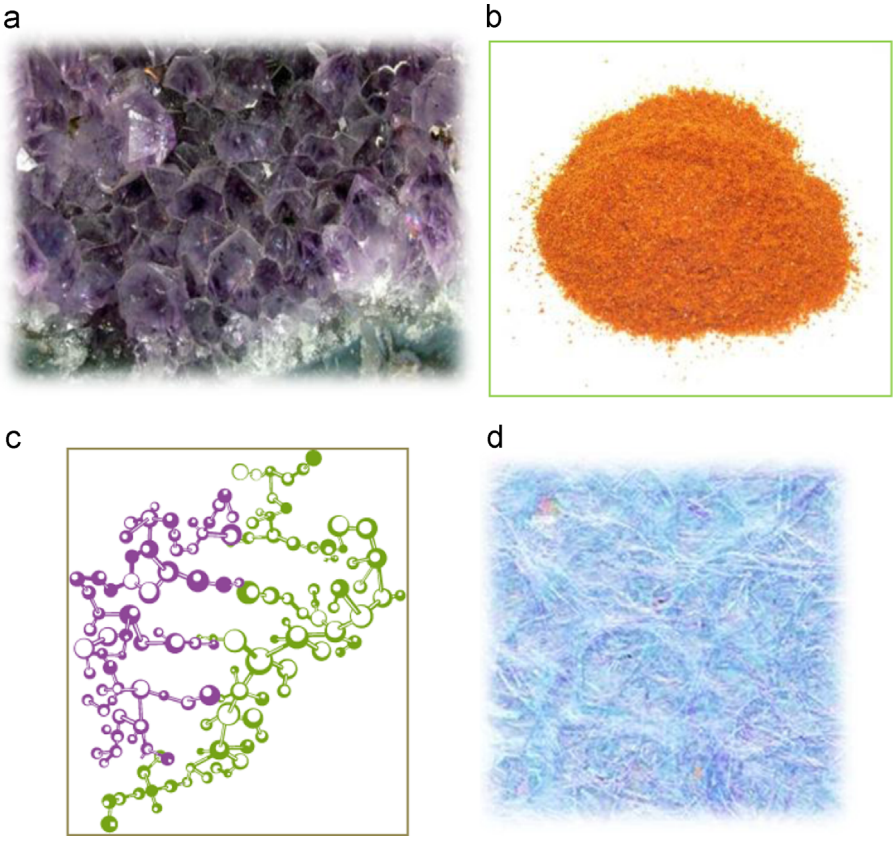

Fig. 1. Chemical process with distributed outputs: (a) crystal, (b) powder, (c) molecular, (d) paper fiber.

2002; Ko and Shang, 2011; Shi et al., 2006; Zeaiter et al., 2006), and the pulp fiber length distribution in paper industries (Smook, 2002; Wang, 2000) are all distributed outputs (Wang, 2000). In polymerization processes, the MWD directly affects many end-use properties, such as thermal properties, stress–strain properties, impact resistance, strength, and hardness (Kiparissides, 2006; Richards and Congalidis, 2006). In crystallization processes, the CSD is critically important in the production of high quality products and for determining the efficiency of downstream operations, such as filtration and washing (Braatz, 2002).

For modeling of product qualities in those chemical processes with distributed outputs, lots of previous research has focused on the single lumped value for the distributed quality variables. For example, the melt index is often considered as the product quality in polymerization processes (Mat Noor et al., 2010; Liu and Gao, 2015; Lou et al., 2012; Zhang et al., 2006). However, recent studies have shown the fact that the economic gain can be further improved by shaping the distributed output since it can significantly influence the product quality and process efficiency (Alhamad et al., 2005; Braatz and Hasebe, 2002; Chang and Liao, 1999; Nagy, 1999; Nagy and Braatz, 2012; Pigeon et al., 2011; Wang, 2000). Generally, such chemical processes are non-Gaussian and may also show strong nonlinearities (Wang, 2000, 2002; Zeaiter et al., 2006). In these cases, the output means and variances are not sufficient to characterize the probabilistic behavior of the stochastic outputs of the process. Instead, the purpose of the controller design should be tracking of the distributed outputs with respect to a desired distributed shape (Pigeon et al., 2011; Wang, 2000). Consequently, online acquisition of reliable and accurate distributed output information (i.e., the product quality) is important for further development of good control schemes.

However, online MWD measurement in polymerization processes and online CSD measurement in crystallization processes are still unsolved (Braatz, 2002; Kiparissides, 2006). For these complex processes with distributed outputs, it is generally difficult to develop a comprehensive first-principle model, although great efforts have been made by many researchers (Braatz, 2002; Hsu and Ward, 2013; Kiparissides, 2006; Richards and Congalidis, 2006). Moreover, a set of partial differential equations for the distributed outputs are only valid to white noise inputs. As a result, the realization of distributed output control for these complicated processes is still challenging. Recently, data-driven modeling methods have been useful alternatives for timely online prediction of product qualities in chemical processes when online analyzers are not available. Current popular methods dealing with the approximation of the distributed output are B-splines and other neural networks (NN) (Guo and Wang, 2005a,b, 2006; Guo et al., 2008; Wang, 2002; Wang and Zhang, 2001; Wang and Wang, 2002; Yue et al., 2008). However, the determination of network topology and the generalization capability of NN for a given modeling task are still unsolved. Additionally, a substantial amount of training examples is often required by NN methods.

Recently, the support vector regression (SVR), least squares-SVR (LSSVR), and other kernel learning (KL) methods have been increasingly used for applications in chemical process modeling (Kadlec et al., 2009; Liu et al., 2012, 2013a). The obtained results indicate that SVR is a promising and alternative method for nonlinear process modeling, especially when training data are limited. Compared with B-splines and NN-based models, one main advantage of SVR is that it can automatically derive the optimal network structure with respect to generalization error (Schölkopf and Smola, 2002; Suykens et al., 2002; Vapnik, 1995). Therefore, the development of suitable SVR/KL-based modeling and prediction method for distributed product outputs is the main aim of this work.

Generally, a single global model can be trained using all observations. However, even if an accurate model has been developed, its estimation performance is likely to deteriorate