Journal of Process Control 62 (2018) 37–43

Contents lists available at ScienceDirect

Journal of Process Control

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

Review

Mixed-effects Gaussian process modeling approach with application in injection molding processes

Linkai Luo [a,*], Yuan Yao [b], Furong Gao [c,*], Chunhui Zhao [d]

[a] Department of Computing, Hang Seng Management College, Siu Lek Yuen, Sha Tin, Hong Kong
[b] Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC
[c] Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
[d] The State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 18 January 2017
Received in revised form
23 September 2017
Accepted 12 December 2017

Keywords:
Mixed-effects model
Gaussian process
Multi-process data
Injection molding

ABSTRACT

We propose a new nonparametric approach for multi-process data analysis, in which each of the process is modeled as a combination of a fixed-effect and a random-effect Gaussian process (GP) regression model, namely, a mixed-effect Gaussian process (ME-GP) model. The ME-GP approach provides a flexible means to combine the common aspects of all processes and describe the heterogeneity among different processes. In particular, we model the mean and covariance structures of both the fixed- and random-effects simultaneously, and predict a future input using probability density distributions. We apply the ME-GP model to predict the melt-flow-length for filling of different molds in injection molding processes. It is shown that the ME-GP model obtains an improved performance against GP model only.

Contents

1. Introduction

In many engineering problems we are not only interested in modeling one process but also multiple processes that share common information [1–4]. When different processes are closely related one can model all processes simultaneously and try to attain improved predictive performance by taking advantage of the common aspects of all processes involved. Often, multi-process data from similar processes are present. Consequently, it is desirable to develop a new model for the associated similar, yet nonidentical process using the available multi-process data.

Take a real-world injection molding process for example. Injection molding is an important polymer processing technique that transforms thermoplastic into various shapes and types of products [5]. As a cyclic process, injection molding consists of three stages: filling, packing-holding and cooling. The motivated example concerns data collected during the filling stage, in which an injection screw moves forward and pushes the polymer melt into the mold cavity. Melt development in the cavity during filling plays a key role in determining the quality of the product. The output of this stage is the trajectory of the melt-flow-length (MFL), represented as the distance that the melt-front has traveled inside the mold

* Corresponding authors.
E-mail addresses: llk1896@gmail.com (L. Luo), kefgao@ust.hk (F. Gao).

https://doi.org/10.1016/j.jprocont.2017.12.003

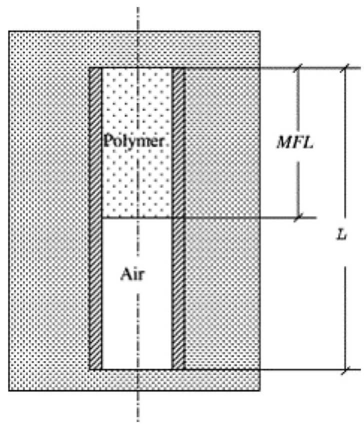

Fig. 1. Illustration of melt flow during filling stage.

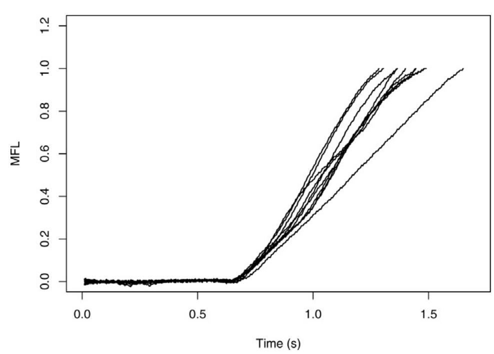

Fig. 2. Ten curves of melt flow length (y-axis, normalized value) against time (x-axis, in seconds). Each curve corresponds to one insert mold.

from the gate; Fig. 1 shows a simplified diagram of the melt flow in a single-gate rectangular mold. MFL is an important parameter reflecting the melt-flow status in the mold cavity; in turn, understanding its dynamic behavior helps us to achieve the perfect filling of the mold, which is always desired in injection molding applications. In reality, it is very difficult to measure the MFL on-line unless some expensive sensor is used and set up in the equipment. Hence, one of the objectives of the investigation is to develop a data-based model for reconstructing the trajectory of MFL by using some easily measured quantities, such as the screw displacement, the injection velocity, the nozzle temperature, etc. Previous studies show that eight such input variables are observed having significant impacts on the trajectory [6–8]. Let $y(t)$ denote the trajectory at time t. The goal is to develop a relationship f to model $y(t)$ with respect to d input variables $\boldsymbol{x}(t) = (x_1(t), \ldots, x_d(t))^T$:

$$y(t) = f(x_1(t), \ldots, x_d(t)) + \epsilon(t), \tag{1}$$

where $\epsilon(\cdot)$ denotes random measurement errors.

In real-world industry, the insert molds are often changed in the injection molding machines, in order to produce various shapes and types of products. Fig. 2 depicts ten curves of MFL trajectory when using ten different but basically similar designed molds; for details see Section 3. For each mold used, both the output variable $y(t)$ and the input variables $\left\{x_1(t), \ldots, x_d(t)\right\}$ are measured until the mold is completely filled by the polymer melt. We refer to $y^m(t)$ the curve when using mth mold ($m = 1, \ldots, 10$).

A quick glance at Fig. 2 shows that the curves of the trajectory behave in a much similar manner. For instance, the melt-flow-length begins to grow after nearly the same time point, because the velocity settings are identical for each process and that the required time for the polymer melt to reach the gate shall also be identical. Additionally, the curves are monotonically increasing with respect to time immediately after the polymer melt begins to fill the mold. These findings reveal that the underlying process is similar regardless the change of mold, and that the curves thereby share similar trend. In this context, common information can be shared, or tied, among all processes that correspond to different molds, such as having similar trajectory.

The objective of this paper is to combine data across multiple processes when a rich source of process data are available, and aims to improve the model accuracy of a single process by taking advantage of all available multi-process data from similar processes. The main idea is that the measurements taken on one process may be informative with respect to the others. In this paper, we try to explore functions to simultaneously model the trend that is common to all processes and the trend that is unique to the individual process. To be specific, we focus on functional mixed-effect model, where each mth process is modeled as a sum of a *fixed* effect shared by all processes and a *random* effect that is interpreted as the specific deviation from the fixed term:

$$y^m(t) = \bar{f}(\boldsymbol{x}) + \tilde{f}^m(\boldsymbol{x}) + \epsilon(t), \tag{2}$$

where $\bar{f}$ is the fixed term, and $\{\tilde{f}^m\}$ is a set of random terms. Eq. (2) is called mixed-effects model in literature [9,10]. The choice of fixed and random terms plays an essential role in fitting the mixed-effects model. A commonly used method is to employ linear regression function, known as linear mixed-effects model; see, [11–13], for example. Such methods, however, may not posses sufficient flexibility to model nonlinear systems. To overcome limitation of linear mixed-effects model, mixed-effects Gaussian processes (GP) model, wherein both the fixed and random terms are assumed to be realizations of Gaussian processes, is proposed [14]. GP has been proven to be a powerful tool to model a variety of complex physical processes that frequently occur in engineering applications [15–17], and is shown to have capability of accounting for the dependency between the linear fixed effect term and the response. Furthermore, the use of GP avoids the trivial treatment of choices of both fixed and random terms, and hence is more flexible than the other methods. Throughout this paper, we will define the proposed model (2) as mixed-effects Gaussian processes (ME-GP) model. Related work can be found in [18–21] and the references therein.

Our main contributions are twofold. (1) We model the multi-subject data with nonparametric fixed- and random-effect models. In [19], for example, the fixed- and random-effects models are linear regression; while in [20], the authors proposed multi-tasks learning algorithm, without modeling the fixed- and random-effects separately. (2) We apply the mixed-effects models in the discipline of process system engineering, while most of related work was in the domain of physical, biological and social sciences.

There is other work that relates to multi-process modeling. In [22] a migration method under certain assumption was presented. The author employed ensemble learning to develop the target model by averaging the other models. In [23], a scale-bias correction methodology to transfer one model to another was proposed. However, the scale and bias are both constant parameters, which cannot express the model nonlinearity. Then, in [24] the author extended the method by adopting Gaussian process function to approximate the scale and bias terms. These methods, however, heavily reply on the performance of the models, or worse, the error in models may be exaggerated. Moreover, they are two-stage modeling strategies, in which the models are first built and then, another similar model is developed by transferring/averaging the available models. The ME-GP technique, however, models the multiple processes simultaneously, but learns the common and unique parts separately. This, in turn, can obviously decrease error exaggeration during multi-process modeling.