European Journal of Operational Research 264 (2018) 212–224

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

Quantile regression metamodeling: Toward improved responsiveness in the high-tech electronics manufacturing industry

Demet Batur[a,*], Jennifer M. Bekki[b], Xi Chen[c]

[a] Supply Chain Management and Analytics, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
[b] The Polytechnic School, Ira A. Fulton Schools of Engineering, Arizona State University, Mesa, AZ 85251, USA
[c] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24601, USA

ARTICLE INFO

Article history:
Received 2 August 2016
Accepted 6 June 2017
Available online 10 June 2017

Keywords:
Manufacturing
Lead-time quotation
Simulation metamodeling
Predictive analytics

ABSTRACT

Both technology and market demands within the high-tech electronics manufacturing industry change rapidly. Accurate and efficient estimation of cycle-time (CT) distribution remains a critical driver of on-time delivery and associated customer satisfaction metrics in these complex manufacturing systems. Simulation models are often used to emulate these systems in order to estimate parameters of the CT distribution. However, execution time of such simulation models can be excessively long limiting the number of simulation runs that can be executed for quantifying the impact of potential future operational changes. One solution is the use of simulation metamodeling which is to build a closed-form mathematical expression to approximate the input–output relationship implied by the simulation model based on simulation experiments run at selected design points in advance. Metamodels can be easily evaluated in a spreadsheet environment "on demand" to answer *what-if* questions without needing to run lengthy simulations. The majority of previous simulation metamodeling approaches have focused on estimating *mean* CT as a function of a single input variable (i.e., throughput). In this paper, we demonstrate the feasibility of a *quantile regression based metamodeling* approach. This method allows estimation of CT *quantiles* as a function of multiple input variables (e.g., throughput, product mix, and various distributional parameters of time-between-failures, repair time, setup time, loading and unloading times). Empirical results are provided to demonstrate the efficacy of the approach in a realistic simulation model representative of a semiconductor manufacturing system.

1. Introduction

The high-tech electronics manufacturing industry faces a number of important challenges. The industry is dominated by consumer goods that are highly driven by customer satisfaction measures. Globalization has made the supply networks more complicated and difficult to manage, while shorter product life cycles and increasing product complexity make it harder for the technological developments to keep up with market demands. In such environments, on-time delivery is often noted as the most important factor in predicting service levels (Boyaci & Ray, 2006; Mönch, Fowler, & Mason, 2013). Accurate quoting of delivery dates also has a significant impact on inventory control, scheduling, and other decision-making problems in manufacturing systems. Consequently, the ability to meet the stated delivery dates across the supply chain is essential to today's highly competitive global market. The importance of this not only to final product manufacturers but also to equipment suppliers for the high-tech electronics industry is highlighted in Atan, de Kok, Dellaert, van Boxsl, and Janssen (2016). Further complicating the industry's ability to obtain high customer service levels is the fact that consumer demand changes rapidly. Therefore, being able to predict delivery dates for *future* operating conditions of the system is as important as generating accurate estimates of delivery dates for *current* operating conditions.

An essential component of providing customers with accurate delivery dates for a given product (both under current and future operating conditions) is the ability to quickly generate accurate estimates of parameters of the product's CT distribution. Here, we consider CT as a random variable representing the time required for a job or a lot to get through the entire manufacturing process (including processing, waiting, and transportation times).

While estimates of *mean* CT are relatively easy to obtain and so are often readily available, using them to generate estimates of

* Corresponding author.
E-mail addresses: dbatur@unl.edu (D. Batur), jennifer.bekki@asu.edu (J.M. Bekki), xchen6@vt.edu (X. Chen).

http://dx.doi.org/10.1016/j.ejor.2017.06.020
0377-2217/

delivery dates ignores inherent variability associated with the CT distribution, and can result in reduced on-time delivery and service to customers. Estimates of *quantiles* of the CT distribution, on the other hand, provide the decision maker with a complete picture of the CT behavior, allowing delivery-date quotes to be made with varying levels of confidence (and associated risk). For example, if the 0.9-quantile of the CT distribution is used in setting up the delivery dates, 90% of the orders will be delivered on time. There is a growing amount of work in the estimation of quantiles (Alexopoulos, Goldsman, & Wilson, 2012; Bekki, Mackulak, Fowler, & Nelson, 2010; Calvin & Nakayama, 2013; Chen & Kelton, 2006; 2008; Jin, Fu, & Xiong, 2003; Raatikainen, 1987; 1990). The importance and relevance of the use and estimation of CT quantiles in manufacturing systems are highlighted by Ankenman et al. (2011), Bekki, Mackulak, Fowler, and Nelson (2010) and Tai, Pearn, and Lee (2012).

Across all manufacturing systems, a key driver of CT is throughput (TH). We define TH to be the overall production rate of the system as a percentage of the system's capacity, and it can be controlled at the operations level by the rate of introduction of new jobs or lots (start rate). Several other controllable operational factors, such as product mix and various distributional parameters of time-between-failures, repair time, setup time, loading and unloading times, have also been identified in the literature as closely tied to CT (Meidan, Lerner, Rabinowitz, & Hassoun, 2011). Therefore, determining the relationship between the CT distribution and these factors is critical in developing accurate CT predictions. If analytical models were available to describe these relationships, they would provide unambiguous functions that can be used in the prediction of CT. However, the use of such analytical models is restricted by the fact that they are only applicable to simple systems that satisfy fairly limiting assumptions. Consequently, in the complex and stochastic environment of the high-tech manufacturing industry, discrete-event simulation models are typically used to predict CTs (Mönch, Fowler, & Mason, 2013).

Though regularly utilized, simulation models of high-tech electronics manufacturing facilities are quite complex, and can take a considerable amount of time to execute for a given set of input variable values. The strain of executing such models is particularly evident when *what-if* scenarios are examined for analyzing alternate or future operating conditions where input variables take various new values. Further elongating the time to perform such analyses is the fact that multiple simulation replications need to be run for each examined *what-if* scenario to mitigate the impact of stochasticity in simulation models. Instead, it would be desirable to develop a regression function approximating the simulation input–output relationship based on the current input variable values and use this regression model to predict how the system would perform at the new input variable values that are considered in the *what-if* scenarios. Such regression models of simulation input–output relationships are known as *metamodels*, and have become an increasingly popular modeling tool (Barton, 2009). Ideally, metamodels provide the fidelity of the full simulation model with the ease of use of a spreadsheet environment.

Metamodels are tools developed to be used in production planning (Ankenman et al., 2011). For example, a practitioner could investigate the impact to a CT quantile of reducing the coefficient-of-variation (CoV) of unloading times, perhaps through increased automation. Similarly, an operations manager could quantify the resulting impact to a CT quantile of increasing the production rate of a product that his facility produces in response to anticipated demand changes for the product. In both cases, this information can be used not only to evaluate the impact and significance of potential, future changes to the production system, but also to quickly adjust lead-time quotations to customers if the changes are ultimately implemented on the manufacturing floor. Some applications of simulation metamodeling in the analysis of production systems can be found in Kesen, Toksari, Güngor, and Güner (2009), MacDonald and Gunn (2011) and Yang (2010).

An important part of the metamodeling process is the selection of a metamodeling (regression) function, $g(\mathbf{x})$, that allows the prediction of the desired output performance measure for a given vector of inputs, $\mathbf{x}$, in a simulation model. The *mean* CT is the most commonly used output measure in the simulation metamodeling literature, and there are several papers that focus on the relationship between the mean CT and TH (Fowler, Park, Mackulak, & Shunk, 2001; Park, Fowler, Mackulak, Keats, & Carlyle, 2002; Veeger, Etman, van Herk, & Rooda, 2010; Yang, 2010). However, as previously discussed, *quantiles* provide a more comprehensive understanding of the CT distribution and are of greater use to decision makers.

When the CT output variable is denoted by Y with distribution F_Y, the q-quantile is defined as

$$y^{[q]} = F_Y^{-1}(q) = \inf\{y : F_Y(y) \geq q\} \text{ for } q \in (0, 1).$$

We propose the use of the Quantile Regression (QR) method (Koenker, 2005) to construct a metamodeling function, $g(\mathbf{x})$, that is suitable for predicting CT quantiles given a vector of inputs, $\mathbf{x}$, in simulation models of manufacturing systems using the following relationship

$$y^{[q]} = g(\mathbf{x}) + \varepsilon,$$

where the randomness of the simulation model is represented by the random error ε. A strength of the proposed method originates from the fact that no distributional assumptions are made for F_Y, and hence ε.

In summary, the goal of this work is to develop a *computational framework* for predicting CT quantiles of high-tech electronics manufacturing facilities. We will develop and demonstrate the QR metamodeling procedure, which is of particular use for predicting CT quantiles of manufacturing systems in which there is more than one input variable of interest. Although the theory of QR modeling is well developed, the application of it to CT quantiles in manufacturing systems has many theoretical and computational challenges. This paper addresses the issues related to specifying the form of the quantile metamodeling function $g(\mathbf{x})$ in the case that multiple input variables are present and finding a parsimonious model fit is technically challenging. Notably, while the proposed approach can theoretically be used for a large number of input variables, the work presented here demonstrates the efficacy of the approach up to only three input variables.

There are only a small number of studies that focus on the relationship between CT quantiles and TH (Chen & Zhou, 2011; Yang, Ankenman, & Nelson, 2008). Both of these techniques cannot be extended to metamodels of input variables other than TH. Our contribution in this paper is to consider multiple input variables simultaneously, not only TH, but also product mix and various distributional parameters of time-between-failures, repair time, setup time, loading and unloading times, etc. To the best of our knowledge, this is the first study that shows how QR modeling can be used to build metamodels of CT quantiles with multiple input variables in the manufacturing environment.

In the upcoming sections of the paper, we more specifically articulate the problem and describe the specific simulation model that was used as an experimental testbed for validating the proposed mechanism of predicting CT quantiles of high-tech electronics manufacturing facilities. We then provide the rationale for the proposed approach and give the details of the QR metamodeling procedure. Later in the paper, experimental results are provided, followed by discussions and future work.