*Proceedings of the 2016 Winter Simulation Conference*
*T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.*

# EXTENDED KERNEL REGRESSION: A MULTI-RESOLUTION METHOD TO COMBINE SIMULATION EXPERIMENTS WITH ANALYTICAL METHODS

Ziwei Lin, Andrea Matta, Na Li

Department of Industrial Engineering & Management
School of Mechanical Engineering
Shanghai Jiao Tong University
800 Dongchuan Road
Shanghai, 200240, CHINA

J. George Shanthikumar

Krannert School of Management
Purdue University
100 S. Grant Street
West Lafayette, IN 47907-2076 , USA

## ABSTRACT

Simulation is widely used to predict the performance of complex systems. The main drawback of simulation is that it is slow in execution and the related compute experiments can be very expensive. On the other hand, analytical methods are used to rapidly provide performance estimates, but they are often approximate because of their restrictive assumptions. Recently, Extended Kernel Regression (EKR) has been proposed to combine simulation with analytical methods for reducing the computational effort. This paper has different purposes. Firstly, EKR is tested on different cases and compared with other techniques. Secondly, two different methods for calculation of confidence band are proposed. Numerical results show that the EKR method provides accurate predictions, particularly when the computational effort is low. Results also show that the performance of the two confidence band methods depends on the case analyzed. Thus, further studies are necessary to develop a robust method for confidence band calculation.

## 1 INTRODUCTION

The detailed design of a system configuration (e.g., production lines in a plant, the emergency department in a hospital), usually requires that a large number of alternative solutions must be analyzed and evaluated. Simulation and analytical methods are two classic and effective tools for system performance evaluation in system design.

Analytical methods are frequently used to estimate the performance of stochastic systems. Queuing theory and Markov chains are the most widely used frameworks under which analytical methods are developed. In literature, a number of analytical methods are available for studying a specific system (Askin and Standridge 1993, Buzacott and Shanthikumar 1993, Gershwin 1994, Papadopoulos et al. 2009, Li and Meerkov 2008). Analytical methods can provide system performance estimates fast. Nevertheless, this advantage is counterbalanced by the bias of the provided estimates. This is because approximations will be introduced either in the system assumptions or in the mathematical derivation of the solution equations to make analytical methods solvable for complex systems. In general, the more complex the system is, the larger the approximation introduced in analytical methods.

Simulation is used when a high-accurate estimate of system performance is required. The accuracy of simulation depends on the detail level of the model and the length of simulation runs. Simulation is also used to quantitatively assess the bias of newly developed analytical methods. In this case, a design of experiments (DOE) is developed by running simulation and analytical methods on a set of points sampled according to some criteria. Then, deviations of analytical methods from simulation are calculated by considering the simulation outputs as the benchmark. Simulation needs considerable effort to build models and it is time-consuming to evaluate all the feasible solutions. To overcome this problem, regression techniques

based on kriging or radial basis function, like Stochastic Kriging (SK) (Ankenman, Nelson, and Staum 2010) and Kernel Regression (KR) (Wand and Jones 1995), are often used. They can provide the estimates of expected system performance at unknown points by using an empirical model built with the simulation outputs of a DOE. Nevertheless, they usually require a large number of design points to perform well when the underlying response surface is not smooth or the dimension of the input is high.

The estimates provided by regression techniques can be improved by combining simulation with analytical methods. High-fidelity models (mainly simulation models) generate accurate measurements of the system performance, whereas low-fidelity models (low accurate simulation models or analytical methods) capture some basic features of the system response. This problem is known as multi-fidelity regression modeling and several methods have been proposed in literature. Co-Kriging (Cressie 1993, Forrester and Keane 2009) merges data from different fidelity sources by extending Kriging estimator. This technique is widely used in engineering. Goh et al. (2013) use a Bayesian approach to combine results from different simulators. Chen et al. (2015) use locally weighted regression and smoothing method (LOESS) to combine simulation results with those provided by a Jackson queuing network.

Extended Kernel Regression model (EKR) (Matta et al. 2015) is a meta-model that uses both simulation and analytical methods to evaluate the system performance. EKR is considered as a non-parametric technique. Because it is not necessary to make any assumptions about the fitted function before the evaluation. Instead, it lets the data show the structure. From another perspective, the EKR model can be seen as a parametric regression method in which the analytical method provides the shape of the function, whereas the rest is locally fitted by kernel functions.

The contribution of this paper is twofold. Firstly, the EKR method is compared with other available techniques, such as SK and KR, by executing numerical experiments on two cases: a $G/G/1$ system and a Closed-Loop Flexible Assembly System (CLFAS). Secondly, the paper proposes two methods to calculate confidence band, which are useful to cut unpromising alternatives during the design of the system. The first method is derived by assuming that the normalized bias of the EKR estimator is a Gaussian process. The second one is derived from basic statistic knowledge.

This paper is organized as follows. Section 2 presents the notation throughout the paper and describes how to build EKR models. Section 3 proposes two methods for confidence band calculation. Section 4 describes the application of the EKR model in a $G/G/1$ system. In section 5, EKR is applied to a CLFAS system. Finally, conclusions and guidelines for future developments are drawn in section 6.

## 2 EXTENDED KERNEL REGRESSION

### 2.1 Notation

We are interested in evaluating the system performance $Y$ as a function of the system description $\boldsymbol{x}$. $Y$ is a univariate random variable and $\boldsymbol{x} \in \mathscr{D} \subset \mathbb{R}^d$ is a $1 \times d$ vector where $\mathscr{D}$ is the set containing all the alternative system configurations.

We assume that a DOE has been developed and two kinds of outputs are available from the experiments: high-fidelity estimates from simulation and low-fidelity estimates from analytical method. There are $n$ design points in the DOE and each point is a system configuration $\boldsymbol{x_i^0} = (x_{i1}^0, \cdots, x_{id}^0) \in \mathscr{D}, i \in \mathscr{N} = \{1, \ldots, n\}$ where $x_{ik}^0$ represents the value of $k$-th dimension (with $k \in \mathscr{K} = \{1, \ldots, d\}$) of the system configuration at the design point $i$. The index 0 in the notation indicates that the variable belongs to the initial design. Let us denote the output of experiment $j$ on the design point $i$ with $y_i^{0j}$. As convention, we use $j = s$ for simulation experiments and $j = a$ for experiments with analytical method.

The problem is stated as follows. We want to estimate the expected system performance $y(\boldsymbol{x})$ at the unobserved point $\boldsymbol{x} = (x_1, \cdots, x_d) \in \mathscr{D}$ using the observations collected from the initial experiment $y_i^{0s}(\boldsymbol{x_i^0}), y_i^{0a}(\boldsymbol{x_i^0})$ and the performance estimates $y^a(\boldsymbol{x})$ provided by analytical method at $\boldsymbol{x}$:

$$y(\boldsymbol{x}) = \mathbb{E}(Y | y^a(\boldsymbol{x}), y_i^{0s}(\boldsymbol{x_i^0}), y_i^{0a}(\boldsymbol{x_i^0}), i \in \mathscr{N}).$$

### 2.2 Model

The main idea of the EKR model is using an analytical method to help kernel regression to estimate the unknown function $y(\boldsymbol{x})$ at the unobserved point $\boldsymbol{x}$.

KR is a non-parametric estimation technique that does not assume any distribution to fit the data. The estimation at an unobserved point is built by calculating the weighted average of the simulation outputs of all the initial design points. The weight of each design point depends on the distance between the point we are interested in (i.e., $\boldsymbol{x}$) and the observation itself (i.e., $\boldsymbol{x_i^0}$) using a spatial correlation function (generally a radial basis function). In EKR model, KR technique is used to average the estimates provided by each initial design point with Nadaraya-Watson estimator (Wand and Jones 1995):

$$\hat{y}_{EK}(\boldsymbol{x}) = \frac{\sum_{i\in\mathscr{N}} K_1(\boldsymbol{x_i^0}-\boldsymbol{x})\hat{y}_i(\boldsymbol{x})}{\sum_{i\in\mathscr{N}} K_1(\boldsymbol{x_i^0}-\boldsymbol{x})} \tag{1}$$

where $K_1(\cdot)$ is a $d$-dimension kernel function and $\hat{y}_i(\boldsymbol{x})$ is the estimate at the unobserved point $\boldsymbol{x}$ provided by the design point $\boldsymbol{x_i^0}$. This work uses the widely applied *Gaussian Kernel* as $K_1(\cdot)$:

$$K_1(\boldsymbol{x_i^0}-\boldsymbol{x}) = \prod_{k\in\mathscr{K}} e^{-\frac{1}{2\theta_{1k}}(x_{ik}^0-x_k)^2} \tag{2}$$

where $\theta_{1k}$ are parameters selected according to some specific criteria. These parameters control the influence of the design points on the prediction. The weight given by *Gaussian Kernel* is large when the design point is close to point $\boldsymbol{x}$, and small for the design point far away. Therefore, it allows each design point to have an importance related to point $\boldsymbol{x}$. The closer the points, the larger their importance.

The knowledge about some structural properties of the function embedded in the analytical method can help to build the estimator, particularly when the point to be evaluated is far from initial design points and the spatial correlation is not likely to hold. The $\hat{y}_i(\boldsymbol{x})$ in equation (1) presents the predictions from simulation outputs combined with analytical method estimates by kernel function $K_2(\cdot)$:

$$\hat{y}_i(\boldsymbol{x}) = K_2(\boldsymbol{x_i^0}-\boldsymbol{x})y_i^{0s}(\boldsymbol{x_i^0}) + (1-K_2(\boldsymbol{x_i^0}-\boldsymbol{x}))\tilde{y}_i^a(\boldsymbol{x}), \forall i \in \mathscr{N} \tag{3}$$

where $y_i^{0s}(\boldsymbol{x_i^0})$ is the simulation output at point $\boldsymbol{x_i^0}$ and $\tilde{y}_i^a(\boldsymbol{x})$ is the response of the analytical method at point $\boldsymbol{x}$ adjusted with the bias evaluated at point $\boldsymbol{x_i^0}$:

$$\tilde{y}_i^a(\boldsymbol{x}) = y^a(\boldsymbol{x}) + y_i^{0s}(\boldsymbol{x_i^0}) - y_i^{0a}(\boldsymbol{x_i^0}), \forall i \in \mathscr{N}. \tag{4}$$

In equation (4), $y^a(\boldsymbol{x})$ is the output of the analytical method at the unknown point $\boldsymbol{x}$ whereas $y_i^{0s}(\boldsymbol{x_i^0}) - y_i^{0a}(\boldsymbol{x_i^0})$ is the bias evaluated at design point $\boldsymbol{x_i^0}$ (i.e., the deviation between the simulation output and the analytical method output at point $\boldsymbol{x_i^0}$). The *Gaussian Kernel* described in equation (2) is also used as $K_2(\cdot)$ with selected parameters $\theta_{2k}$:

$$K_2(\boldsymbol{x_i^0}-\boldsymbol{x}) = \prod_{k\in\mathscr{K}} e^{-\frac{1}{2\theta_{2k}}(x_{ik}^0-x_k)^2}. \tag{5}$$

The weight given by *Gaussian Kernel* is between 0 and 1. Hence, in equation (3), the estimator $\hat{y}_i(\boldsymbol{x})$ gives more importance to the simulation output $y_i^{0s}(\boldsymbol{x_i^0})$ when the design point $\boldsymbol{x_i^0}$ is close to $\boldsymbol{x}$, whereas more importance is given to the adjusted analytical method response $\tilde{y}_i^a(\boldsymbol{x})$ when the design point is far from $\boldsymbol{x}$.

For simplicity, we assume that each dimension of the system configuration is equally important for predicting the system performance. Thus, the parameters $\theta_{1k}$ and $\theta_{2k}$ have the same value along all the dimension $k$ (i.e., $\theta_{1k} = \theta_1, \theta_{2k} = \theta_2, \forall k \in \mathscr{K}$). All the variables of the system configuration (i.e., $x_k, x_{ik}, i \in \mathscr{N}, k \in \mathscr{K}$) are normalized into $[0,1]$.

Algorithm 1 describes in detail how to implement the EKR method.

**Algorithm 1**

**Step 1 Initialization**:
$\boldsymbol{x_i^0}, \forall i \in \mathscr{N} = \{1,\ldots,n\} \leftarrow$ *System configuration of initial design points (DOE)*
$\boldsymbol{x} \leftarrow$ *System configuration to evaluate*
$\theta_{1k} = \theta_1, \theta_{2k} = \theta_2, \forall k \in \mathscr{K} = \{1,\cdots,d\} \leftarrow$ *Assigned parameters values*
**Step 2 Execute DOE**:
**for** $i = 1,\ldots,n$ **do**
  Execute simulation at $\boldsymbol{x_i^0}$ and collect $y_i^{0s}(\boldsymbol{x_i^0})$
  Execute analytical model at $\boldsymbol{x_i^0}$ and collect $y_i^{0a}(\boldsymbol{x_i^0})$
**end for**
**Step 3 Estimation**:
Execute analytical model at $\boldsymbol{x}$ and collect $y^a(\boldsymbol{x})$
Normalize $\boldsymbol{x}, \boldsymbol{x_i^0}, \forall i \in \mathscr{N}$ into $[0,1]^d$
**for** $i = 1,\ldots,n$ **do**
  Calculate $\tilde{y}_i^a(\boldsymbol{x})$ using equation (4)
  Calculate $K_2(\boldsymbol{x_i^0} - \boldsymbol{x})$ using equation (5)
  Calculate $\hat{y}_i(\boldsymbol{x})$ using equation (3)
  Calculate $K_1(\boldsymbol{x_i^0} - \boldsymbol{x})$ using equation (2)
**end for**
Calculate $\hat{y}_{EK}(\boldsymbol{x})$ using equation (1)

## 3 CONFIDENCE BAND

Confidence band is used to present the uncertainty in an estimate of a function based on limited data. Two approximate methods to calculate the confidence band around the EKR estimates are presented in this section. The first method, derived from Wasserman (2006), assumes that the normalized bias between the estimates $\hat{y}_{EK}(\boldsymbol{x})$ and expected estimates $\mathbb{E}[\hat{y}_{EK}(\boldsymbol{x})]$ is a Gaussian process. The second method is derived from basic statistic knowledge by assuming that the estimates $\hat{y}_i(\boldsymbol{x}), i \in \mathscr{N}$, follow a normal distribution.

### 3.1 Method A

We consider the $\hat{y}_i(\boldsymbol{x})$ in equation (1) as the observation of each design point, then the EKR model can be seen as a smooth regression model:

$$\hat{y}_{EK}(\boldsymbol{x}) = \sum_{i \in \mathscr{N}} w_i(\boldsymbol{x}) \hat{y}_i(\boldsymbol{x})$$

where $\boldsymbol{x}$ is the point to be evaluated, $\hat{y}_i(\boldsymbol{x})$ is the evaluation at point $\boldsymbol{x}$ calculated by equation (3) and the form of $w_i(\boldsymbol{x})$ is presented here:

$$w_i(\boldsymbol{x}) = \frac{K_1(\boldsymbol{x_i^0} - \boldsymbol{x})}{\sum_{i \in \mathscr{N}} K_1(\boldsymbol{x_i^0} - \boldsymbol{x})}, \forall i \in \mathscr{N}. \tag{6}$$

Let $W(\boldsymbol{x})$ denote the normalized bias between the estimates $\hat{y}_{EK}(\boldsymbol{x})$ and expected estimates $\mathbb{E}[\hat{y}_{EK}(\boldsymbol{x})]$:

$$W(\boldsymbol{x}) = \frac{\hat{y}_{EK}(\boldsymbol{x}) - \mathbb{E}[\hat{y}_{EK}(\boldsymbol{x})]}{\sigma(\boldsymbol{x}) \|\boldsymbol{w}(\boldsymbol{x})\|}$$

where $\sigma(\boldsymbol{x})$ is the standard deviation of the bias and $\boldsymbol{w}(\boldsymbol{x})$ is the weight vector $(w_1(\boldsymbol{x}), w_2(\boldsymbol{x}), \cdots, w_n(\boldsymbol{x}))^{\mathrm{T}}$. Assuming $W(\boldsymbol{x})$ is a Gaussian process, then an approximate $1-\alpha$ confidence band for $\hat{y}_{EK}(\boldsymbol{x})$ can be calculated as follows by extending Wasserman (2006) method to the multi-dimension case:

$$\hat{y}_{EK}(\boldsymbol{x}) \pm c \hat{\sigma}(\boldsymbol{x}) \|\boldsymbol{w}(\boldsymbol{x})\| \tag{7}$$

where $\hat{\sigma}(\boldsymbol{x})$ is estimated by calculating the error of estimates at design points through a regression method as follows. Let $\hat{\sigma}^2(\boldsymbol{x}) = e^{\hat{q}(\boldsymbol{x})}$, where $\hat{q}(\boldsymbol{x})$ is an estimate of $\log \sigma^2(\boldsymbol{x})$ and can be calculated by regressing $Z_i$ with function $K_1(\cdot)$:

$$\hat{q}(\boldsymbol{x}) = \frac{\sum_{i\in\mathcal{N}} K_1(\boldsymbol{x_i^0} - \boldsymbol{x}) Z_i}{\sum_{i\in\mathcal{N}} K_1(\boldsymbol{x_i^0} - \boldsymbol{x})}$$

where $Z_i = \log(y_i^{0s}(\boldsymbol{x_i^0}) - \hat{y}_{EK}(\boldsymbol{x_i^0}))^2$. The parameter $c$ in equation (7) can be calculated by solving

$$\mathbb{P}(|T_m| > c) + \frac{k_0}{\pi}\left(1 + \frac{c^2}{m}\right)^{-m/2} = \alpha$$

where $T_m$ is $t$-distributed with $m = n - \mathrm{tr}(\boldsymbol{L})$ degrees of freedom. $\boldsymbol{L}$ is an $n \times n$ matrix whose $i$-th row is $\boldsymbol{w}(\boldsymbol{x_i^0})^{\mathrm{T}}$. An approximation for $k_0$ is given by:

$$k_0 \approx \frac{1}{\sqrt{\theta_1}} \cdot \frac{\|\boldsymbol{K_1'}\|}{\|\boldsymbol{K_1}\|}$$

where $\boldsymbol{K_1}$ is a vector containing the kernel value of each design point $\boldsymbol{x_i^0}$ and $\boldsymbol{K_1'}$ is the sum of the partial derivatives of $\boldsymbol{K_1}$ calculated as $\frac{\partial \boldsymbol{K_1}}{\partial x_1} + \frac{\partial \boldsymbol{K_1}}{\partial x_2} + \cdots + \frac{\partial \boldsymbol{K_1}}{\partial x_d}$.

### 3.2 Method B

Method A is complicated and time consuming, especially when the DOE size is large. The main reason is the regression used for $\sigma(\boldsymbol{x})$'s estimation and the use of derivative. Therefore, a simple but approximate method is provided here for fast calculation of the confidence band.

In this method, we assume that the estimators $\hat{y}_i(\boldsymbol{x}), i \in \mathcal{N}$, calculated by equation (3) are independent and follow the normal distribution $N(y(\boldsymbol{x}), \sigma^2(\boldsymbol{x}))$, where $y(\boldsymbol{x})$ is the true value of the system performance at the new point $\boldsymbol{x}$ and $\sigma^2(\boldsymbol{x})$ is an unknown function of $\boldsymbol{x}$. Under this assumption, we can write:

$$\frac{\hat{y}_{EK}(\boldsymbol{x}) - y(\boldsymbol{x})}{S\sqrt{\sum_{i\in\mathcal{N}} w_i(\boldsymbol{x})^2}} \sim t(n-1)$$

where $S$ is the sample standard deviation of $\hat{y}_i(\boldsymbol{x})$, $w_i(\boldsymbol{x})$ are the weights defined in equation (6) and $t(n-1)$ is the $t$-distribution with $n-1$ degrees of freedom. Thus, we can calculate the confidence band by using the quantile of the $t$-distribution:

$$\mathbb{P}\left(\hat{y}_{EK}(\boldsymbol{x}) - t_{\frac{\alpha}{2}}(n-1)S\sqrt{\sum_{i\in\mathcal{N}} w_i(\boldsymbol{x})^2} < y(\boldsymbol{x}) < \hat{y}_{EK}(\boldsymbol{x}) + t_{\frac{\alpha}{2}}(n-1)S\sqrt{\sum_{i\in\mathcal{N}} w_i(\boldsymbol{x})^2}\right) = 1 - \alpha$$

where $1-\alpha$ is the confidence level and $t_{\frac{\alpha}{2}}(n-1)$ is the upper quantile of the $t$-distribution.

In order to obtain a more accurate confidence band, a weighted sample standard deviation $S_w$, rather than the sample standard deviation $S$, is used as the estimate for the standard deviation of $\hat{y}_i(\boldsymbol{x})$:

$$\hat{\sigma} = S_w = \sqrt{\sum_{i\in\mathcal{N}} w_i(\boldsymbol{x})(\hat{y}_i(\boldsymbol{x}) - \hat{y}_{EK}(\boldsymbol{x}))^2}.$$

The confidence band has the following form:

$$\hat{y}_{EK}(\boldsymbol{x}) \pm t_{\frac{\alpha}{2}}(n-1)S_w\sqrt{\sum_{i\in\mathcal{N}} w_i(\boldsymbol{x})^2}.$$

## 4 A SIMPLE EXAMPLE: $G/G/1$ SYSTEM

### 4.1 System Description

In this section, the EKR model is applied to a $G/G/1$ system in which inter-arrival times and service times are independently and normally distributed. The performance measure to be evaluated is the mean waiting time in the queue. This measure is denoted with $W_q$.

### 4.2 Experimental Design

The system configuration consists of the mean inter-arrival time $T_a$ and the mean service time $T_s$ (i.e., $\boldsymbol{x} = (T_a, T_s)$). The domain of interest is $T_a \in [10, 15]$ and $T_s \in [5, 10]$. The standard deviation of the inter-arrival time and service time are all equal to 3.

Kingman's formula calculates the mean waiting time in the queue of the $G/G/1$ system as follows (Hopp and Spearman 2011):

$$y^a(T_a, T_s) \approx \frac{(c_a^2 + c_s^2)}{2} \cdot \frac{u}{1-u} \cdot T_s$$

where $u = T_s/T_a$ and $c_a, c_s$ are the coefficients of variation of the inter-arrival time and of the service time, respectively.

Several DOEs with different sizes (from 10 points to 500 points) are developed. The simulations are executed in Arena with $100,000$ warm-up time and $1,000,000$ simulation length. EKR models are built using the simulation outputs at each point of the DOEs and the outputs of the Kingman's formula.

The parameters $\theta_1, \theta_2$ are selected as follows. First, we construct a DOE with 10 new points $(\boldsymbol{x_i'}, i = 1, ..., 10)$ and collect the simulation outputs $y^s(\boldsymbol{x_i'})$. Second, a complete factorial design with two factors $(\theta_1, \theta_2)$ and nine levels $(0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1)$ is developed. Third, the performance $W_q$ at the 10 new points are estimated using the values of $\theta_1, \theta_2$ from the complete factorial design (i.e., $\hat{y}_{EK}(\boldsymbol{x_i'}|\theta_1, \theta_2)$). Finally, we select the parameters $\theta_1, \theta_2$ that minimize the Root Mean Square Error (RMSE) of the 10 new points:

$$RMSE(\theta_1, \theta_2) = \sqrt{\frac{1}{10}\sum_{i=1}^{10}(\hat{y}_{EK}(\boldsymbol{x_i'}|\theta_1, \theta_2) - y^s(\boldsymbol{x_i'}))^2}.$$

The proposed method is compared with SK and KR methods. To evaluate the prediction performance of the models, $R = 1000$ checkpoints $\boldsymbol{x_i}$ are sampled. The Mean Absolute Relative Error (MARE) is used to assess the prediction performances of these models:

$$MARE = \frac{1}{R}\sum_{i=1}^{R}\frac{|\hat{y}(\boldsymbol{x_i}) - y_i^s(\boldsymbol{x_i})|}{y_i^s(\boldsymbol{x_i})}$$

where $\hat{y}(\boldsymbol{x_i})$ and $y_i^s(\boldsymbol{x_i})$ are the outputs of the regression models (EKR or KR or SK) and the simulation outputs at the 1000 checkpoints $\boldsymbol{x_i}$, respectively. MARE may be not useful when the response is very small. In this case, even small absolute deviations leads to large MARE. Therefore, another useful measure of the prediction accuracy is the RMSE:

$$RMSE = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{y}(\boldsymbol{x_i}) - y_i^s(\boldsymbol{x_i}))^2}.$$

The points of all DOEs as well as the checkpoints are sampled by Latin Hypercube Sampling (LHS) (Helton and Davis 2003).

### 4.3 Numerical Results

Figure 1 shows the *MARE* (left) and the *RMSE* (right) of the three models with different DOE sizes (i.e., the points we use to build regression models). The relative error decreases with the increase of the DOE size for all the three models. When the DOE size is small, EKR model performs best while KR model has the largest relative error. For large DOE size, the relative errors of EKR, KR and SK models are close.

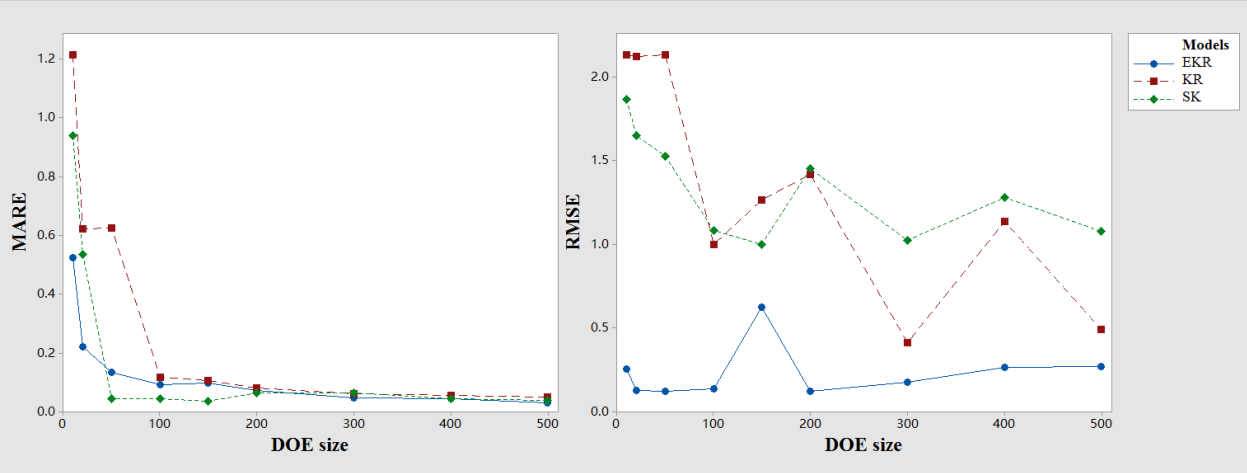


Figure 1: The *MARE* (left) and the *RMSE* (right) of the estimates provided by EKR, KR and SK models at the 1000 checkpoints with DOE size from 10 points to 500 points.

The EKR model always performs better than the other two methods according to *RMSE*. The reason is that most of the initial design points have small simulation outputs. Indeed, as shown in Figure 2, SK and KR have large errors when the simulation output is large, whereas EKR model is very robust and follows the true path due to the contribution of the Kingman's formula. That means *RMSE* is more representative than *MARE* in this case.

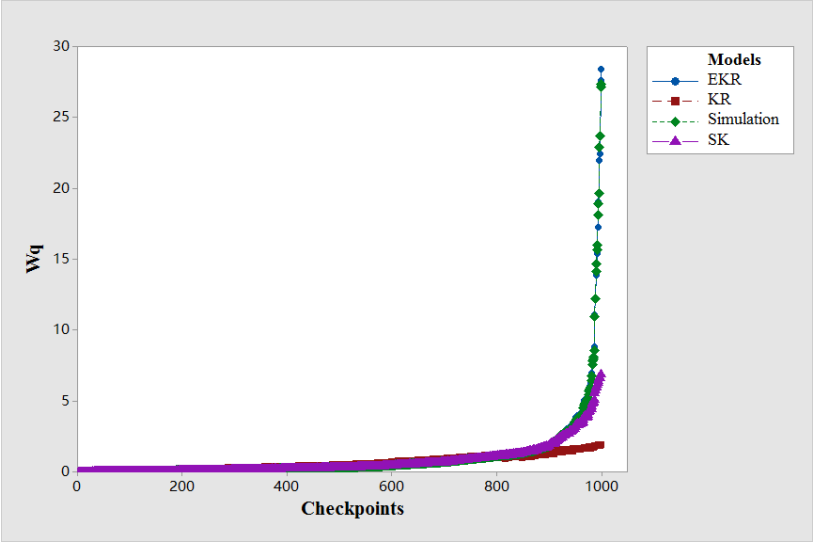


Figure 2: The simulation outputs of the 1000 checkpoints as well as their estimates provided by KR, SK and EKR models with DOE size equal to 50 points. The checkpoints are sorted by simulation outputs.

The experiment is repeated 50 times with DOE size equal to 20 points for robustness analysis. Figure 3 shows boxplots of the *MARE* (left) and the *RMSE* (right) of the three models. Variability of *MARE* for SK is larger than EKR and KR, whereas for *RMSE* the variabilities of these three models are quite close.

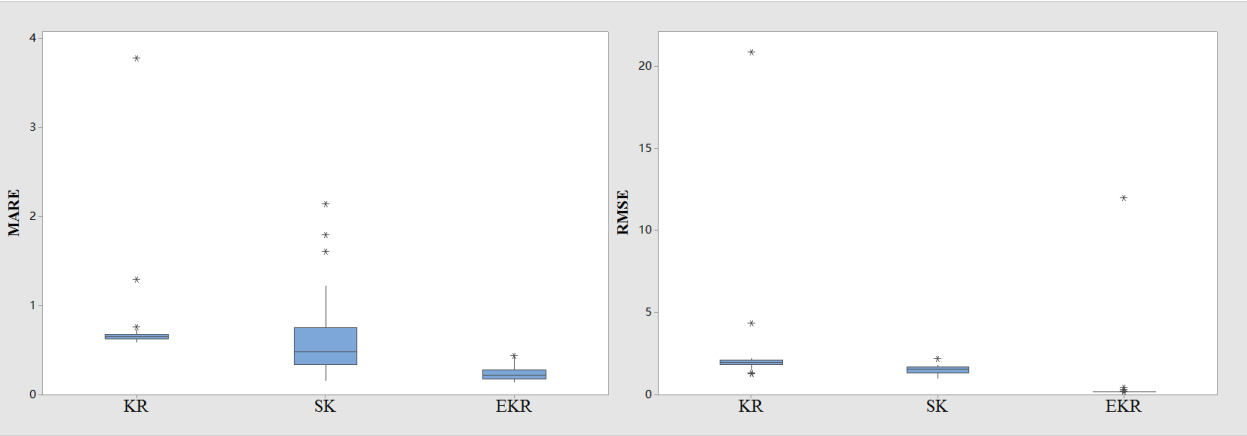


Figure 3: The *MARE* (left) and the *RMSE* (right) of the estimates provided by KR, SK and EKR at the 1000 checkpoints with DOE size equal to 20 points. 50 replicates are executed.

Method B provides a much faster calculation of the confidence band than method A. It takes about 0.13 seconds to calculate the confidence bands among 1000 checkpoints with DOE size equal to 20 points, while method A needs about 21 seconds. Table 1 and Table 2 present the proportion of simulation outputs within the confidence band calculated by the method A and method B among the 1000 checkpoints, respectively, of different values for confidence level. The confidence band calculated by method A is too large. Indeed, almost all the simulation data fall into the confidence band regardless of the confidence level. As far as using method B, the proportion is a little smaller than the confidence level. This is probably caused by the assumption of independence among $\hat{y}_i(\boldsymbol{x}), i \in \mathcal{N}$. Nevertheless, the proportion becomes stable when the DOE size increases.

Table 1: The proportion of simulation outputs within the confidence band provided by method A(%).

| DOE size | | 10 | 20 | 50 | 100 | 150 | 200 | 300 | 400 | 500 |
|---|---|---|---|---|---|---|---|---|---|---|
| $1-\alpha$ | 80% | 99.8 | 99.3 | 98.8 | 92.5 | 95.7 | 99.3 | 98.7 | 99.8 | 99.0 |
| | 90% | 99.9 | 99.4 | 99.2 | 93.3 | 96.2 | 99.4 | 98.8 | 99.8 | 99.0 |
| | 95% | 99.9 | 99.4 | 99.3 | 93.5 | 96.7 | 99.5 | 98.9 | 99.8 | 99.2 |

Table 2: The proportion of simulation outputs within the confidence band provided by method B(%).

| DOE size | | 10 | 20 | 50 | 100 | 150 | 200 | 300 | 400 | 500 |
|---|---|---|---|---|---|---|---|---|---|---|
| $1-\alpha$ | 80% | 49.5 | 75.0 | 75.7 | 50.9 | 69.4 | 62.4 | 68.6 | 69.8 | 70.9 |
| | 90% | 57.4 | 83.0 | 82.2 | 60.3 | 81.5 | 73.2 | 77.2 | 79.2 | 79.7 |
| | 95% | 63.2 | 87.6 | 86.5 | 65.2 | 88.0 | 79.5 | 82.7 | 84.9 | 85.9 |

Figure 4 presents the confidence band by showing all the checkpoints with service time $T_s \in [7.4, 7.6]$. The horizontal axis is the inter-arrival time $T_a$. The estimate of each point is calculated by EKR model with 20 points in the DOE. The confidence band is calculated by method B with 95% confidence level.

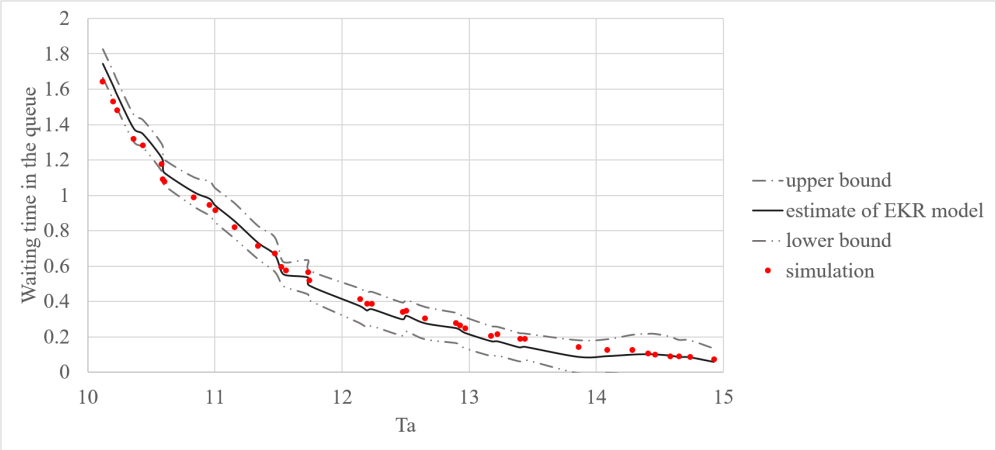


Figure 4: The simulation outputs, EKR estimates and confidence band at the checkpoints with service time $T_s \in [7.4, 7.6]$. The confidence band is provided by method B with confidence level 95%. The EKR estimates are calculated with 20 points in the DOE.

## 5 A REALISTIC EXAMPLE: A CLOSED-LOOP FLEXIBLE ASSEMBLY SYSTEM

### 5.1 System Description

The closed-loop flexible assembly system (CLFAS) described in Suri and Leung (1987) is considered in this section. The same system was studied by Chen et al. (2013). In the analyzed system, there are several workstations connected together in a loop. The buffer capacities between adjacent workstations are limited. The first workstation is for loading and unloading of workpieces on pallets. All the workpieces enter or leave the system through the load/unload station. The workpieces are assembled on the pallets. The number of pallets in the whole system is fixed. It is assumed there is a large number of workpieces waiting outside system. Once a finished workpiece leaves the system, a new workpiece will enter the system to be processed.

As shown in Figure 5, there are six workstations in the system and only one buffer slot between adjacent workstations. The number of the pallets in the system is also six, which means there will be always six workpieces in the system.

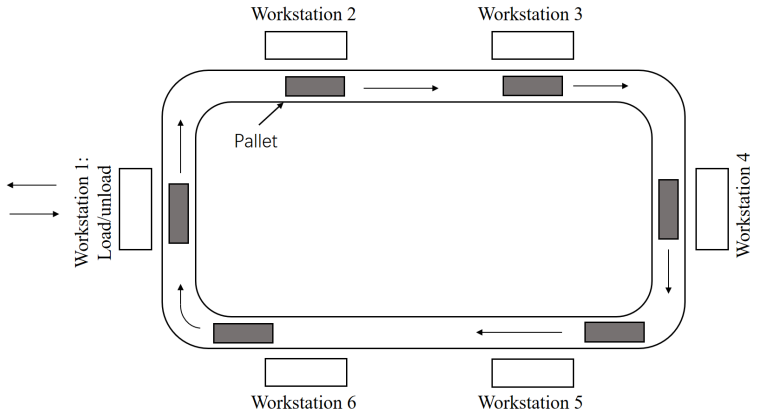


Figure 5: Closed-Loop Flexible Assembly System with six workstations and six pallets.

The transfer time between two workstations is assumed negligible. The operation time at each workstation $T_r$ (with $r = 1, 2, ..., 6$) consists of two parts: a deterministic machine cycle time $x_r$ and an

additional random time $R_r$ modeling machine jams. Let $R_r$ be $i.i.d.$ and follow the uniform distribution $U(0.1,1.1)$. The probability that the machine $r$ will jam is $\alpha_r = 0.005$. Therefore, the operation time $T_r = x_r + I_r\{machine\ r\ jams\}R_r$, where $I_r\{\cdot\}$ is the indicator function. The total assembly time equals to the total operation time $\sum_{r=1}^{6} T_r$. We want to evaluate the mean throughput of the system.

### 5.2 Experimental Design

The system configuration is the machine cycle time $\boldsymbol{x} = (x_1, x_2, ..., x_6)$ and the domain of interest is $[0.05, 0.15]^6$. Two kinds of analytical methods are applied to this case. In the first case a Continuous Time Markov Chain (CTMC), which assumes that the operation time at each workstation $T_r$ follows an exponential distribution, is developed. In the second case a Closed Queuing Network (CQN) Jackson model (Tempelmeier and Kuhn 1993), which further assumes that the buffer capacity is infinite, is developed.

LHS and complete factorial design are used as in section 4 for sampling of design points and selection of parameters, respectively. The simulation outputs are collected by simulating $100,000$ workpieces with $50,000$ workpieces of warm-up period.

Two designs are developed (17 and 25 points in the DOE) and 1000 checkpoints are generated for evaluation of the prediction performance of the models by comparing the two synthetic indicators *MARE* and *RMSE*. The experiment is replicated 50 times.

### 5.3 Numerical Results

Figure 6 shows the results of different models. The results of *MARE* and *RMSE* are similar. The estimates have lower variability with 25 points in the DOE compared to that with 17 points for all the models. The performance of the SK model fluctuates widely while EKR model performs better (both for mean error and variability). The use of the CTMC slightly improves the prediction performance of the EKR model compared to that with CQN, although the CTMC has larger error. This indicates that the smoothness of the bias of the analytical method is more important than the mean error of that in EKR model. However the margin is not significant. The reason is due to the fact that only six pallets circulate in the system, thus the buffer capacity is not critical.

Table 3 shows the proportion of simulation outputs within the confidence band calculated by different methods for different confidence levels with 17 and 25 points in the DOE. The EKR models with CTMC and CQN have similar results, here only the results of the EKR model with CTMC are presented. Differently from the previous case, the confidence band calculated by the method A is more reliable than method B.

Table 3: The proportion of simulation outputs within the confidence band.(%).

| | 17 points | | 25 points | |
|---|---|---|---|---|
| $1-\alpha$ | Method A | Method B | Method A | Method B |
| 80% | 77.1 | 41.5 | 69.2 | 39.6 |
| 90% | 83.5 | 51.6 | 74.7 | 49.4 |
| 95% | 88.7 | 59.5 | 79.3 | 57.2 |

## 6 CONCLUSION

In this paper, the Extended Kernel Regression model is presented in detail and tested numerically on a $G/G/1$ system and a Closed-Loop Flexible Assembly System. The proposed method is compared with SK and KR methods and results show that EKR provides better predictions in the analyzed cases. Based on the numerical analysis reported in this paper, the KR and SK models provide good estimate when the system performance is smooth or the DOE size is large and EKR model can improve the performance of the regression technique with the involvement of the analytical method.

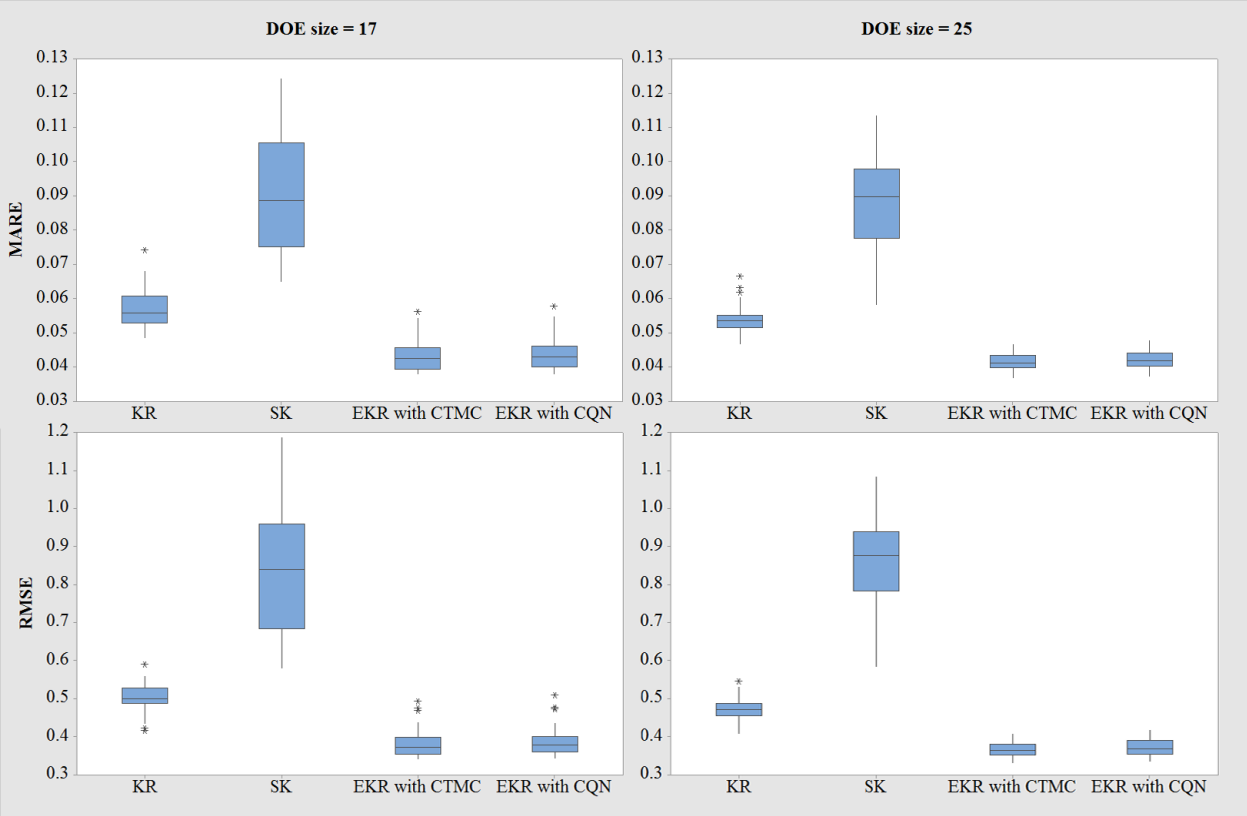


Figure 6: The *MARE* (above) and the *RMSE* (below) of the estimates provided by KR, SK, EKR (with CTMC and CQN) at the 1000 checkpoints with DOE size equals to 17 (left) and 25 (right) points.

Besides, two methods for calculation of confidence band are proposed and assessed numerically. Method B performs better than method A in the $G/G/1$ system. While for CLFAS, the confidence band calculated by the method A is more reliable than that calculated by method B. Future research will be devoted into different directions. First, more effort will be put in to improve the calculation of confidence band. Second, the EKR method will be compared with other techniques (e.g., Co-Kriging). Third, more than one analytical methods with different fidelities will be considered in the EKR model.

## REFERENCES

Ankenman, B., B. L. Nelson, and J. Staum. 2010. "Stochastic Kriging for Simulation Metamodeling". *Operations Research* 58 (2): 371–382.

Askin, R. G., and C. R. Standridge. 1993. *Modeling and Analysis of Manufacturing Systems*. Wiley.

Buzacott, J. A., and J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Prentice Hall.

Chen, R., J. Xu, S. Zhang, C.-H. Chen, and L. H. Lee. 2015. "An Effective Learning Procedure for Multi-Fidelity Simulation Optimization with Ordinal Transformation". In *Automation Science and Engineering (CASE), 2015 IEEE International Conference*, 702–707. IEEE.

Chen, X., B. E. Ankenman, and B. L. Nelson. 2013. "Enhancing Stochastic Kriging Metamodels with Gradient Estimators". *Operations Research* 61 (2): 512–528.

Cressie, N. 1993. *Statistics for Spatial Data*. Wiley.

Forrester, A. I., and A. J. Keane. 2009. "Recent Advances in Aurrogate-Based Optimization". *Progress in Aerospace Sciences* 45 (1): 50–79.

Gershwin, S. B. 1994. *Manufacturing Systems Engineering*. Prentice Hall.

Goh, J., D. Bingham, J. P. Holloway, M. J. Grosskopf, C. C. Kuranz, and E. Rutter. 2013. "Prediction and Computer Model Calibration Using Outputs from Multifidelity Simulators". *Technometrics* 55 (4): 501–512.

Helton, J. C., and F. J. Davis. 2003. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems". *Reliability Engineering & System Safety* 81 (1): 23–69.

Hopp, W. J., and M. L. Spearman. 2011. *Factory Physics*. Waveland Press.

Li, J., and S. M. Meerkov. 2008. *Production Systems Engineering*. Springer.

Matta, A., N. Li, Z. Lin, and J. G. Shanthikumar. 2015, June 01-06. "Operational Learning of Approximate Analytical Methods for Performance Evaluation of Manufacturing Systems". In *10th Conference on Stochastic Models of Manufacturing and Service Operations SMMSO 2011*. Volos (Greece).

Papadopoulos, C. T., M. E. O'Kelly, M. J. Vidalis, and D. Spinellis. 2009. *Analysis and Design of Discrete Part Production Lines*. Springer.

Suri, R., and Y. T. Leung. 1987. "Single Run Optimization of a SIMAN Model for Closed Loop Flexible Assembly Systems". In *Proceedings of the 1987 Winter Simulation Conference*, edited by H. Grant, W. D. Kelton, and A. Thesen, 738–748. Atlanta, Georgia: Association for Computing Machinery.

Tempelmeier, H., and H. Kuhn. 1993. *Flexible Manufacturing Systems: Decision Support for Design and Operation*. Wiley.

Wand, M. P., and M. C. Jones. 1995. *Kernel Smoothing*. Monographs on Statistics & Applied Probability. Chapman and Hall/CRC.

Wasserman, L. 2006. *All of Nonparametric Statistics*. Springer.

## AUTHOR BIOGRAPHIES

**ZIWEI LIN** is a Ph.D student in Industrial Engineering & Management at Shanghai Jiao Tong University. Her research interest is performance evaluation of stochastic discrete event systems. Her email address is linziwei@sjtu.edu.cn.

**ANDREA MATTA** is Distinguished Professor at the Department of Industrial Engineering & Management at Shanghai Jiao Tong University. He currently teaches stochastic models and simulation. His research area includes analysis and design of manufacturing and healthcare systems. His email address is matta@sjtu.edu.cn.

**NA LI** is an Associate Professor in the Department of Industrial Engineering, Shanghai Jiao Tong University. Her current research interests are stochastic system modeling and analysis, both for production system and service system. Her email address is na-li03@sjtu.edu.cn

**J. GEORGE SHANTHIKUMAR** is Distinguished Professor of Management at Purdue University. His research interests are in integrated inter-disciplinary decision making, model uncertainty and learning, production systems modeling and analysis. His email address is shanthikumar@purdue.edu