APEM journal

Advances in Production Engineering & Management
Volume 12 | Number 3 | September 2017 | pp 285–295
https://doi.org/10.14743/apem2017.3.259

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Predicting the availability of production lines by combining simulation and surrogate model

Jia, Y.[a], **Tian, H.**[a], **Chen, C.**[a,*], **Wang, L.**[a,b]

[a]School of Mechanical Science and Engineering, Jilin University, Changchun, P.R. China
[b]School of Mechanical Engineering, Dalian University of Technology, Dalian, P.R. China

A B S T R A C T

The availability analysis plays a significant part in both the design and operations management of production lines. In this paper, a method combining discrete event simulation (DES) and surrogate model is presented to predict the availability of production lines with unreliable workstations and finite intermediate buffers. The DES can conduct computer experiments for production lines with the help of design of experiments (DOE) under the Matlab environment. The surrogate model is constructed by using Kriging model integrated with Latin hypercube sampling (LHS), which can predict the responses based on a limited set of simulation results. The major advantages of the proposed approach are its flexibility and convenience. Also, it is the first time to investigate Kriging opportunities in predicting the performance of production lines. Finally, an application in a crankshaft production line is presented, and the results indicate that the proposed approach can achieve higher prediction accuracy than the other methods.

A R T I C L E I N F O

Keywords:
Production lines
Availability prediction
Discrete event simulation (DES)
Kriging model
Latin hypercube sampling (LHS)

**Corresponding author:*
cchchina@jlu.edu.cn
(Chen, C.)

Article history:
Received 5 May 2017
Revised 22 June 2017
Accepted 10 July 2017

1. Introduction

A production line, also known as a transfer line or a flow line, is one of the most important and common types of manufacturing systems employed for high-volume low-variety production of industrial components. Unlike flexible manufacturing systems (FMS) or manufacturing cells (FMC), production lines play a significant role in processing the main products of a plant, and usually require high capital investment. They are often organized with a predetermined sequence of equipment [1] and intermediate buffers arranged in a serial structure and connected by a material handling system. Because of high sensibility to failures, the improvement of production lines' availability is an important issue for the designers or operators to resolve. It is therefore necessary to research the methods for evaluating or predicting the availability of production lines.

A review focusing on availability analysis techniques of production lines is given as follows, which can be divided into three groups: exact analytical methods, approximate analytical methods and simulation. Exact analytical methods, which are generally on the basis of queueing models and Markov chain [2], can obtain the exact solutions of steady-state probability, thus providing insight into the qualitative performance of production lines. But they are only suitable for small lines (no more than three-stage) because of the state explosion problem.

Based on the two-stage exact models, approximate analytical methods are developed for lines with more machines. The most representative methods are decomposition and aggregation approaches. The common intention of decomposition methods is to decompose an N-machine system into a cluster of N-1 subsystems which contain two pseudo-machines and one original buffer, and to get the result by solving simultaneous equations [3]. Gershwin [4] developed an efficient decomposition method for synchronous lines and it can get accurate results with the help of the Dallery-David-Xie (DDX) algorithm [5]. Afterwards, Burman [6] modified the continuous model for asynchronous lines and presented a new accelerate DDX (ADDX) algorithm. Other extensions of decomposition method can be found in [7-9]. Compared with decomposition methods, aggregation methods are more straightforward and simpler. The general idea is to aggregate the original line into one unique equivalent machine by iteratively replacing a two-machine one-buffer subline [10]. Meerkov and his group have obtained some achievements on aggregation methods. Detailed descriptions are summarized in [11], and a corresponding software called PSE Toolbox is developed. Although approximate analytical methods have been fully investigated, there is a limitation for wide application: all the distributions have to be limited to special forms, such as geometric or exponential.

Compared with analytical approaches, simulation can build the models of production lines at any requested level of detail [12] without being restricted by assumptions such as specified distributions. Consequently, discrete event simulation (DES) has been proved to be the ideal tool for exhibiting the dynamics of complex manufacturing process, and meanwhile, the accuracy can be controlled. Because of the applicability and practicability, DES has been widely used for predicting and optimizing the performance of production lines [13-15]. Furthermore, some authors [16-18] investigated the real production lines by means of case studies. A comprehensive discussion of many important aspects of discrete event simulation is given by Law [19] from fundamentals to applications. In addition, many commercial software packages (e.g., Flexsim, Witness, Plant Simulation, Arena, etc.) have been designed specifically to simulate manufacturing systems, thus increasing the popularity of simulation in recent years. Despite the modelling flexibility and great ease of use, DES is usually time-consuming, particularly at the initial design stage when lots of system parameters are indeterminate. Although high-performance computers are developed, a lot of computing time and resources are still necessary to obtain statistically significant results.

To overcome the limitations of the above methods, an integrated simulation-surrogate model methodology is presented to predict the availability of production lines in this paper. For reasons of generality, our research is limited to the analysis of discrete serial-parallel lines with unreliable workstations and finite intermediate buffers. The rest of the article is organized as follows. Production line description, assumptions and symbols are described in Section 2. Section 3 proposes a general DES to simulate the production process and obtain the production lines' availability. Section 4 outlines a surrogate model combined with LHS and Kriging model for prediction. An application in a rough machining production line for crankshaft is presented in Section 5. Finally, conclusions and prospects close the paper in Section 6.

2. Production line model, assumptions and symbols

2.1 Production line description

A discrete production line is often organized with workstations connected in product-flow layout and separated by intermediate buffers. A workstation may be composed of several machines in series/parallel or just one machine (in this case the terms "workstation" and "machine" are used interchangeably). The graphical-based structural model of an N-workstation production line is given in Fig. 1, where workstations are represented by squares and buffers are represented by circles. In a production line, workpieces from outside enter the system from the first workstation WS_1. Each workpiece is processed by WS_1 within operation time T_1, after which it is transferred to the first buffer B_1. Then it moves in the direction of arrows until it is finished by the last workstation WS_N, and exits the system.

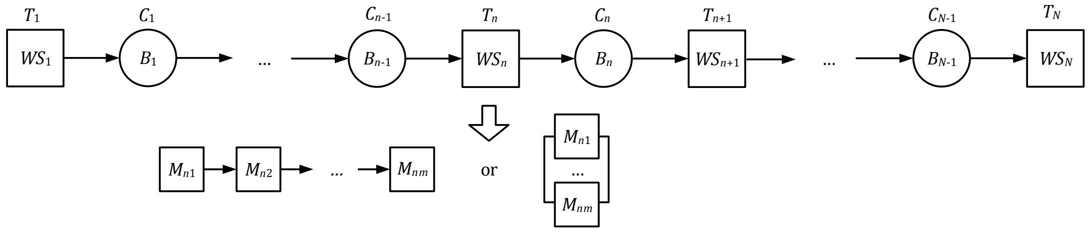

Fig. 1 The block diagram of an N-workstation production line

In the real production operations, the machines always experience random breakdowns. Consequently, failure of one workstation may affect all other workstations upstream and downstream. This complex phenomenon is generally regarded as perturbation propagation [11], which makes the analysis of the production line difficult. To limit the propagation of disruptions, buffers are usually placed between workstations. In fact, buffers are capable to provide continuity by means of saving parts from the upstream subsystem and releasing parts to the downstream subsystem. The buffer inventory can provide a period of isolation time [20] for maintenance actions before the buffer becomes empty or full without bringing down the entire system immediately. From this point of view, buffers alleviate this mutual interference by decoupling adjacent workstations from "rigid" connection to "elastic" connection.

Except the failures of machines, another reason for line inefficiency is the workstations' interference: starvation and blocking. A workstation is called *starved* when its upstream buffer is empty (buffer inventory is 0). It is said to be *blocked* when its downstream buffer is full (buffer inventory is its maximum capacity C_n). Taking WS_n as an example, starvation is the phenomenon that when WS_n has finished a workpiece it is forced to wait because B_{n-1} is empty. When a workstation is *up*, it is said to be *busy* when it is processing a workpiece, and is said to be *idle* when it is either starved or blocked. Generally, uptimes (including busy times and idle times) and downtimes exhibit statistical regularity, and can be expressed as independent and identically distributed random variables. Thus, the state is summarized as follows:

$$\text{the state of } WS_n = \begin{cases} up & \begin{cases} busy \\ idle & \begin{cases} starved \\ blocked \end{cases} \end{cases} \\ down \end{cases}$$

As mentioned above, the parallel machines are simplified to one workstation in this research (see Fig. 1). They are generally used to balance the production line, and have the same operation as well as configuration in most situations. They are therefore assumed to have identical operation time as well as parameter distributions. Thus the workstation's operation time equals to the machines' operation time divided by the number of parallel machines. Moreover, we defined "degradation ratio" as the production capacity coefficient of the workstation when one of the parallel machines is down. For example, the degradation ratio is 2/3 when the workstation contains three parallel machines. Without loss of generality, let it be 0 in the series case.

2.2 Assumptions

The following additional assumptions are also used:

(a) As the supply and storage of production line are beyond the scope of this research and they are considered to be infinite. In other words, WS_1 is never starved and WS_N is never blocked.
(b) Scheduled downtimes such as breaks, meetings, and preventative maintenance are not concerned in this paper.
(c) Operation time of each workstation which contains transfer time and setup time is constant because most gantry robots and machines are controlled by predetermined NC code.
(d) Failures don't destroy workpieces. Therefore, the workpieces remain at the machines during maintenance, and processing resumes when the machines are up.

(e) The two-parameter Weibull distribution is employed to model uptimes and downtimes. Although the proposed DES model is appropriate for any distributed workstations, the Weibull distribution is one of the most common distributions in reliability engineering and can be conveniently transformed into exponential distribution, which is widely used in analytical approaches.

2.3 Symbols of system parameters

It is necessary to present a summary of the symbols as well as their explanations used in this research. The system input and output parameters are listed in Table 1 and Table 2, respectively.

Table 1 The input parameters of model

Notations	Explanation of the notations
N	Total number of workstations
T_n	Operation time of WS_n, and $\boldsymbol{T} = (T_1, T_2, \cdots, T_N)^T$
C_n	Maximum capacity of B_n (including the space at WS_{n+1}), and $\boldsymbol{C} = (C_1, C_2, \cdots, C_{N-1})^T$
D_n	Degradation ratio of WS_n, and $\boldsymbol{D} = (D_1, D_2, \cdots, D_N)^T$, $D_n \in [0, 1)$
ut_{ni}	Uptime of WS_n before the i^{th} failure, and $UT_n = \{ut_{n1}, ut_{n2}, \cdots, ut_{ni}, \cdots\}$
dt_{ni}	Downtime of WS_n during the i^{th} failure, and $DT_n = \{dt_{n1}, dt_{n2}, \cdots, dt_{ni}, \cdots\}$
α_n	Scale parameter of Weibull distribution for WS_n, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_N)^T$
β_n	Shape parameter of Weibull distribution for WS_n, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_N)^T$
t_{sim}	Simulation time
t_{warmup}	Warmup period
r	Number of independent replications of the simulation

Table 2 The output parameters of model

Notations	Explanation of the notations
$k_n(t)$	Inventory level of B_n at time t, and $0 \leq k_n(t) \leq C_n$
$ct_n(t)$	Cumulative busy time of WS_n at time t
$op_n(t)$	Output of WS_n at time t
$s_n(t)$	State of WS_n at time t, and $s_n = D_n$ during downtime, $s_n = 1$ during busy time, s_n = "*starved*" or "*blocked*" during idle time
OP	Total amount of output, and $OP = op_N(t_{sim})$
$OPh(t)$	Hourly output of system at time t
$MTBF$	Mean time between failures, and $MTBF_n = \overline{ut}_n$
$MTTR$	Mean time to repair, and $MTTR_n = \overline{dt}_n$
A	Availability of the whole line, defined as the probability of system being processing. Actually, in the steady state, if no failures occur, the system can process a workpiece in every bottleneck operation time T_{max}, which is the maximum operation time of workstations. Thus, the total processing time of system is $OP \times T_{max}$, and $A = OP \times T_{max}/t_{sim}$.

3. Discrete event simulation

In this section a general DES is developed under Matlab environment to simulate the manufacturing process of production lines. The simulation model can provide real-time information on operating characteristics, and evaluate the availability of production lines with various input parameters. Meanwhile, it has a good compatibility with subsequent surrogate model programs.

3.1 Simulation process

The flow chart of the proposed DES is shown in Fig. 2 and the main steps are described as follows:

(a) Set the input parameters for the DES model: N, $\boldsymbol{T}$, $\boldsymbol{C}$, $\boldsymbol{D}$, t_{sim}, and t_{warmup}.
(b) Initialization of $k_n(0)$, $ct_n(0)$ and $op_n(0)$ for every buffer and workstation with the default values all zero.
(c) Generate sample sets UT_n and DT_n with required distribution by Monte Carlo technique for every workstation. Accumulate and sort these data in chronological order until t_{sim} ends.

(d) Scan the real-time state for every workstation and buffer at each time unit by means of a nested loop, of which the outer loop runs from t = 1 to t = t_{sim} and the inner loop runs from n = 1 to n = N. In this process, discriminate WS_n between up and down according to the samples obtained from step (c). If WS_n is up, discriminate WS_n between busy and idle according to $k_{n\text{-}1}(t\text{-}1)$ and $k_n(t\text{-}1)$. Then record $k_n(t)$, $ct_n(t)$, $op_n(t)$ and $s_n(t)$.

(e) Calculate OP, OPh and A at the end of simulation.

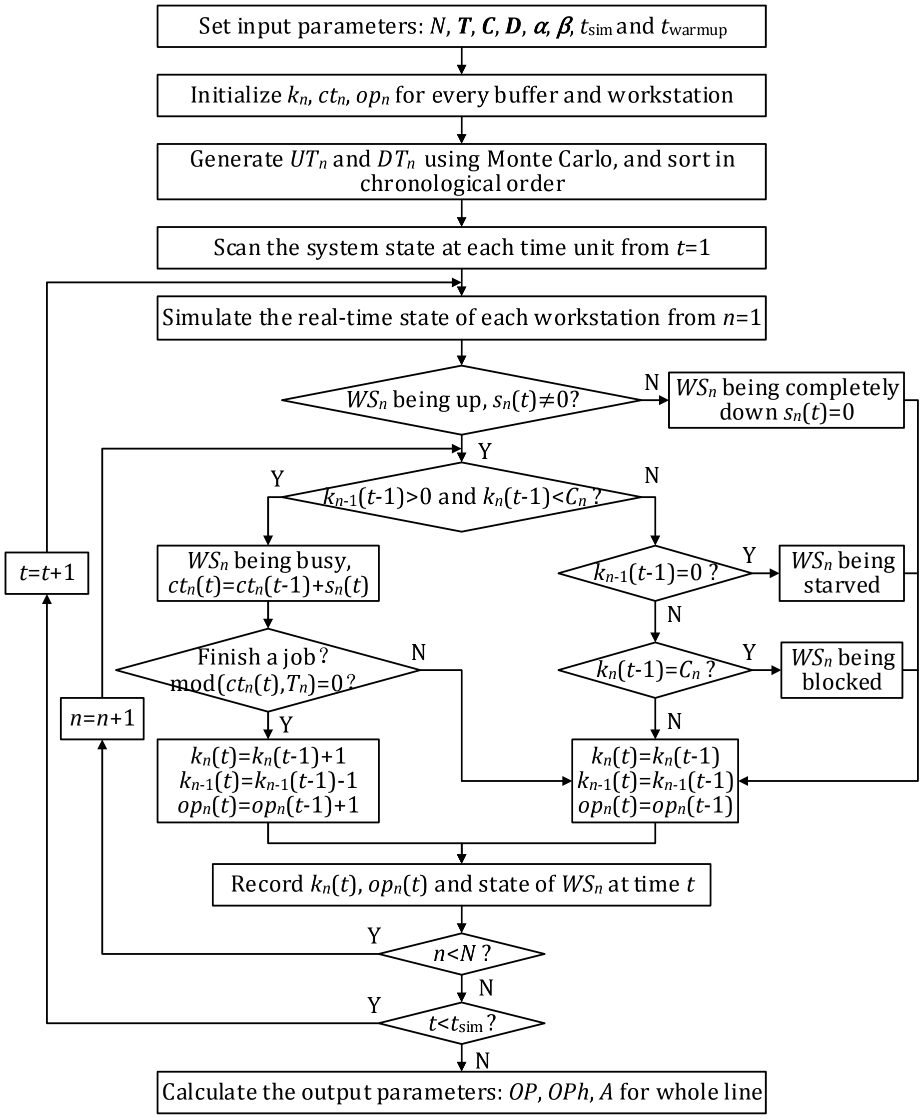

Fig. 2 The flow chart of DES model for production lines

3.2 Validation

The validation of proposed DES model was conducted by comparing the results with Plant Simulation software in different lines. We investigated 4 cases as follows, with the configuration details in Table 3:

Case 1: Synchronous series line with same workstations;
Case 2: Synchronous series line with different workstations;
Case 3: Asynchronous series-parallel line, also be described in Section 5;
Case 4: Asynchronous series-parallel line with different buffers and Weibull workstations.

We performed r = 100 independent repeated trials with each of simulation length t_{sim} = 1440 h (3 m × 30 d × 16 h). The average results, which are summarized in Table 4, indicate that the percentage errors of OP are extremely small. That means the proposed DES model is practicable.

Table 3 Configuration details of 4 cases

Parameters		Case 1	Case 2	Case 3	Case 4
N		5	5	5	5
T (s)		(200,200,200,200,200)	(200,200,200,200,200)	(180,200,190,180,190)	(180,200,190,180,190)
K		(10,10,10,10)	(10,10,10,10)	(10,10,10,10)	(10,15,10,15)
D		(0,0,0,0,0)	(0,0,0,0,0)	(0,0,0.5,0,0)	(0,0,0.5,0,0)
UT (h)	α	(400,400,400,400,400)	(400,600,350,300,450)	(400,600,350,300,450)	(400,600,350,300,450)
	β	(1,1,1,1,1)	(1,1,1,1,1)	(1,1,1,1,1)	(1.2,0.8,1.2,0.8,1.2)
DT (h)	α	(2,2,2,2,2)	(1.5,2,2.5,3,1.5)	(1.5,2,2.5,3,1.5)	(1.5,2,2.5,3,1.5)
	β	(1,1,1,1,1)	(1,1,1,1,1)	(1,1,1,1,1)	(0.8,1.2,0.8,1.2,0.8)

Table 4 Simulation results of Plant Simulation and the DES in proposed approach

Parameters	Case 1			Case 2			Case 3			Case 4		
	Matlab	PS	Error, %	Matlab	PS	Error, %	Matlab	PS	Error, %	Matlab	PS	Error, %
OP	25351	25369	-0.0710	25294	25301	-0.0277	25480	25486	-0.0235	25505	25511	-0.0235

3.3 Warm-up

The start-up or initial transient problem is a common problem in simulation process. In order to ensure that observations can represent steady-state behaviour, the warming up or initial-data deletion technique is often suggested. In this research, a graphical procedure proposed by Welch is employed to choose the warm-up period (see [19]). This procedure can smooth out the plot of observations based on r independent replications of the simulation and moving average with w (where w is the *window*, a parameter to adjust the smoothness). Output parameters $OPh(t)$ is selected as the observation, because A is closely related to OP. Taking case 3 as an example, the moving averages for $OPh(t)$ with w = 50 h are shown Fig. 3. From the plot we chose a warmup period of t_{warmup} = 80 h (5 d × 16 h).

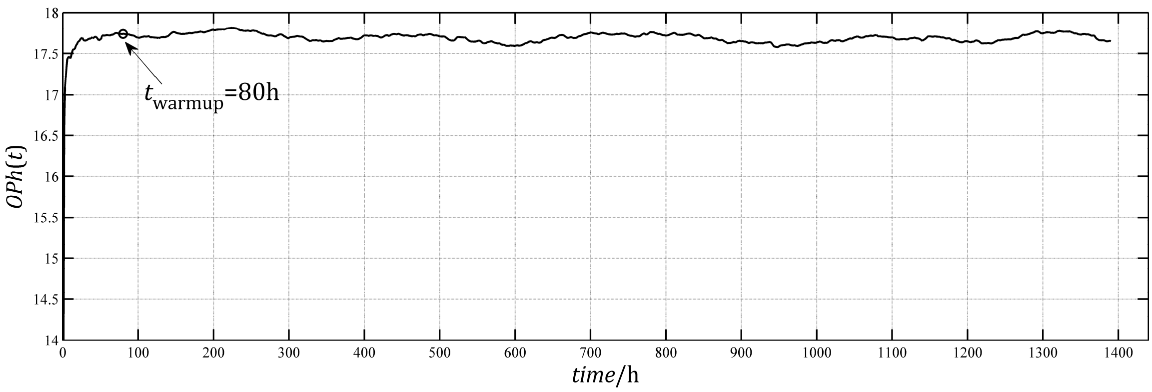

Fig. 3 Moving averages for $OPh(t)$ with w = 50 h

4. Prediction based on Surrogate Model

To increase the efficiency, design of experiments (DOE) [21] technique is usually integrated into the simulation, which also be referred to as computer experiments or simulation experiments [22]. This technique can help us to explore the relationship between input parameters (factors) and performance measures (responses) with the least amount of simulating. Another use of DOE is to construct a surrogate model, also known as metamodel or response surfaces, which is a simplified model of the simulation model for representing the quantitative relationship between factors and responses [23]. Fig. 4 illustrates the relationship of different models. Surrogate model provides one approach to predict the responses from a limited set of simulated factor-level configurations. In this section, a prediction method based on surrogate model, which combines Latin hypercube sampling (LHS) and Kriging model, is presented to predict the availability of production lines.

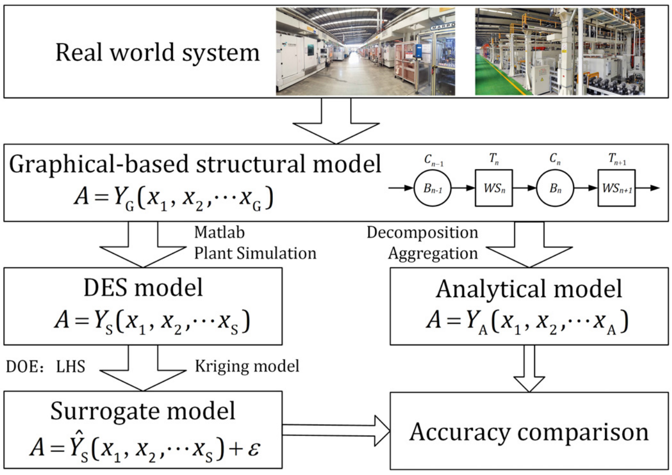

Fig. 4 The relationship of different models

4.1 Latin hypercube sampling

As a kind of space filling design, LHS is widely used for computer experiments because its stratification property allows it to cover the input domain uniformly in a relatively small sample size. A LHS with p sample points in q dimensions is written as an $p \times q$ matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p]^{\mathrm{T}}$, in which each column represents a factor and each row $\boldsymbol{x}_j = [x_{j(1)}, x_{j(2)}, \cdots, x_{j(q)}]$ represents a sample. Firstly, a LHS divides each dimension into p equal levels and gets p^q grids. Then select p of them to make that exactly one is selected at each level. Finally, generate one point randomly in every grid and get p sample points. To get better uniformity, various criterions are available to optimize the design.

In this paper, the LHS experiment plan was created by Matlab function *lhsdesign* with *maxmin* criterion (maximize minimum distance between points) in 40 iterations. For example, Fig. 5 shows a LHS with 100 sample points for variation of workstations' reliability parameters, which is in the range of $MTBF \in$ [-200, 200] h and $MTTR \in$ [-1, 1] h.

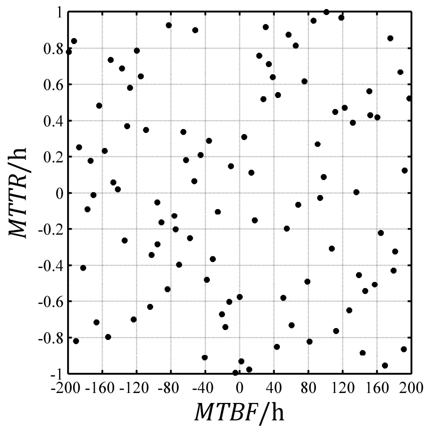

Fig. 5 A 100 × 2 LHS experiment plan

4.2 Kriging model

Kriging is a popular interpolation methodology for computer experiments to construct a cost-effective model as a surrogate to the tedious and time-consuming engineering simulation. Actually, the Kriging is the best linear unbiased interpolation and has the ease of immediate validation by measuring its uncertainty [24]. Compared with traditional polynomial regression, Kriging can give better global predictions because it assumes that the prediction errors are correlated, i.e., gives more weight to 'neighbouring' observations [25].

In Kriging model, the simulation output at design point $\boldsymbol{x}$ is defined as:

$$Y(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\lambda} + Z(\boldsymbol{x}) \tag{1}$$

This model is built by adding up two terms: The first term, which represents the global trend, is a linear combination where $\boldsymbol{b}(\boldsymbol{x})$ is a vector of given basis functions of $\boldsymbol{x}$, and $\boldsymbol{\lambda}$ is a vector of unknown coefficients need to be estimated from the simulation results. The second one $Z(\boldsymbol{x})$ is a local bias expressed by a second-order stationary random process with mean zero and covariance $\mathrm{Cov}[Z(\boldsymbol{x}), Z(\boldsymbol{x}')] = \sigma^2 R(\|\boldsymbol{x}-\boldsymbol{x}'\|; \boldsymbol{\theta})$, where σ^2 can be elucidated as the variance of $Z(\boldsymbol{x})$ for all $\boldsymbol{x}$, and R is the spatial correlation function (SCF) that depends on the 'distances' $\|\boldsymbol{x}-\boldsymbol{x}'\|$ (Euclidean norm). For simplification, the form $R(\|\boldsymbol{x}-\boldsymbol{x}'\|; \boldsymbol{\theta}) = \prod_{v=1}^{q} R\left(\left|x_{(v)} - x'_{(v)}\right|; \theta_{(v)}\right)$ is used for studying q-dimensional problems. The parameters σ^2 and $\boldsymbol{\theta}$ are estimated from the experimental data.

There are two steps to build a surrogate model by Kriging: modelling and prediction. Firstly, model $Y(\boldsymbol{x})$, i.e., estimate the unknown parameters based on the simulated points. Secondly, predict the availability of production lines for given points. We performed these tasks by using a Matlab toolbox called DACE (Design and Analysis of Computer Experiments) [26]. The basis functions $\boldsymbol{b}(\boldsymbol{x})$ are commonly expressed as polynomials of orders d = 0, 1 or 2. For SCF, the toolbox provides seven types and we choose three of them: exponential, Gaussian and spherical, which are defined as follows, respectively:

$$R_{\mathrm{exp}}(|x-x'|; \theta) = \exp(-\theta|x-x'|) \tag{2}$$

$$R_{\mathrm{Gauss}}(|x-x'|; \theta) = \exp(-\theta|x-x'|^2) \tag{3}$$

$$R_{\mathrm{sph}}(|x-x'|; \theta) = 1 - 1.5\xi^2 + 0.5\xi^3, \qquad \xi = \min\{1, \theta|x-x'|\} \tag{4}$$

4.3 Prediction accuracy

The goodness of prediction fit for different models is measured by their mean absolute percentage error (*MAPE*), root-mean-square error (*RMSE*) and R-square (R^2), expressed as:

$$MAPE = \frac{1}{p}\sum_{j=1}^{p}\left|\frac{y_j - \hat{y}_j}{y_j}\right| \times 100\% \tag{5}$$

$$RMSE = \sqrt{\frac{1}{p}\sum_{j=1}^{p}\left(y_j - \hat{y}_j\right)^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{p}\left(y_j - \hat{y}_j\right)^2}{\sum_{j=1}^{p}\left(y_j - \bar{y}\right)^2} \tag{7}$$

where y_j is the simulated value and $\hat{y}_j$ is the forecast value. For the first two indicators, a value closer to zero means a better fit. With respect to R^2, it can take any value between zero and one, and the higher the value, the better accuracy of predictions will be.

5. Case study

In this section, a rough machining production line before heat treatment for car crankshaft is given to illustrate application of the proposed method. It is organized with five workstations, among which the third one is composed of two parallel machines (see Fig. 6). The operating content is described in Table 5, and the design parameter is the same as Case 3 in Section 3.2.

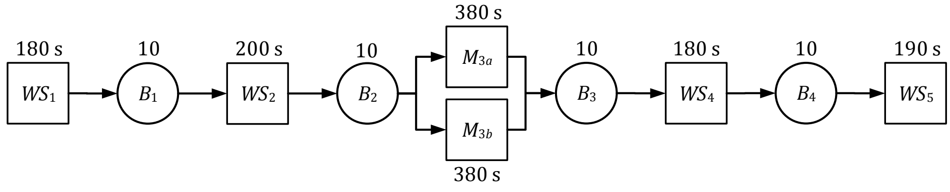

Fig. 6 The block diagram of crankshaft production line

Table 5 The operating content of crankshaft production line

No.	Operating content
OP1	Milling two end faces and location seam, drilling center hole
OP2	Turning rear and front journals
OP3	Finishing turning rear and front journals, main journals and undercuts
OP4	Milling pin journals, undercuts and outer rings of balancers
OP5	Drilling vertical/oblique oil holes and chamfers

To investigate how equipment management influences the performance of the whole line, the availability was predicted at different levels of workstations' reliability and maintainability. We made r = 100 independent replications of the proposed DES with t_{sim} = 1440 h and t_{warmup} = 80 h for p = 100 sample points obtained via experiment plan in Section 4.1. These average simulated values of A are shown by black scatter points in Fig. 7. Then we built the Kriging model from above simulation result, and made predictions for grid points with step sizes of $MTBF$ = 40 h and $MTTR$ = 0.2 h. Another simulation was performed for these grid points as verification group by Plant Simulation software. Fig. 7 displays the predictions obtained from Kriging model with second-order basis functions and Gaussian correlation function. We can see that the surface basically overlaps that of simulated values, which demonstrates that the prediction accuracy is high enough. Meanwhile, the response surface proves that A is an increasing function of $MTBF$ and a decreasing function of $MTTR$.

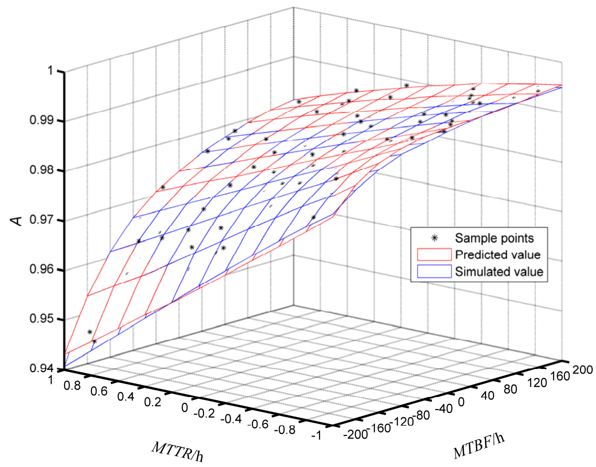

Fig. 7 Response surface of availability A

We also compared various Kriging models with four other methods to show the accuracy of the proposed method. They are moving least square (MLS) method with Gaussian weight function [27], Matlab function *griddata* with *v4* method, aggregation approximate method [11] and semi-analytical simulation [28]. The first two methods interpolate the same sample points and act like Kriging model. The results (listed in Table 6) indicate that Kriging models have the lowest $MAPE$, $RMSE$ and the highest R^2, which means that the proposed method provides better predictions than the other four methods. The methods combining simulation and surrogate model are generally better than analytical methods. Furthermore, for Kriging models, Gaussian correlation function delivers better performance, and the other two perform basically the same. The ideal basis function is second-order polynomial.

Table 6 Prediction accuracy of availability A

Method	$MAPE$ ($\times 10^{-2}$)			$RMSE$ ($\times 10^{-4}$)			R^2		
	$d = 0$	$d = 1$	$d = 2$	$d = 0$	$d = 1$	$d = 2$	$d = 0$	$d = 1$	$d = 2$
Exponential Kriging	6.9253	6.7465	5.4850	10.835	9.4687	7.2949	0.99034	0.99262	0.99562
Gaussian Kriging	4.2619	4.3389	**4.1437**	5.8173	5.9956	**5.7200**	0.99721	0.99704	**0.99731**
Spherical Kriging	6.6904	6.3328	5.6112	10.596	8.7454	7.3571	0.99076	0.99370	0.99554
Gaussian MLS	46.527	17.823	6.5864	67.015	23.145	8.8208	0.63028	0.95590	0.99359
Griddata with v4		8.5439			12.922			0.98625	
Aggregation approximate		65.843			76.846			0.51385	
Semi-analytical simulation		11.139			15.356			0.98059	

6. Conclusion and further research

In this paper, a new procedure has been proposed to predict the availability of discrete production lines with unreliable workstations and finite buffers. The main advantages are its flexibility and property that it is not restricted by specified distributions. This procedure consists of two phases: DES and surrogate model. The first phase involves conducting computer experiments to obtain the availability of production lines under different system parameters with the help of DOE. These input parameters can describe production lines with different structures, operation times, uptimes, downtimes and buffer capacities. The second phase aims to predict the availability based on the simulated results by surrogate model. It is constructed by combining Kriging model and LHS.

A case study of crankshaft production line has shown that the proposed method was practical. It provided better predictions than other interpolation methods, approximate analytical method and semi-analytical simulation. For Kriging model, Gaussian correlation function and second-order basis function predicted with the best performance. We also got the response surface of whole line availability versus workstations' *MTBF* and *MTTR*. The main contribution of this paper are developing a new DES under Matlab environment and importing Kriging model to the domain of modeling the availability of production lines. Although the proposed method was flexible and accurate, the downside is that its efficiency is still lower than analytical methods. Since the method is based on DES, the time-consuming problem will inevitably emerge. Even though LHS and Kriging help a lot, there is still room for improvement, especially for the long lines.

In the future, we will investigate the prediction of other performance measures with more input parameters for higher dimensions, such as speed losses, defect losses or process failure. Moreover, sensitivity analysis methods should be helpful to gain a deeper understanding of the relationships between factors and responses. It would also be interesting to observe the performance of other DOE techniques and surrogate models.

Acknowledgement

This paper is based upon work supported by: 1. National Science and Technology Major Project of China (Grant No. 2014ZX04002031, Grant No. 2014ZX04014010); 2. National Natural Science Foundation of China (Grant No. 51505186); 3. Graduate Innovation Fund of Jilin University (Project 2016027). The authors would like to thank the editors and referees for their helpful comments and insightful advice.

References

[1] Lv, Y., Zhang, J., Qin, W. (2017). A genetic regulatory network-based sequencing method for mixed-model assembly lines, *Advances in Production Engineering & Management*, Vol. 12, No. 1, 62-74, doi: 10.14743/apem2017.1.240.

[2] Govil, M.K., Fu, M.C. (1999). Queueing theory in manufacturing: A survey, *Journal of Manufacturing Systems*, Vol. 18, No. 3, 214-240, doi: 10.1016/S0278-6125(99)80033-8.

[3] Dallery, Y., Gershwin, S.B. (1992). Manufacturing flow line systems: A review of models and analytical results, *Queueing Systems*, Vol. 12, No. 1-2, 3-94, doi: 10.1007/BF01158636.

[4] Gershwin, S.B. (1987). An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking, *Operations Research*, Vol. 35, No. 2, 291-305, doi: 10.1287/opre.35.2.291.

[5] Dallery, Y., David, R., Xie, X.-L. (1988). An efficient algorithm for analysis of transfer lines with unreliable machines and finite buffers, *IIE Transactions*, Vol. 20, No. 3, 280-283, doi: 10.1080/07408178808966181.

[6] Burman, M.H. (1995). *New results in flow line analysis*, PhD Thesis, Operations Research Center, Massachusetts Institute of Technology, USA.

[7] Xia, B., Xi, L., Zhou, B. (2012). An improved decomposition method for evaluating the performance of transfer lines with unreliable machines and finite buffers, *International Journal of Production Research*, Vol. 50, No. 15, 4009-4024, doi: 10.1080/00207543.2011.587842.

[8] Xia, B., Xi, L., Zhou, B., Du, S. (2013). An efficient analytical method for performance evaluation of transfer lines with unreliable machines and finite transfer-delay buffers, *International Journal of Production Research*, Vol. 51, No. 6, 1799-1819, doi: 10.1080/00207543.2012.713137.

[9] Colledani, M., Gershwin, S.B. (2013). A decomposition method for approximate evaluation of continuous flow multi-stage lines with general Markovian machines, *Annals of Operations Research*, Vol. 209, No. 1, 5-40, doi: 10.1007/s10479-011-0961-9.

[10] Li, J., Blumenfeld, D.E., Huang, N., Alden, J.M. (2009). Throughput analysis of production systems: Recent advances and future topics, *International Journal of Production Research*, Vol. 47, No. 14, 3823-3851, doi: 10.1080/00207540701829752.

[11] Li, J., Meerkov, S.M. (2009). *Production Systems Engineering*, Springer, New York, USA, doi: 10.1007/978-0-387-75579-3.

[12] Kuhn, W. (2006). Digital factory – Simulation enhancing the product and production engineering process, In: *Proceedings of the 2006 Winter Simulation Conference*, Monterey, USA, 1899-1906.

[13] Betterton, C.E., Cox, J.F. (2012). Production rate of synchronous transfer lines using Monte Carlo simulation, *International Journal of Production Research*, Vol. 50, No. 24, 7256-7270, doi: 10.1080/0207543.2011.645081.

[14] Dhouib, K., Gharbi, A., Ayed, S. (2009). Simulation based throughput assessment of non-homogeneous transfer lines, *International Journal of Simulation Modelling*, Vol. 8, No. 1, 5-15, doi: 10.2507/IJSIMM08(1)1.111.

[15] Dhouib, K., Gharbi, A., Ayed, S. (2008). Availability and throughput of unreliable, unbuffered production lines with non-homogeneous deterministic processing times, *International Journal of Production Research*, Vol. 46, No. 20, 5651-5677, doi: 10.1080/00207540701294635.

[16] Padhi, S.S., Wagner, S.M., Niranjan, T.T., Aggarwal, V. (2013). A simulation-based methodology to analyse production line disruptions, *International Journal of Production Research*, Vol. 51, No. 6, 1885-1897, doi: 10.1080/00207543.2012.720389.

[17] Heshmat, M., El-Sharief, M.A., El-Sebaie, M.G. (2013). Simulation modeling of automatic production lines with intermediate buffers, *Journal of Engineering Sciences,* Vol. 41, No. 6, 2175-2189.

[18] Supsomboon, S., Hongthanapach, K. (2014). A simulation model for machine efficiency improvement using reliability centered maintenance: case study of semiconductor factory, *Modelling and Simulation in Engineering*, Vol. 2014, 1-9, doi: 10.1155/2014/956182.

[19] Law, A.M. (2015). *Simulation modeling and analysis,* 5th edition, McGraw-Hill, New York, USA.

[20] Macchi, M., Kristjanpoller, F., Garetti, M., Arata, A., Fumagalli, L. (2012). Introducing buffer inventories in the RBD analysis of process production systems, *Reliability Engineering & System Safety*, Vol. 104, 84-95, doi: 10.1016/j.ress.2012.03.015.

[21] Montgomery, D.C. (2013). *Design and analysis of experiments,* 8th edition, Wiley, New Jersey, USA.

[22] Barton, R.R. (2013). Designing simulation experiments, In: *Proceedings of the 2013 Winter Simulation Conference*, Washington, USA, 342-353, doi: 10.1109/WSC.2013.6721432.

[23] Kleijnen, J.P.C. (2015). *Design and analysis of simulation experiments,* 2nd edition, Springer International Publishing, Switzerland, doi: 10.1007/978-3-319-18087-8.

[24] Kleijnen, J.P.C. (2009). Kriging metamodeling in simulation: A review, *European Journal of Operational Research*, Vol. 192, No. 3, 707-716, doi: 10.1016/j.ejor.2007.10.013.

[25] Matta, A., Pezzoni, M., Semeraro, Q. (2012). A Kriging-based algorithm to optimize production systems approximated by analytical models, *Journal of Intelligent Manufacturing*, Vol. 23, No. 3, 587-597, doi: 10.1007/s10845-010-0397-0.

[26] Lophaven, S.N., Nielsen, H.B., Søndergaard, J. (2002). *DACE: A Matlab Kriging toolbox, Version 2.0,* Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

[27] Jia, J. Moving Least Square(MLS2D), from *http://cn.mathworks.com/matlabcentral/fileexchange/34547-moving-least-square-mls2d-*, accessed January12, 2012.

[28] Jia, Y., Yang, Z., Li, G., Du, X., He, J., Wang, L., Li, Q. (2016). An efficient semi-analytical simulation for availability evaluation of discrete production lines with unreliable machines, In: *2016 International Conference on System Reliability and Science*, Paris, France, 115-121, doi: 10.1109/ICSRS.2016.7815849.