



# A two-stage stochastic model for designing cellular manufacturing systems with simultaneous multiple processing routes and subcontracting

M. Mahootchi<sup>a,\*</sup>, K. Forghani<sup>a</sup>, and M. Abdollahi Kamran<sup>a,b</sup>

a. *Department of Industrial Engineering & Management Systems, Amirkabir University of Technology, Tehran, Iran.*

b. *Department of Industrial Engineering, Urmia University of Technology, Urmia, Iran.*

Received 15 September 2016; received in revised form 28 December 2016; accepted 17 June 2017

## KEYWORDS

Cellular manufacturing system;  
 Cell formation;  
 Stochastic programming;  
 Sample average approximation;  
 Multiple processing routes;  
 Subcontracting.

**Abstract.** In recent decades, many researchers have studied the cellular manufacturing system with consideration of various issues such as scheduling, production planning, layout, reliability, etc. However, limited research papers have investigated this problem in an uncertain environment. The present paper addresses a stochastic problem in cellular manufacturing systems considering simultaneous multiple routings and subcontracting. In the developed problem, each part can be simultaneously produced in multiple processing routes. It is also assumed that the unsatisfied part demands as a result of limited machine capacity or high manufacturing cost could be outsourced. A two-stage stochastic programming approach is employed to take the uncertainty into consideration and to formulate the problem. The objective function is to minimize the summation of production, subcontracting, material handling, and machine idleness costs. A sample average approximation method is applied as a solution method. Also, for further illustration of the problem, a numerical example is solved and sensitivity analyses are conducted. Finally, through some numerical examples extracted from related literature, the advantages of constructing a stochastic optimization model for the problem are demonstrated.

© 2018 Sharif University of Technology. All rights reserved.

## 1. Introduction

Cellular Manufacturing System (CMS), as a manufacturing strategy, emanates from group technology concept. In CMSs, each cell is composed of a group of machines that are dedicated to the production of a specific subset of parts called part family. The main advantages expected from production using CMS include a reduction in work-in-process inventories,

setup times, lead times, material handling costs, and tool requirements. Production using CMS also causes a noticeable improvement in productivity, product quality, and production control [1-3]. According to what has been presented by Wemmerlöv and Hyer [1], the following decisions could be taken in the design stage of CMSs:

1. Cell Formation (CF): Grouping parts with similar processing requirements or design features (i.e., making part families) and categorizing machines into machine cells based on the needed operations for part families such that the inter-cell movement of parts is minimized;
2. Group layout: Designing intra-cell layout (layout of

\*. *Corresponding author. Tel.: +98 21 64545387  
 E-mail addresses: mmahootchi@aut.ac.ir (M. Mahootchi);  
 kamran21f@gmail.com (K. Forghani); kamran.m@uut.ac.ir  
 (M. Abdollahi Kamran).*

machines inside cells) and inter-cell layout (layout of cells regarding each other);

3. Group scheduling: Scheduling parts and part families for production;
4. Resource allocation: Assigning tools, manpower, materials, and other resources.

These issues have been investigated in many research papers. For instance, the CF problem has been addressed in [4-7]. The layout of CMS has been discussed in [3,8-11]. The cell scheduling problem has been considered in [12-15].

In reality, the aforementioned decisions should be taken in an uncertain environment. However, limited studies have investigated the uncertainty issue in designing CMSs. The uncertainty could appear in demands, operational costs, resource capacities, etc. Dealing with this issue, Wang et al. [16] surveyed a facility layout problem in CMSs in which the demand rate varied over product life cycle. In their problem, the objective was to find the inter- and intra-cell layouts such that the total material handling cost was minimized. A simulated annealing algorithm was applied to solve the presented problem. Jeon and Leep [17] developed a two-phase procedure to configure a CMS. In the first phase, they used a genetic algorithm to obtain part families using a new similarity coefficient. The proposed similarity coefficient considered the possibility of alternative processing routes during machine failure. In the second phase, considering the periodic demand changes, machines were assigned to part families using multiple optimization processes. In these optimization processes, sequential and simultaneous mixed integer programming models were employed to minimize the total costs that were in connection with the operational and scheduling aspects. Tavakkoli-Moghaddam et al. [18] surveyed a facility layout problem in CMSs with normally distributed part demands. The objective was to optimize the total inter- and intra-cell material handling costs. Schaller [19] presented a model for the CF problem considering periodic demand variability. In their problem, the composition of cells was allowed to change from period to period. Heuristic algorithms were implemented to minimize the sum of three cost components including the production cost of parts, the amortized cost of machines, and the relocating cost of machines during periods. Safaei et al. [20] assumed that part demands and machine capacities were fuzzy numbers. Then, they formulated a dynamic CMS model. The objective of this model was to determine the optimal cell configuration in each period in such a way that the sum of material handling, machine constant/variable, and reconfiguration costs was minimized. They used a fuzzy programming based method to solve the developed problem. Arıkan and Güngör [21] presented a multi-objective CMS model

in a fuzzy environment. In this study, part demands, machine capacities, and Exceptional Elements (EEs) elimination costs were considered as fuzzy numbers. The objective functions that they considered included the minimization of EEs elimination cost, the minimization of inter-cell movements, and the maximization of utilized machine capacity. Ghezavati and Saidi-Mehrabad [22] addressed an integrated mathematical model of CF and group scheduling problems in an uncertain environment. It was assumed that the processing time of parts on machines was a stochastic parameter, which was represented by discrete scenarios. The main goal of their model was to minimize the total expected costs of maximum tardiness, EEs subcontracting, and resource under-utilization. A hybrid genetic-simulated annealing algorithm was employed as a solution method. Das and Abdul-Kader [23] presented a bi-objective integer-programming model for designing a CMS by considering dynamic changes in machine reliability and parts demands. The first objective function was to minimize the total system costs including the manufacturing, inter-cell material handling, machine under-utilization, and machine duplication costs. The second objective function was to maximize the total system reliability. An  $\varepsilon$ -constraint solution method was used to solve the problem. Ghezavati and Saidi-Mehrabad [24] applied a queuing theory approach to design a CMS with exponentially distributed service and arrival times. It was assumed that each machine worked as a server and each part was a customer that should be served by machines. They formulated a mathematical model to maximize the average utilization level of machines. A hybrid method based on genetic and simulated annealing algorithms was exerted to solve the problem. Rabbani et al. [25] proposed a bi-objective CF problem in which part demands were expressed by some probabilistic scenarios. A two-stage stochastic programming model was presented to undertake the uncertain demand of parts. The expected variable cost of all machines and the expected inter-cell material handling cost were considered in the first objective function. The total expected cell load variation was considered as the second objective function. They applied a two-phase fuzzy linear programming approach to solve the presented problem. Forghani et al. [26] applied an interval robust optimization approach to take the uncertainty of part demands into consideration. Then, an integrated CF and layout problem was formulated to minimize the inter- and intra-cell material handling costs.

Based on the above survey, the following shortcomings in the developed models can be more investigated:

- Using normal probability distributions: In practice,

uncertain parameters are not necessarily normally distributed and may follow other distributions. However, for simplicity, most papers assume that random variables are normally distributed;

- Considering single process routing and other simplifying assumptions: Most related papers usually assume that each part has a unique process routing. While, in practice, each part may have multiple processing routes. Consideration of multiple processing routes in the design of CMSs may enhance planning flexibility and throughput rates [27]. Furthermore, other manufacturing parameters such as operation sequences, processing times, resource capacities, and operational costs are not usually addressed in the existing mathematical models;
- Limitation on considering machine duplication: In most current studies, the number of each machine type is known a priori and there is not any capacity limitation on machines. In other words, most developed models do not make a decision about the number of machines in the cell design process. In reality, the capacity of machines is limited. Therefore, each machine can be duplicated so as to cope with the production requirement of parts;
- Neglecting outsourcing costs: Producing all parts in CMSs may not be possible due to either limitation on resource capacities or high manufacturing costs. Accordingly, outsourcing some parts to external suppliers may be preferred under certain circumstances. However, this issue has been addressed only in few research papers.

To deal with the shortcomings mentioned above, in this paper, we propose a stochastic CMS model considering subcontracting and simultaneous multiple processing routes. Furthermore, practical design parameters such as operation sequences, processing times, and machines capacities are taken into account. It is assumed that part demands and outsourcing costs are stochastic parameters with known probability distributions. In the suggested problem, each part can be simultaneously produced in multiple processing routes. Unsatisfied part demands as a result of limited machine capacity or high manufacturing cost are outsourced. A two-stage stochastic programming approach is applied to cope with the uncertainty and formulate the problem. By considering the cell size, machine capacity, and budget constraints, the first stage decisions are the assignment of each machine to a cell and the number of purchased machines of each type. The second stage variables, which are dependent on the various realizations of uncertain parameters and the first stage decisions, are the amount of part demands that should be outsourced and produced in each processing route. The objective function is to

minimize the total variable cost, including the production, subcontracting, material handling, and machine idleness costs. A solution method based on Sample Average Approximation (SAA) approach is suggested to find suitable solutions to this stochastic model. To illustrate the problem, a numerical example is solved and sensitivity analyses are conducted. Finally, numerical examples extracted from the related literature are solved to illustrate the efficiency of the model and to demonstrate its advantages over other developed models.

The remainder of this paper is organized as follows: In Section 2, the proposed problem is explained in detail and a non-linear two-stage stochastic model is presented. Furthermore, a linearization method is applied to linearize the model. In Section 3, the SAA method is presented. In Section 4, in order to clarify the proposed problem, an illustrative numerical example is solved and sensitivity analyses are carried out. In Section 5, by means of solving ten numerical examples extracted from the literature, the effectiveness of the solution method, as well as the advantages of the constructed model, is examined. Finally, the conclusions and hints for future studies are given in Section 6.

## 2. Problem statement

In this section, a two-stage stochastic mathematical programming model is developed for designing a CMS. It is assumed that a set of parts,  $i = 1, \dots, P$ , each having an uncertain demand with a known probability distribution, should be produced by a set of machine types,  $k = 1, \dots, M$ . Each machine type has a limited capacity, which is known a priori. It is assumed that there are  $R_i$  processing routes for the production of part  $i$ . Each of these routings can be independently implemented in the production of part  $i$ . In each routing, the sequences of operations and processing times are known a priori. Unsatisfied part demands, which can result from limited machine capacity or high manufacturing costs, are outsourced. The suggested subcontracting approach is similar to that proposed by Mohammadi and Forghani [11]. The outsourcing price of each part is also a random variable with a known probability distribution. Machine duplication (i.e., purchasing machines) is allowed and machines of the same type are allocated to the same cell. Given the budget constraint and cell size limit (the maximum number of machines that can be assigned to a cell), machines are grouped into a maximum of  $C^{\max}$  cells. A schematic illustration of the proposed problem is given in Figure 1.

Two-stage stochastic programming is one of the methods which can be used to take the uncertainty into consideration. It has two different types of decision:

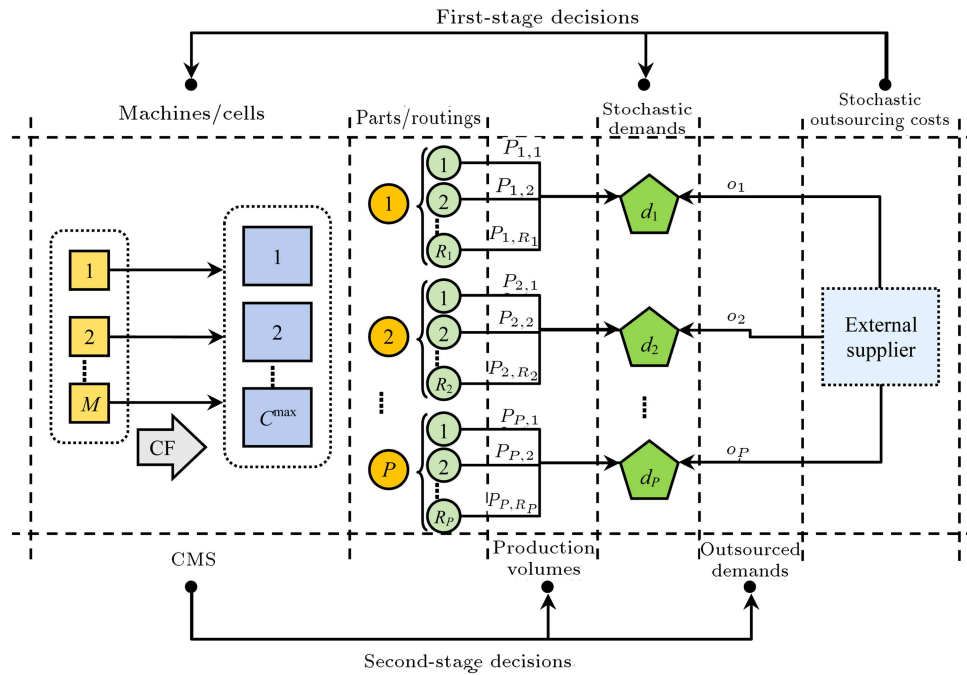


Figure 1. An illustration of the proposed problem.

the first- and second-stage decisions. The first-stage decisions, also called strategic decisions, are made before the realization of the uncertain parameters. In the second stage, where the uncertain parameters are realized, operational and tactical decisions are made [28]. In the suggested problem, we use this method to tackle the uncertainty, which arises from part demands and outsourcing costs.

2.1. Notations

The following notations are applied throughout this paper.

Indices

- $i$  Index for parts ( $i = 1, \dots, P$ , where  $P$  is the number of parts),
- $j$  Index for processing routes ( $j = 1, \dots, R_i$ , where  $R_i$  is the number of processing routes of part  $i$ ),
- $k$  Index for machine types ( $k = 1, \dots, M$ , where  $M$  is the number of machine types),
- $l$  Index for cells ( $l = 1, \dots, C^{\max}$ , where  $C^{\max}$  is the maximum number of cells allowed),
- $s$  Index for scenarios ( $s = 1, \dots, S$ , where  $S$  is the number of scenarios).

Parameters

- $\pi_{[s]}$  Probability of scenario  $s$ ,
- $d_{i[s]}$  Demand of part  $i$  in scenario  $s$ ,
- $c_{i[s]}^O$  Unit outsourcing cost of part  $i$  in scenario  $s$ ,

- $c_{ij}^P$  Unit production cost of part  $i$  using processing route  $j$ ,
  - $c_{ikk'}^A$  Unit intra-cell material handling cost for transporting part  $i$  between machine types  $k$  and  $k'$ ,
  - $c_{ikk'}^E$  Unit inter-cell material handling cost for transporting part  $i$  between machine types  $k$  and  $k'$  ( $c_{ikk'}^E \leq c_{ikk'}^A$ ),
  - $f_{ijkk'}$  Number of times that part  $i$  in processing route  $j$  is transported between machine types  $k$  and  $k'$ ,
  - $t_{ijk}$  Processing time of part  $i$  on machine type  $k$  in processing route  $j$ ,
  - $c_k^M$  Purchase price of one unit of machine type  $k$ ,
  - $T_k$  Available time on machine type  $k$ ,
  - $c_k^I$  Idleness cost of machine type  $k$  per unit time,
  - $N_k^{\max}$  Maximum number for machine type  $k$  allowed to be purchased,
  - $NM$  Maximum number of machine types allowed in a cell,
  - $B$  Budget available for purchasing machines,
  - $w_S$  Estimated total variable cost based on  $S$  scenarios.
- First-stage decision variables**
- $z_{kl}$  If machine type  $k$  is assigned to cell  $l$ ; otherwise 0,

$n_k$  Number of machine type  $k$  to be purchased.

**Second-stage decision variables**

$p_{ij[s]}$  Amount of part  $i$  to be produced using processing route  $j$  in scenario  $s$ ,

$o_{i[s]}$  Amount of part  $i$  to be outsourced in scenario  $s$ ,

$u_{k[s]}$  Unused capacity (time) of machine type  $k$  in scenario  $s$ .

**2.2. Mathematical model**

The two-stage mathematical model of the proposed problem is given below:

$$w_s = \min \sum_{s=1}^s \pi[s] \left( \sum_{i=1}^P \sum_{j=1}^{R_i} c_{ij}^P p_{ij[s]} + \sum_{i=1}^P c_{i[s]}^O o_{i[s]} + \sum_{k=1}^M c_k^I u_{k[s]} + \sum_{k=1}^{M-1} \sum_{k'=k+1}^M \sum_{i=1}^P \sum_{j=1}^{R_i} f_{ijkk'} p_{ij[s]} \left[ c_{ikk'}^A \sum_{l=1}^{C^{\max}} z_{kl} z_{k'l} + c_{ikk'}^E \left( 1 - \sum_{l=1}^{C^{\max}} z_{kl} z_{k'l} \right) \right] \right), \tag{1}$$

subject to:

$$\frac{n_k}{N_k^{\max}} \leq \sum_{l=1}^{C^{\max}} z_{kl} \leq 1, \quad \forall k, \tag{2}$$

$$\sum_{l=1}^{C^{\max}} z_{kl} \leq n_k, \quad \forall k, \tag{3}$$

$$\sum_{k=1}^M z_{kl} \leq NM, \quad \forall l, \tag{4}$$

$$o_{i[s]} + \sum_{j=1}^{R_i} p_{ij[s]} = d_{i[s]}, \quad \forall i, s, \tag{5}$$

$$\sum_{i=1}^P \sum_{j=1}^{R_i} t_{ijk} p_{ij[s]} + u_{k[s]} = T_k n_k, \quad \forall k, s, \tag{6}$$

$$\sum_{k=1}^M c_k^M n_k \leq B, \tag{7}$$

$$n_k \in \{0, \dots, N_k^{\max}\}, \quad \forall k, \tag{8}$$

$$z_{kl} \in \{0, 1\}, \quad \forall k, l, \tag{9}$$

$$p_{ij[s]}, o_{i[s]}, u_{k[s]} \geq 0, \quad \forall i, j, k, s. \tag{10}$$

In the proposed model, objective function (1) minimizes the estimation of the expected total variable cost, which is composed of production, subcontracting, idleness, and material handling costs (i.e., intra- and inter-cell material handling costs), according to different realizations of uncertain parameters. Constraints (2) and (3) jointly represent that if a machine type is purchased, it is only assigned to a single cell. Constraint (4) restricts the number of purchased machines in each cell. Constraint (5) ensures that under various scenarios, the demand of parts is satisfied using production and outsourcing. Constraint (6) guarantees that machine capacity is not exceeded. To be more precise, the amount of used and unused time of machine type,  $k$ , should be equal to its total available time. Constraint (7) presents the budgetary limitations on purchasing machines. Constraint (8) states that the number of each machine type is an integer variable; also, it limits the number of each machine type. Finally, Constraints (9) and (10) indicate the types of other decision variables.

**2.3. Mathematical model linearization**

The mathematical model presented in Subsection 2.2 is a mixed-integer nonlinear program (MINLP) due to the existence of a nonlinear term in the fourth cost component of objective function (1). Therefore, a linearization method is applied to transform the model into a Mixed-Integer linear Program (MIP), which is very efficient to optimally solve using high-performance solvers such as GURUBI, XPRESS, or CPLEX. In doing so, objective function (1) is rewritten as Eq. (11):

$$\min \sum_{s=1}^s \pi[s] \left( \sum_{i=1}^P \sum_{j=1}^{R_i} c_{ij}^P p_{ij[s]} + \sum_{i=1}^P c_{i[s]}^O o_{i[s]} + \sum_{k=1}^M c_k^I u_{k[s]} + \sum_{k=1}^{M-1} \sum_{k'=k+1}^M \sum_{i=1}^P \sum_{j=1}^{R_i} f_{ijkk'} c_{ikk'}^E p_{ij[s]} - \sum_{k=1}^{M-1} \sum_{k'=k+1}^M \sum_{l=1}^{C^{\max}} \left( Z_{kl} Z_{k'l} \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi[s] f_{ijkk'} (c_{ikk'}^E - c_{ikk'}^A) p_{ij[s]} \right) \right). \tag{11}$$

Then, a new auxiliary variable,  $\lambda_{kk'l}$ , is defined to replace with the nonlinear term:

$$z_{kl} z_{k'l} \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi[s] f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) p_{ij[s]},$$

in Eq. (11). Eq. (11) aims at minimizing the negative coefficient of this nonlinear term; therefore, inequality:

$$\lambda_{kk'l} \leq \min \left\{ BMz_{kl}, BMz_{k'l}, \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) p_{ij[s]} \right\}, \forall k' > k,$$

is valid, where  $BM$  is an upper bound on  $\sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) p_{ij[s]}$  (it is reasonable to let  $BM = \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) d_{i[s]}$ , because  $p_{ij[s]} \leq d_{i[s]}, \forall j$ , see Constraint (5)). On the other hand, according to Constraints (2) and (9), we know that  $\sum_{l=1}^{C^{\max}} z_{kl} z_{k'l} \leq 1, \forall k' > k$ . Therefore, inequality:

$$\sum_{l=1}^{C^{\max}} \lambda_{kk'l} \leq \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) p_{ij[s]}, \forall k' > k,$$

is also valid. Furthermore, we know that  $c_{ikk'}^E \geq c_{ikk'}^A$ ; thus, it is correct to derive inequality  $\lambda_{kk'l} \geq 0, \forall k' > k$ . Now, based on these three sets of inequalities, Constraints (13)-(16) are added to the model in order to complete the linearization. The linearized mathematical model of the problem is given below:

$$\begin{aligned} W_S = \min \sum_{s=1}^s \pi_{[s]} & \left( \sum_{i=1}^P \sum_{j=1}^{R_i} c_{ij}^P p_{ij[s]} \right. \\ & + \sum_{i=1}^P c_{i[s]}^O o_{i[s]} + \sum_{k=1}^M c_k^I u_k[s] \\ & + \sum_{k=1}^{M-1} \sum_{k'=k+1}^M \sum_{i=1}^P \sum_{j=1}^{R_i} f_{ijkk'} c_{ikk'}^E p_{ij[s]} \left. \right) \\ & - \sum_{k=1}^{M-1} \sum_{k'=k+1}^M \sum_{l=1}^{C^{\max}} \lambda_{kk'l}, \end{aligned} \tag{12}$$

subject to Eqs. (2)-(10):

$$\sum_{l=1}^{C^{\max}} \lambda_{kk'l} \leq \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) p_{ij[s]}, \forall k' > k, l, \tag{13}$$

$$\lambda_{kk'l} \leq \left( \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) d_i \right) z_{kl}, \forall k' > k, l, \tag{14}$$

$$\lambda_{kk'l} \leq \left( \sum_{i=1}^P \sum_{j=1}^{R_i} \sum_{s=1}^S \pi_{[s]} f_{ijkk'} \left( c_{ikk'}^E - c_{ikk'}^A \right) d_i \right) z_{k'l}, \forall k' > k, l, \tag{15}$$

$$\lambda_{kk'l} \geq 0, \forall k' > k, l. \tag{16}$$

It should be mentioned that the model above, which is also called the SAA problem, includes  $3/2 M(M - 1) C^{\max} + 3M + C^{\max} + S(P + M) + 1$  constraints,  $S(M + P + \sum_i R_i) + M/2(M - 1)C^{\max}$  positive variables,  $MC^{\max}$  binary variables, and  $M$  integer variables.

### 3. The SAA method

SAA is a Monte Carlo simulation-based solution method in which large numbers of generated scenarios are used to find bounds of the objective function of a stochastic problem. In the proposed model, outsourcing costs and part demands are stochastic parameters with known probability distributions. In this section, the stochastic model is formulated in a concise form such that the implementation of the SAA method can be easily explained. To do so, by ignoring the indices of the parameters and the decision variables, they are presented in a bold form (e.g.,  $z_{kl}$  is denoted by  $\mathbf{z}$ ). The stochastic parameters are presented by  $\boldsymbol{\xi}(\omega) = (\mathbf{c}^o(\omega), \mathbf{d}(\omega))$  where  $\omega \in \Omega$  ( $\Omega$  is the set of all scenarios with a known probability distribution  $\mathcal{P}$ ) denotes a *scenario* which becomes known while making second-stage recourse decision  $(\mathbf{p}, \mathbf{o}, \mathbf{u})$ . Also, the objective value under decision  $(\mathbf{z}, \mathbf{p}, \mathbf{o}, \mathbf{u})$  and a particular realization of uncertain parameter,  $\mathbf{c}^o$ , is indicated by  $w(\mathbf{z}, \mathbf{p}, \mathbf{o}, \mathbf{u}, \mathbf{c}^o(\omega))$ . Now, assume that  $X_1$  denotes the set of constraints that are not affected by the uncertainty; this kind of constraints is called the first-stage constraints. Similarly, assume that  $X_2(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega))$  denotes the set of constraints affected by the uncertainty and the first-stage decision; this kind of constraints is also called the second-stage constraints. According to these definitions, the concise form of the proposed stochastic problem is as follows:

$$w^* = \min_{(\mathbf{z}, \mathbf{n}) \in X_1} \mathbb{E}_{\mathcal{P}}[Q(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega))], \tag{17}$$

where:

$$Q(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega)) =$$

$$\min\{w(\mathbf{z}, \mathbf{p}, \mathbf{o}, \mathbf{u}, \mathbf{c}(\omega)) | (\mathbf{p}, \mathbf{o}, \mathbf{u}) \in X_2(\mathbf{z}, \mathbf{n}, \mathbf{d}(\omega))\}. \tag{18}$$

In Eq. (17), the vector of binary and integer variables,  $\mathbf{z}$  and  $\mathbf{n}$ , corresponds to the first-stage strategic decisions, and  $Q(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega))$  represents the optimal total variable cost associated with the given strategic decision  $(\mathbf{z}, \mathbf{n})$  and a particular realization of uncertain parameter  $\boldsymbol{\xi}(\omega)$ .

Given the probability distribution  $\mathcal{P}, S$  sample scenarios are generated from  $\Omega$  using Monte Carlo simulation (the samples are denoted by  $\omega^1, \omega^2, \dots, \omega^S$ ) in order to approximate the expected value function  $\mathbb{E}[Q(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega))]$ . Then, a deterministic optimization problem (also called the SAA problem) specified by the generated samples is solved. The following is the SAA problem in a concise form:

$$w_S = \min_{(\mathbf{z}, \mathbf{n}) \in X_1} \frac{1}{S} \sum_{s=1}^S Q(\mathbf{z}, \mathbf{n}, \boldsymbol{\xi}(\omega^s)). \tag{19}$$

In fact, the problem above (i.e., Eq. (19)) is a concise form of the model presented in Subsection 2.3 (where  $X_1$  is the set of Constraints (2)-(4) and (7)-(9);  $X_2(\mathbf{z}, \mathbf{n}, \mathbf{d}(\omega))$  is the set of Constraints (5), (6), (10), and (13)-(16); and  $\pi_{[s]} = 1/S$ ).

It is worth mentioning that to provide a good statistical estimation of  $w^*$ , the approximation process should be repeated several times using different independent sample scenarios. At each experiment, the best solution, as well as the estimation of the upper and lower bounds of the objective function, is calculated. This process is repeated until a satisfactory optimality gap is achieved. According to Shapiro and Homem-de-Mello [29], the steps of the SAA method can be summarized as follows:

**Step 1.** Randomly generate  $T$  independent samples of stochastic parameters each with  $S$  scenarios, i.e.,  $\omega_t^1, \omega_t^2, \dots, \omega_t^S$  for  $t = 1, \dots, T$ . Solve the two-stage problem (i.e., Eq. (19))  $T$  times with  $S$  scenarios to obtain objective values  $w_S^1, w_S^2, \dots, w_S^T$  and candidate solutions  $(\mathbf{z}^1, \mathbf{n}^1), (\mathbf{z}^2, \mathbf{n}^2), \dots, (\mathbf{z}^T, \mathbf{n}^T)$ .

**Step 2.** Calculate the average of  $T$  optimal objective values by:

$$\bar{w}_S = \frac{1}{T} \sum_{t=1}^T w_S^t, \tag{20}$$

where  $\bar{w}_S$  is an unbiased estimator of  $\mathbb{E}[w_S]$ . It is well-known that the expected value of  $w_S$  is less than or equal to the optimal objective value of the true problem, that is,  $\mathbb{E}[w_S] \leq w^*$  (for example, see [30]). Thus,  $\bar{w}_S$  provides a statistical estimate for a lower bound on the optimal value of the true problem. The standard deviation of  $\bar{w}_S$  can be estimated by:

$$\hat{\sigma}_{T,S} = \sqrt{\frac{1}{(T-1)T} \sum_{t=1}^T (w_S^t - \bar{w}_S)^2}. \tag{21}$$

An approximate  $100(1-\alpha)\%$  confidence lower bound on  $\mathbb{E}[w_S]$  or  $w^*$  can be calculated by:

$$w_L = \bar{w}_S - t_{\alpha, T-1} \hat{\sigma}_{T,S}, \tag{22}$$

where  $t_{\alpha, T-1}$  is the  $\alpha$ -critical value of the  $t$ -distribution with  $T-1$  degrees of freedom.

**Step 3.** Randomly produce  $S'$  ( $S'$  should be quite larger than  $S$ ) independent samples of stochastic parameters, i.e.,  $\omega^1, \omega^2, \dots, \omega^{S'}$ . Based on these  $S'$  scenarios, compute the estimated objective value of candidate solution  $(\mathbf{z}^t, \mathbf{n}^t), \forall t = 1, \dots, T$ , using the following estimator:

$$\hat{w}_{S'}(\mathbf{z}^t, \mathbf{n}^t) = \frac{1}{S'} \sum_{s=1}^{S'} Q(\mathbf{z}^t, \mathbf{n}^t, \boldsymbol{\xi}(\omega^s)), \tag{23}$$

$$\forall t = 1, \dots, T.$$

Note that for each candidate solution, this step involves solving  $S'$  independent second-stage sub-problems given in Eq. (18). Choose  $(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*)$  as one of the candidate solutions  $(\mathbf{z}^1, \mathbf{n}^1), (\mathbf{z}^2, \mathbf{n}^2), \dots, (\mathbf{z}^T, \mathbf{n}^T)$  which has the smallest estimated objective value, that is:

$$(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*) \in \{(\mathbf{z}^{t^*}, \mathbf{n}^{t^*}) | t^* \in \arg \min_{t=1, \dots, T} \hat{w}_{S'}(\mathbf{z}^t, \mathbf{n}^t)\}.$$

Calculate an approximate  $100(1-\alpha)\%$  confidence upper bound on the true objective value of solution  $(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*)$ , that is,  $w(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*)$  by:

$$w_U = \hat{w}_{S'}(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*) + \Phi^{-1}(1-\alpha) \hat{\sigma}_{S'}, \tag{24}$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution and  $\hat{\sigma}_{S'}$  is computed by:

$$\hat{\sigma}_{S'} = \sqrt{\frac{1}{(S'-1)S'} \sum_{s=1}^{S'} \left( Q(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*, \boldsymbol{\xi}(\omega^s)) - \hat{w}_{S'}(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*) \right)^2}. \tag{25}$$

**Step 4.** Obtain a statistically valid bound with the confidence of at least  $100(1-2\alpha)\%$  on the true optimality gap of solution  $(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*)$  by:

$$\hat{g}(\hat{\mathbf{z}}^*, \hat{\mathbf{n}}^*) = w_U - w_L. \tag{26}$$

If the estimated optimality gap is small enough, the obtained solution is almost optimal for the first-stage decision on the true stochastic problem. Otherwise, the values of  $S$  or  $T$  should be increased and all the above steps should be repeated until a satisfactory optimality gap is achieved.

**4. An illustrative example**

In this section, a numerical example adopted from [2] is used to illustrate the proposed problem. Also, by using this example, sensitivity analyses are carried out to investigate the behavior of the solution in terms of

the budget constraint and the SAA parameters. This numerical example includes 20 parts, 10 machine types, and 36 processing routes. Tables 1 and 2 contain the original data plus the additional ones that we have generated. The maximum number of cells ( $C^{max}$ ) is assumed to be 2, and the maximum number of

**Table 1.** Data related to the parts in the illustrative example\*.

<i>i</i>	<i>j</i>	$\bar{d}_i^\dagger$	$c_{ij}^P$ (\$)	$c_{ikk'}^A$ (\$)	$c_{ikk'}^E$ (\$)	Machine type (processing time)						
						<i>Op</i> <sub>1</sub>	<i>Op</i> <sub>2</sub>	<i>Op</i> <sub>3</sub>	<i>Op</i> <sub>4</sub>	<i>Op</i> <sub>5</sub>		
1	1	150	13	0.5	1.5	2 (2)	3 (3)	4 (2)				
	2		11			0.5	1.5	3 (1)	4 (3)	9 (2)		
2	1	226	10	0.5	1.5	1 (2)	2 (3)	3 (4)	7 (3)			
3	1	335	15	0.5	1.5	3 (2)	6 (3)	8 (4)	9 (2)			
4	1	446	12	0.5	1.5	3 (3)	4 (2)	5 (4)	9 (2)			
	2		12			0.5	1.5	3 (3)	4 (3)	9 (2)	10 (2)	
5	1	274	8	0.5	1.5	1 (2)	8 (3)	10 (4)				
	2		8			0.5	1.5	2 (3)	8 (3)	10 (4)		
6	1	171	14	0.5	1.5	1 (2)	2 (3)	7 (2)	8 (1)	10 (5)		
	2		12			0.5	1.5	1 (2)	2 (3)	8 (2)	10 (5)	
7	1	218	11	0.5	1.5	3 (3)	4 (2)	9 (3)	7 (1)	10 (2)		
	2		11			0.5	1.5	3 (3)	4 (2)	5 (2)	9 (3)	10 (2)
	3		10			0.5	1.5	3 (3)	5 (2)	9 (2)	7 (1)	10 (2)
8	1	273	8	0.5	1.5	6 (2)	1 (3)	9 (1)	10 (2)			
	2		8			0.5	1.5	1 (3)	6 (2)	9 (1)	10 (2)	
9	1	307	14	0.5	1.5	1 (3)	6 (1)	8 (4)	9 (2)	10 (3)		
	2		13			0.5	1.5	1 (3)	6 (2)	9 (3)	10 (3)	
10	1	414	10	0.5	1.5	1 (3)	6 (2)	9 (2)	10 (1)			
	2		11			0.5	1.5	6 (3)	8 (2)	9 (2)	10 (2)	
11	1	223	11	0.5	1.5	1 (2)	2 (3)	9 (4)	10 (1)			
	2		11			0.5	1.5	1 (2)	2 (3)	8 (5)	10 (1)	
12	1	378	6	0.5	1.5	2 (3)	3 (2)	9 (1)	7 (1)			
	2		6			0.5	1.5	2 (3)	3 (2)	4 (1)	7 (1)	
13	1	328	7	0.5	1.5	8 (3)	9 (2)	10 (4)				
	2		7			0.5	1.5	6 (2)	9 (3)	10 (4)		
14	1	280	8	0.5	1.5	1 (1)	9 (3)	10 (2)				
	2		9			0.5	1.5	2 (2)	9 (3)	10 (2)		
15	1	270	6	0.5	1.5	3 (1)	5 (2)	9 (1)	7 (1)			
	2		7			0.5	1.5	3 (1)	5 (2)	9 (1)	10 (2)	
16	1	182	9	0.5	1.5	1 (2)	2 (3)	8 (2)	10 (3)			
17	1	244	8	0.5	1.5	1 (3)	2 (2)	3 (3)	10 (4)			
18	1	152	7	0.5	1.5	1 (5)	2 (4)	3 (2)	4 (2)			
	2		9			0.5	1.5	3 (4)	4 (1)	9 (1)	10 (1)	
19	1	366	10	0.5	1.5	3 (4)	4 (1)	9 (1)	10 (1)			
	2		9			0.5	1.5	3 (3)	4 (1)	9 (1)	7 (1)	
20	1	226	7	0.5	1.5	4 (1)	9 (1)	10 (2)				
	2		6			0.5	1.5	3 (1)	4 (1)	9 (1)		

\* $d_i \sim \text{Normal}(\bar{d}_i, \bar{d}_i/3)$  and  $c_i^O \sim \text{Uniform}(1.25 \times \max_j c_{ij}^P, 1.75 \times \max_j c_{ij}^P)(\$)$ .

†  $\bar{d}_i$  is the part demands given in [2].



**Table 2.** Data related to the machine types in the illustrative example.

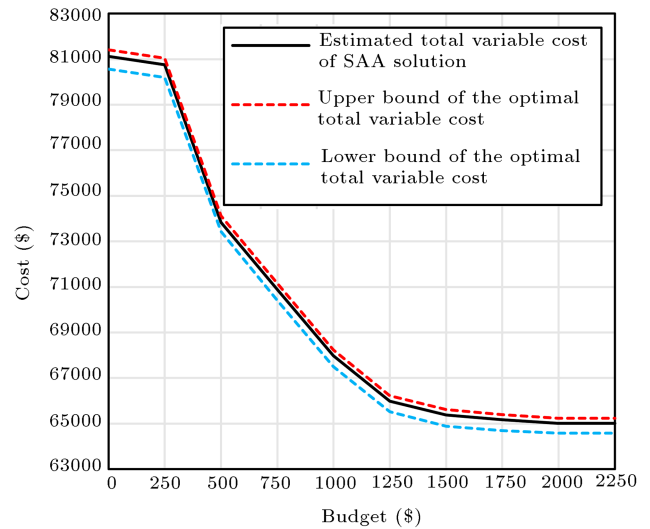
$k$	$T_k$	$c_k^M$ (\$)	$c_k^I$ (\$)	$N_k^{\max}$
1	3000	106	0.11	3
2	3000	136	0.14	3
3	3000	65	0.07	3
4	3000	140	0.14	3
5	3000	103	0.10	3
6	3000	61	0.06	3
7	3000	126	0.13	3
8	3000	93	0.09	3
9	3000	94	0.09	3
10	3000	70	0.07	3

machines allowed in a cell ( $NM$ ) is assumed to be 5. The available budget for purchasing machines ( $B$ ) is limited to \$1500. In order to show the capability of the proposed problem in considering various distributions, it is assumed that the demand of part  $i$  is normally distributed with mean  $\bar{d}_i$  and standard deviation  $\bar{d}_i/3$ . Also, it is assumed that the outsourcing cost of part  $i$  is uniformly distributed over the interval  $[1.25 \max_j c_{ij}^P, 1.75 \max_j c_{ij}^P]$ . This example is solved using the SAA method at confidence level of 95% ( $\alpha = 0.025$ ) assuming  $S = 30$ , and  $S' = 2000$ . A summary of the results is provided in Table 3. It should be noted that the mathematical model of the SAA problem is coded in the GAMS 24.5, and the GUROBI 6.0 is selected as the default MIP and LP solver. Computations are performed on a PC having Microsoft Windows 10 operating system with Intel(R) Core (TM) i7-4790K 4.00 GHz CPU and 16GB of RAM. In order to use all the available CPU cores (4 cores and 8 threads), we set ‘threads = 0’ in the GAMS option (this can lead to considerable saving of the CPU time).

It took almost 351 seconds for the SAA method

to solve the illustrative example, with an estimated total variable cost ( $\hat{w}_{S'}(\hat{z}^*, \hat{n}^*)$ ) equal to \$65385.243. Refer to Table 3 to see the machine cells and the number of each machine type in each cell. According to Eqs. (22) and (24), the lower bound on the optimum total variable cost ( $w_L$ ) and the upper bound on the true total variable cost of the SAA solution ( $w_U$ ) are equal to \$64886.883 and \$65617.360, respectively. This also means that the optimal total variable cost of the stochastic model ( $w^*$ ) lies in the interval  $[65100.815, 65226.736]$  with 95% confidence. Based on these bounds, the estimated optimality gap ( $\hat{g}(\hat{z}^*, \hat{n}^*)$ ) is obtained as 730.478. Also, by dividing  $\hat{g}(\hat{z}^*, \hat{n}^*)$  by  $w_U$ , the relative estimated optimality gap is 1.11%.

In order to verify the solution sensitivity to the available budget, the attempted example is investigated considering different budgets (starting from \$0 to \$2250 by increments of \$250) and the results are plotted in Figure 2. As it can be seen in this figure, the total



**Figure 2.** Solution sensitivity in terms of the available budget.

**Table 3.** Summary of the results at confidence level of 95% for the illustrative example ( $B = \$1500$ ).

Cell 1	1(2), 6(1), 8(1), 9(3), 10(3)
Cell 2	2(1), 3(2), 4(1), 5(1), 7(1)
Estimated production cost	\$44855.988
Estimated subcontracting cost	\$9441.630
Estimated idleness cost	\$769.085
Estimated material handling cost	\$10318.540
Estimated total variable cost ( $\hat{w}_{S'}(\hat{z}^*, \hat{n}^*)$ )	\$65385.243
$\hat{\sigma}_{S'}$	158.415
$\hat{\sigma}_{T,S}$	118.429
Estimated upper bound on $w(\hat{z}^*, \hat{n}^*)$ ( $w_U$ )	\$65617.360
Estimated lower bound on $w^*$ ( $w_L$ )	\$64886.883
Estimated optimality gap ( $\hat{g}(\hat{z}^*, \hat{n}^*)$ )	730.478
Relative estimated optimality gap ( $\hat{g}(\hat{z}^*, \hat{n}^*)/w_U$ )	1.11%
CPU time	351s

variable cost is very sensitive to the budget within the interval [250,1250]. Such a plot can provide a useful outlook to make decision on the budget constraint.

The solution quality and the computational time are highly influenced by the number of scenarios in the SAA problem ( $S$ ) and the number of samples ( $T$ ). Therefore, by solving the attempted example, a sensitivity analysis is conducted based on different values of these two parameters. The CPU times and the relative optimality gaps (regarding  $S=5, 10, 15, 20, 25$ , and  $30$ , and  $T=5, 10, 15, 20, 25$ , and  $30$ ) are depicted in Figures 3 and 4, respectively. From Figure 4, it is seen that at  $T=5$ , there is a significant difference between the relative optimality gaps obtained for  $S=5$  and  $S=30$ . However, as  $T$  increases,  $S$  loses its significance in the solution quality. On the other hand, it can be seen in Figure 3 that at  $T=5$ , the computation times

in terms of various values of  $S$  are not much different, while by increasing  $T$  to 30 samples, the difference in the computation times increases. As a conclusion, by considering a trade-off between the solution quality and the computation time, it is reasonable to solve the mentioned example with  $S=5$  or  $10$  and  $T=30$ .

### 5. Computational results

In this section, 10 numerical examples extracted from the literature are considered to demonstrate the suitability of the proposed stochastic model and the efficiency of the final solution compared to those derived from the literature. The missing data in the source papers are generated and added to the original data set. In these problems, we assume that the stochastic parameters (i.e., the demands and outsourcing costs) are independent random variables with uniform distribution. Table 4 includes the specifications of the problems, as well as the parameter values of the SAA. It should be noted that all the computations are performed on the same PC mentioned in Section 4.

#### 5.1. SAA method versus the expected value problem

In order to highlight the merit of solving the proposed stochastic problem against the expected value problem (the problem in which the stochastic parameters are substituted with their expected values), a common measure called Value of Stochastic Solution ( $VSS$ ) is taken into consideration. Let  $(\bar{z}, \bar{n})$  denote the optimal solution to the expected value problem. Thus, in our problem,  $VSS$  is defined by  $VSS = w(\bar{z}, \bar{n}) - w^*$ , where  $w(\bar{z}, \bar{n})$  is the true objective value of solution  $(\bar{z}, \bar{n})$  and  $w^*$  is the optimum objective value of the true stochastic problem. Now, we can apply the same  $S'$  scenarios (used in the SAA method) to estimate  $VSS$ . Let  $vss$  denote the estimated value of  $VSS$ ; thus,  $vss = \hat{w}_{S'}(\bar{z}, \bar{n}) - \hat{w}_{S'}(\hat{z}^*, \hat{n}^*)$  where  $(\hat{z}^*, \hat{n}^*)$  is the solution derived from the SAA method. Based on these explanations, we made a comparison between these two approaches by solving the numerical examples given in Table 4. A summary of the comparison results, as well as the number of constraints, and positive and discrete variables (including binary and integer variables) in each model is provided in Table 5. In this table, validation model' refers to the model with  $S'$  scenarios, which is used to estimate the true objective value of a given solution. The estimated optimality gap reported in column ' $\hat{g}$ ' has been obtained at confidence level of 95%, that is,  $\alpha = 0.025$ . The CPU time of the SAA method is the overall time of solving  $T$  number of SAA and validation problems. For the sake of time saving, in some problems, we use smaller number of scenarios (to see the SAA parameter values, refer to Table 4). Also, the CPU time of the expected value

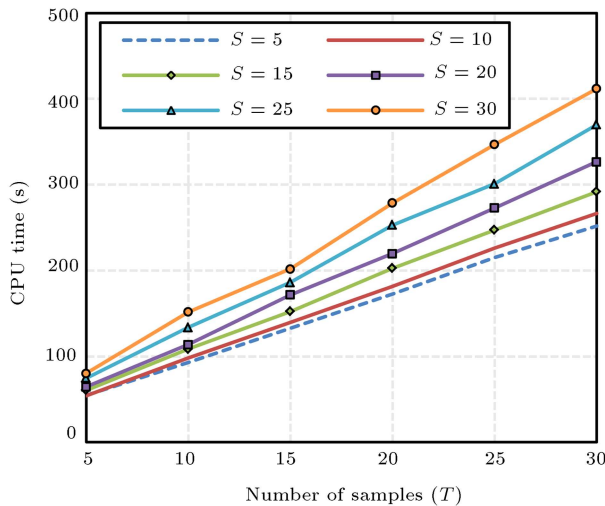


Figure 3. CPU time sensitivity in terms of the number of samples and scenarios.

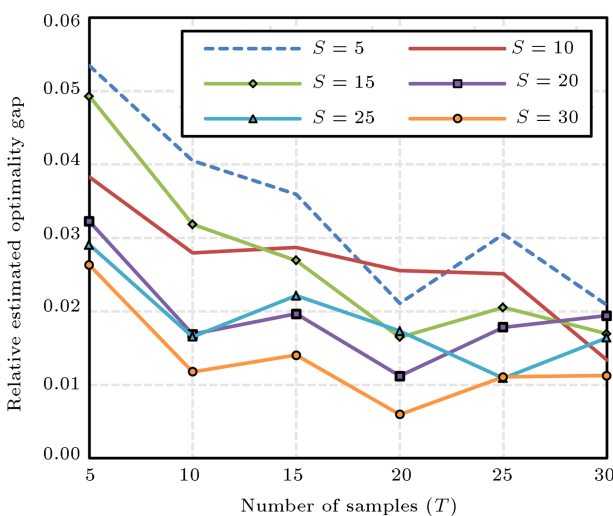


Figure 4. Relative estimated optimality gap sensitivity in terms of the number of samples and scenarios.

**Table 4.** Specifications of the selected problems from the literature.

Problem <sup>†</sup>		SAA parameters							Added parameters to the original data set*									
#	Source	Size (M × P)	∑ <sub>i</sub> R <sub>i</sub>	C <sup>max</sup>	NM	T	S	S'	N <sub>k</sub> <sup>max</sup>	$\bar{d}_i$	$c_{ij}^P$	$c_k^I$	$c_{ikk'l}^A$	$c_{ikk'l}^E$	T <sub>k</sub>	$c_k^M$	B	t <sub>ijk</sub>
1	[5]	7 × 10	23	3	3	30	30	2000	3	<sup>A</sup>	a + U(1,3)	0.01 × c	0.5	3 × c <sup>A</sup> <sub>ikk'l</sub>	<sup>A</sup>	3 × 10 <sup>4</sup>	1 × b	<sup>A</sup>
2	[31]	8 × 13	26	3	3	30	30	2000	3	<sup>A</sup>	a + U(1,3)	0.01 × c	0.2	3 × c <sup>A</sup> <sub>ikk'l</sub>	100	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)
3	[2]	10 × 20	36	1	10	30	30	2000	4	<sup>A</sup>	a + U(1,3)	3 × c	0.3	3 × c <sup>A</sup> <sub>ikk'l</sub>	<sup>A</sup>	<sup>A</sup>	2164	<sup>A</sup>
4	[10]	12 × 20	26	3	5	30	20	2000	3	<sup>A</sup>	a + U(1,3)	0.01 × c	0.2	3 × c <sup>A</sup> <sub>ikk'l</sub>	250	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)
5	[10]	14 × 20	45	3	5	30	10	2000	3	100	a + U(1,3)	0.01 × c	0.25	3 × c <sup>A</sup> <sub>ikk'l</sub>	150	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)
6	[32]	15 × 15	27	3	6	30	30	2000	3	100 <sup>‡</sup>	a + U(1,3)	0.01 × c	0.2	3 × c <sup>A</sup> <sub>ikk'l</sub>	30	3 × 10 <sup>4</sup>	2.5 × b	<sup>A</sup>
7	[33]	17 × 16	402	2	10	30	5	1000	3	<sup>A</sup>	a + U(1,3)	0.01 × c	0.5	3 × c <sup>A</sup> <sub>ikk'l</sub>	200	3 × 10 <sup>4</sup>	2 × b	<sup>A</sup>
8	[10]	17 × 30	61	4	5	30	15	2000	3	<sup>A</sup>	a + U(1,3)	0.01 × c	0.35	3 × c <sup>A</sup> <sub>ikk'l</sub>	250	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)
9	[10]	18 × 30	59	3	7	30	5	2000	3	100 <sup>‡</sup>	a + U(1,3)	0.01 × c	0.4	3 × c <sup>A</sup> <sub>ikk'l</sub>	150	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)
10	[34]	20 × 20	51	5	5	30	5	2000	3	100	a + U(1,3)	0.01 × c	0.5	3 × c <sup>A</sup> <sub>ikk'l</sub>	100	3 × 10 <sup>4</sup>	2 × b	U(0.1,0.9)

<sup>A</sup>The data of this field is available in the source paper.

\* a = number of operations in route j of part i, b = ∑<sub>k</sub> c<sub>k</sub><sup>M</sup>, and c = c<sub>k</sub><sup>M</sup>/T<sub>k</sub>.

† For all the problems, d<sub>i[s]</sub> = U(0, 2 $\bar{d}_i$ ) and c<sub>i[s]</sub><sup>O</sup> = U(1.25, 1.75) × max<sub>j</sub> c<sub>ij</sub><sup>P</sup>, ∀i, s.

‡ In this problem, the demands in the original data set are multiplied by 100.

**Table 5.** Comparison between the solution of the stochastic model and the solution of the expected value model.

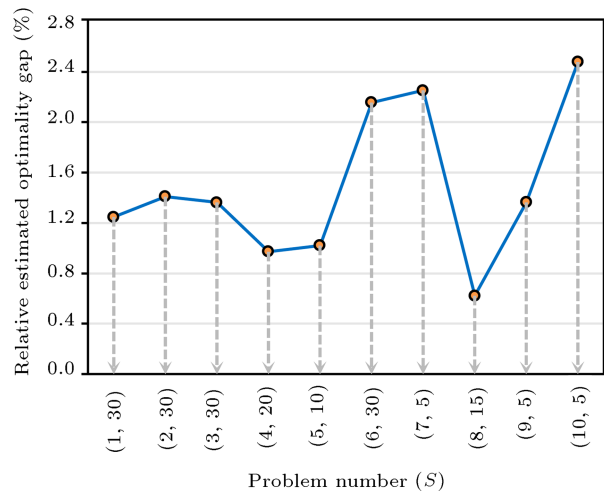
Problem #	Expected value model			Two-stage model			Validation model		SAA method						Expected value approach			
	nc <sup>a</sup>	np <sup>b</sup>	nd <sup>c</sup>	nc	np	nd	nc	np	$\bar{w}_S$	$\hat{w}_{S'}^{SAA}$	$\hat{\sigma}_{T,S}$	$\hat{\sigma}_{S'}^{SAA}$	$\hat{g}$	CPU time	$\hat{w}_{S'}^{EV}$	$\hat{\sigma}_{S'}^{EV}$	CPU time	vss
1	232	104	28	725	1264	28	34001	80001	1880.1	1870.1	9.7	7.0	23.5	142	1870.1	7.0	2	0.0
2	302	132	32	911	1495	32	42001	94001	7594.0	7585.7	31.3	26.5	107.7	174	7680.7	26.5	2	94.9
3	198	112	20	1068	2026	20	60001	132001	57382.2	57447.0	185.4	175.5	788.0	242	57778.2	184.5	4	331.2
4	667	257	48	1275	1359	48	64001	116001	17275.8	17197.9	67.7	54.9	168.2	674	17288.6	51.3	4	90.7
5	900	353	56	1206	1064	56	68001	158001	16608.0	16479.0	101.5	46.3	169.2	2444	16523.5	44.4	10	44.5
6	1025	373	60	1895	2026	60	60001	114001	33580.5	33743.4	158.3	125.7	732.9	3463	34075.6	112.8	33	332.2
7	904	708	51	1036	2448	51	33001	435001	3622.9	3607.5	31.9	16.4	82.0	5496	3850.5	11.8	30	243.1
8	1736	653	85	2394	2165	85	94001	216001	26589.8	26437.9	99.1	58.3	165.2	7635	26590.3	57.0	14	152.4
9	1484	567	72	1676	995	72	96001	214001	24442.9	24317.6	169.8	56.9	333.4	10358	24486.7	51.5	28	169.1
10	2957	1042	120	3117	1406	120	80001	182001	15136.2	15202.5	115.5	38.9	378.8	5473	15407.0	33.6	19	204.5

<sup>a</sup>nc: number of constraints; <sup>b</sup>np: number of positive variables; <sup>c</sup>nd: number of discrete variables.

approach is the overall time of solving the expected value and validation problems. According to the results given in Table 5, we can see that the SAA method has solved the problems with satisfactory estimated gap in relatively acceptable time (even when a smaller number of scenarios were used in the SAA problem). The relative estimated optimality gap of the SAA solution for each problem is also plotted in Figure 5 (relative optimality gap =  $\hat{g}/w_U$ ,  $w_U$  is obtained by Eq. (24)). In this plot, the first and the second numbers in the horizontal axis denote the problem number and the number of scenarios used in the SAA problem, respectively. According to this plot, we can see that the relative estimated optimality gap of all the problems is below 2.5%. On the other hand, according to Table 5, vss in all the problems is a positive number, except for Problem 1. In order to ensure that the true value of the optimality gap (i.e., VSS) is also a positive number, we examine null hypothesis  $H_0 : VSS > 0$  against  $H_1 : VSS \leq 0$ , by obtaining the p-value of test statistic

$$z_0 = vss / \sqrt{\left( (\hat{\sigma}_{S'}^{SAA})^2 + (\hat{\sigma}_{S'}^{EV})^2 \right) / (S' - 1)}.$$

to the test, except for Problem 1, the p-value of the remaining problems was almost 0. This implies that solving the stochastic problem using the SAA method gives a better solution in terms of the expected total



**Figure 5.** Relative estimated optimality gap of the SAA solution.

variable cost than using the expected value approach does.

**5.2. Comparison with the solutions in the literature**

One of the differences between the proposed approach and the ones in the literature is that in this study, there is not any *routing selection* concept; in fact, parts can be produced using multiple routings. On the other hand, in similar problems extracted from the literature, even though there exist multiple routes in producing a particular part, a single route is chosen through a routing selection procedure and, finally, parts are produced by means of a single routing. Therefore, in order to conduct a comparison, the following two approaches are considered:

- A) *Single routing approach based on the solution in the literature:* In this approach, the cell formation result ( $\bar{z}$ ) and the routing selection result are extracted from the literature, and then according to this information and the mean values of the stochastic parameters ( $\mathbf{c}$  and  $\mathbf{d}$ ), a mathematical model is solved to determine the number of each machine type in each cell ( $\bar{n}$ ). Afterwards, based on  $S'$  scenarios used in the SAA method, the validation model is solved to estimate the objective value of  $(\bar{z}, \bar{n})$ , which is denoted by  $w_{S'}^{LS}$ .
- B) *Multiple routings approach based on the solution in the literature:* In this approach, according to the cell formation result ( $\bar{z}$ ) of the solution in the literature and the mean value of the stochastic parameters ( $\mathbf{c}$  and  $\mathbf{d}$ ), a mathematical model (in which the production through multiple routings is allowed) is solved to determine the number of each machine type in each cell ( $\bar{n}$ ). Afterwards, the estimated objective value of solution  $(\bar{z}, \bar{n})$ , (for this approach, let it be denoted by  $w_{S'}^{LM}$ ) is obtained by solving the validation model.

Now, a comparison is carried out between the approach proposed in this research, and approaches

A and B. The solutions of the 10 numerical examples as well as ensuing improvements in consequence of using any of the intended approaches are reported in Table 6. As it can be inferred from the table, the results of solving the problems using the proposed approach show considerable improvements in comparison with approaches A and B, especially A (see columns ‘Imp<sup>1</sup>’ and ‘Imp<sup>2</sup>’ in Table 6). Also, we can see that allowing production through multiple routings (approach B) resulted in improvement in the estimated total variable cost in comparison with single routing approach (see ‘Imp<sup>3</sup>’ in Table 6).

**6. Conclusions**

In this paper, a new problem was attempted to design a CMS considering stochastic part demands and stochastic outsourcing costs. The idea of simultaneous multiple processing routes along with subcontracting was addressed in the proposed problem. According to this idea, each part can be produced simultaneously in multiple processing routes, and unsatisfied part demands (as a result of limited machine capacity or high manufacturing cost) are outsourced. The problem was formulated as a two-stage stochastic program and a solution procedure based on the SAA method was suggested. The objective function aimed at minimizing the summation of the production, outsourcing, material handling, and machine idleness costs. To clarify the problem, an illustrative example was investigated in the case that uniform and normal distributions were used for the outsourcing costs and the part demands, respectively. Based on this example, a sensitivity analysis was carried out to study the behavior of the resulting solution in terms of various available budgets. Furthermore, some experiments were performed to evaluate the solution quality and the computation time in terms of the number of samples and scenarios in the SAA method. Then, we used 10 numerical examples from the literature to demonstrate

**Table 6.** Comparison results between the proposed approach and the conventional approaches.

Problem #	Proposed approach							Literature solution						Imp <sup>1</sup> (%)	Imp <sup>2</sup> (%)	Imp <sup>3</sup> (%)					
								Multiple routings approach			Single routing approach										
	TP	TO	TU	TH	$w_{S'}^{SAA}$	$\hat{\sigma}_{S'}$		TP	TO	TU	TH	$\hat{w}_{S'}^{LM}$	$\hat{\sigma}_{S'}$				TP	TO	TU	TH	$\hat{w}_{S'}^{LS}$
1	1070	383	157	260	1870	7	1076	381	156	295	1908	7	975	572	348	178	2071	7	1.99	9.72	7.89
2	5267	1341	346	632	7586	27	5003	1641	439	597	7681	26	4885	2298	803	409	8395	27	1.24	9.64	8.51
3	50171	1668	1127	4481	57447	176	49979	1753	2073	4510	58315	172	49950	2150	2160	4474	58734	173	1.49	2.19	0.71
4	9990	4879	445	1884	17198	55	11632	2530	957	2222	17341	48	10963	3739	1041	2005	17748	50	0.82	3.10	2.30
5	10310	2961	566	2642	16479	46	10387	2871	559	3206	17023	46	8918	5915	798	2344	17974	50	3.20	8.32	5.29
6	15811	13030	915	3987	33743	126	18375	9133	2013	4639	34160	112	16885	11966	1547	4353	34751	123	1.22	2.90	1.70
7	1945	981	124	557	3607	16	2190	406	538	962	4096	12	1428	2108	880	612	5028	18	11.93	28.25	18.53
8	18766	2024	918	4730	26438	58	18687	2468	1167	4267	26590	57	18470	3155	1908	3977	27510	56	0.57	3.90	3.34
9	11176	9111	400	3631	24318	57	13753	5267	772	4831	24624	52	13398	6298	1133	4373	25201	53	1.24	3.50	2.29
10	8272	3841	762	2328	15203	39	9140	2517	1417	2514	15587	33	8525	3951	1869	1900	16245	35	2.47	6.42	4.05

Imp<sup>1</sup> = 100 ( $\hat{w}_{S'}^{LM} - \hat{w}_{S'}^{SAA}$ ) /  $\hat{w}_{S'}^{LM}$ ; Imp<sup>2</sup> = 100 ( $\hat{w}_{S'}^{LS} - \hat{w}_{S'}^{SAA}$ ) /  $\hat{w}_{S'}^{LS}$ ; Imp<sup>3</sup> = 100 ( $\hat{w}_{S'}^{LS} - \hat{w}_{S'}^{LM}$ ) /  $\hat{w}_{S'}^{LS}$ ;

the performance of the solution method. The computation results indicated that the SAA method could produce efficient solutions with a satisfactory gap in a relatively reasonable computational time. Also, a common measure called *VSS* was used to highlight the advantage of the proposed stochastic approach over the expected value approach. The results demonstrated that solving the stochastic problem using the SAA method was reasonably justified. Moreover, through these numerical examples, a comparison was executed between the obtained solutions and the ones reported in the literature. The comparison results showed that allowing production through multiple routings could lead to an improvement in the estimated total variable cost as compared to the single routing approach.

Finally, to provide some directions for future research, the following issues are recommended:

- In this research, we used the SAA method to solve a stochastic CF problem. Even though this method is efficient in solving small and medium-size instances, for larger problems, the computational time is a concern. To overcome such a difficulty, addressing the Bender's decomposing algorithm in the SAA method could be a possible remedy;
- Subcontracting, simultaneous multiple process routings, and machine duplication were the main issues that we addressed in the CMS design problem. However, these issues could be considered in an integrated problem in which the layout and scheduling problems are also incorporated;
- In order to deal with the uncertainty, we used a two-stage stochastic programming model. However, multi-stage stochastic programming methods could also be used to formulate the problem in an uncertain and dynamic environment.

## References

1. Wemmerlöv, U. and Hyer, N. "Procedures for the part family/machine group identification problem in cellular manufacture", *Journal of Operations Management*, **6**, pp. 125-147 (1986).
2. Solimanpur, M., Vrat, P., and Shankar, R. "A multi-objective genetic algorithm approach to the design of cellular manufacturing systems", *International Journal of Production Research*, **42**, pp. 1419-1441 (2004).
3. Forghani, K., Mohammadi, M., and Ghezavati, V. "Integrated cell formation and layout problem considering multi-row machine arrangement and continuous cell layout with aisle distance", *The International Journal of Advanced Manufacturing Technology*, **78**, pp. 687-705 (2015).
4. Wu, T., Chung, S., and Chang, C. "Hybrid simulated annealing algorithm with mutation operator to the cell formation problem with alternative process routings", *Expert Systems with Applications*, **36**, pp. 3652-3661 (2009).
5. Kao, Y. and Lin, C. "A PSO-based approach to cell formation problems with alternative process routings", *International Journal of Production Research*, **50**, pp. 4075-4089 (2012).
6. Arkat, J., Abdollahzadeh, H., and Ghahve, H. "A new branch and bound algorithm for cell formation problem", *Applied Mathematical Modelling*, **36**, pp. 5091-5100 (2012).
7. Boutsinas, B. "Machine-part cell formation using bi-clustering", *European Journal of Operational Research*, **230**, pp. 563-572 (2013).
8. Wu, X., Chu, Ch-H., Wang, Y., and Yue, D. "Genetic algorithms for integrating cell formation with machine layout and scheduling", *Computers and Industrial Engineering*, **53**, pp. 277-289 (2007).
9. Krishnan, K., Mirzaei, S., Venkatasamy, V., and Pillai, V. "A comprehensive approach to facility layout design and cell formation", *International Journal of Advanced Manufacturing Technology*, **59**, pp. 737-753 (2012).
10. Chang, C., Wu, T., and Wu, C. "An efficient approach to determine cell formation, cell layout and intracellular machine sequence in cellular manufacturing systems", *Computers and Industrial Engineering*, **66**, pp. 438-450 (2013).
11. Mohammadi, M. and Forghani, K. "A novel approach for considering layout problem in cellular manufacturing systems with alternative process routings and subcontracting approach", *Applied Mathematical Modelling*, **38**, pp. 3624-3640 (2014).
12. Solimanpur, M., Vrat, P., and Shankar, R. "A heuristic to minimize makespan of cell scheduling problem", *International Journal of Production Economics*, **88**, pp. 231-241 (2004).
13. Elmi, A., Solimanpur, M., Topaloglu, S., and Elmi, A. "A simulated annealing algorithm for the job shop cell scheduling problem with intercellular moves and reentrant parts", *Computers and Industrial Engineering*, **61**, pp. 171-178 (2011).
14. Arkat, J., Farahani, M., and Ahmadizar, F. "Multi-objective genetic algorithm for cell formation problem considering cellular layout and operations scheduling", *International Journal of Computer Integrated Manufacturing*, **25**, pp. 625-635 (2012).
15. Li, D., Meng, X., Li, M., and Tian, Y. "An ACO-based intercell scheduling approach for job shop cells with multiple single processing machines and one batch processing machine", *Journal of Intelligent Manufacturing*, **27**, pp. 283-296 (2016).
16. Wang, T., Wu, K., and Liu, Y. "A simulated annealing algorithm for facility layout problems under variable demand in cellular manufacturing systems", *Computers in Industry*, **46**, pp. 181-188 (2001).
17. Jeon, G. and Leep, H. "Forming part families by using genetic algorithm and designing machine cells under demand changes", *Computers and Operations Research*, **33**, pp. 263-283 (2006).

18. Tavakkoli-Moghaddam, R., Javadian, N., Javadi, B., and Safaei, N. "Design of a facility layout problem in cellular manufacturing systems with stochastic demands", *Applied Mathematics and Computation*, **184**, pp. 721-728 (2007).
19. Schaller, J. "Designing and redesigning cellular manufacturing systems to handle demand changes", *Computers and Industrial Engineering*, **53**, pp. 478-490 (2007).
20. Safaei, N., Saidi-Mehrabad, M., and Babakhani, M. "Designing cellular manufacturing systems under dynamic and uncertain conditions", *Journal of Intelligent Manufacturing*, **18**, pp. 383-399 (2007).
21. Arıkan, F. and Güngör, Z. "Modeling of a manufacturing cell design problem with fuzzy multi-objective parametric programming", *Mathematical and Computer Modelling*, **50**, pp. 407-420 (2009).
22. Ghezavati, V. and Saidi-Mehrabad, M. "Designing integrated cellular manufacturing systems with scheduling considering stochastic processing time", *International Journal of Advanced Manufacturing Technology*, **48**, pp. 701-717 (2010).
23. Das, K. and Abdul-Kader, W. "Consideration of dynamic changes in machine reliability and part demand: a cellular manufacturing systems design model", *International Journal of Production Research*, **49**, pp. 2123-2142 (2011).
24. Ghezavati, V. and Saidi-Mehrabad, M. "An efficient hybrid self-learning method for stochastic cellular manufacturing problem: A queuing-based analysis", *Expert Systems with Applications*, **38**, pp. 1326-1335 (2011).
25. Rabbani, M., Jolai, F., Manavizadeh, N., Radmehr, F., and Javadi, B. "Solving a bi-objective cell formation problem with stochastic production quantities by a two-phase fuzzy linear programming approach", *International Journal of Advanced Manufacturing Technology*, **58**, pp. 709-722 (2012).
26. Forghani, K., Mohammadi, M., and Ghezavati, V. "Designing robust layout in cellular manufacturing systems with uncertain demands", *International Journal of Industrial Engineering Computations*, **4**, pp. 215-226 (2013).
27. Heragu, S. and Chen, J. "Optimal solution of cellular manufacturing system design: Benders' decomposition approach", *European Journal of Operational Research*, **107**, pp. 175-192 (1998).
28. Ruszczyński, A. and Shapiro, A. *Handbooks in Operation Research and Management Science: Stochastic Programming*, Elsevier, Amsterdam (2003).
29. Shapiro, A. and Homem-de-Mello, T. "On rate of convergence of Monte Carlo approximations of stochastic programs", *SIAM Journal of Optimization*, **11**, pp. 70-86 (2000).
30. Mak, W.K., Morton, D.P., and Wood, P.K. "Monte Carlo bounding techniques for determining solution quality in stochastic programs", *Operations Research Letters*, **24**, pp. 47-56 (1999).
31. Kazerooni, M., Luong, H., and Abhary, K. "A genetic algorithm based cell design considering alternative routing", *International Journal of Computer Integrated Manufacturing Systems*, **10**, pp. 93-107 (1997).
32. Ramabhatta, F.V. and Nagi, R. "An integrated formulation of manufacturing cell formation with capacity planning and multiple routings", *Annals of Operations Research*, **77**, pp. 79-95 (1998).
33. Yin, Y. and Yasuda, K. "Manufacturing cells' design in consideration of various production factors", *International Journal of Production Research*, **40**, pp. 885-906 (2002).
34. Chan, F.T.S., Lau, K.W., Chan, P.L.Y., and Choy, K.L. "Two-stage approach for machine-part grouping and cell layout problems", *Robotics and Computer-Integrated Manufacturing*, **22**, pp. 217-238 (2006).

## Biographies

**Masoud Mahootchi** received BSc and MSc degrees from the Industrial Engineering Department of Amirkabir University of Technology in 1993 and 1997, respectively. He also received his PhD degree in Systems Design Engineering from the University of Waterloo in 2009 and is currently a Faculty Member at Amirkabir University of Technology, Tehran, Iran. He focused on storage management under uncertainty using nonlinear and soft computing methods during his PhD. His research areas include stochastic optimization, portfolio optimization, dynamic programming, and reinforcement learning.

**Kamran Forghani** obtained his BSc degree in Industrial Engineering, in 2009, from the Islamic Azad University of Bonab, East Azarbaijan, Iran, and his MSc degree from Kharazmi University, Tehran, Iran, in 2013, in the same subject. He is currently a PhD candidate in Industrial Engineering at Amirkabir University of Technology in Tehran, Iran. His research interests include cellular manufacturing systems, facility layout, production planning, scheduling, hybrid algorithms, and stochastic programming.

**Mehdi Abdollahi Kamran** received his MSc and BSc degree in Industrial Engineering from Amirkabir University of Technology. He is a Faculty Member of Urmia University of Technology (UUT). He is also a PhD candidate in Industrial Engineering at Amirkabir University of Technology (AUT) in Tehran. His interests include facilities location planning, optimization algorithms, MCDM, supply chain, simulation, and statistics.