Stochastic cellular manufacturing system design subject to maximum acceptable risk level

Gokhan Egilmez, Gürsel A. Süer *, Jing Huang

ISE Department, Ohio University, Athens, OH 45701, USA

A R T I C L E   I N F O

Article history:
Received 2 July 2011
Received in revised form 8 May 2012
Accepted 27 May 2012
Available online 13 June 2012

Keywords:
Cellular manufacturing system design
Non-linear modeling
Stochastic optimization
Simulation
Model validation

A B S T R A C T

In this study, a non-linear mathematical model is proposed to solve the stochastic cellular manufacturing system (CMS) design problem. The problem is observed in both machine and labor-intensive cells, where operation times are probabilistic in addition to uncertain customer demand. We assume that processing times and customer demand are normally distributed. The objective is to design a CMS with product families that are formed with most similar products and minimum number of cells and machines for a specified risk level. Various experiments are carried out to study the impact of risk level on CMS design. As the risk level increases, lower number of cells and product families are formed and average cell utilization increases. However, this leads to poor performance in cells, where standard deviations of capacity requirements are high. Later, the deterministic approach proposed by Suer, Huang, and Sripathi (2010) and the proposed stochastic model with various risk levels are compared. Both of the models’ results are simulated with Arena Simulation Software. Simulation is performed to validate models and assess the performance of designed CMSs with respect to following measures: cell utilization, WIP, total waiting time and total number waiting. Stochastic CMS design with 10% risk formed a better CMS in all of the performance measures according to the results obtained from simulation experiments.

1. Introduction

Manufacturing strategies and approaches have significantly changed in recent decades in comparison to the first half of the 20th century. One of the changes that has been employed and implemented by numerous companies in recent decades is the Group Technology (GT) and Cellular Manufacturing (CM) concepts. Group technology is a general philosophy, where similar items are brought together considering a critical attribute and the same solution is applied to the entire group thus improving the productivity of the system.

Cellular manufacturing is an application of GT to the manufacturing world. In a cellular manufacturing system (CMS), similar products are grouped into product families and the required machines are assigned to manufacturing cells to produce the corresponding product families. In this respect, a cell is a small manufacturing unit designed to have people, dissimilar equipment and machines together to produce like products resulting in lower leadtimes, work-in-process inventory (WIP), setup times and work-force (Wemmerlov & Johnson, 1997). See Burbidge and Wei (1992) for more detailed explanation of the benefits of implementing CMS.

Although there are significant benefits that can be achieved when CMS is employed, there are some disadvantages of CMS implementation such as being less flexible to rapid changes in product mix and demand (Ang & Willey, 1984; Wei & Gaither, 1990; Satoglu & Suresh, 2009). In addition to these cons, the major concern about Cellular Manufacturing (CM) is the reduced machine utilization due to the dedication of machines and cells to certain product families (Suer & Ortega, 1996). Moreover, overutilization or underutilization of cells can be another complicated issue when demand of each product is uncertain. Due to such difficulties as inefficient cell and machine utilization and poor production control associated with highly probabilistic demand (Suer et al., 2010), stochastic behavior of demand should be taken into consideration prior to CMS design.

Manufacturing cells can be defined as either machine-intensive cell or labor-intensive. In machine-intensive cells, the operator involvement is limited and the operation is mostly influenced by the machine performance. The operators usually load the raw material or semi-product, unload it from the machine and perform quality control. In these environments, processing times may not greatly vary from one unit of the job to the next unit of the same job as the machines increasingly have a better repeatability feature. On the other hand, in labor-intensive manufacturing cells, operations are mainly carried out by the operators and the processing time of an operation can vary significantly from one unit of the job to the next unit depending on the operator and even for the same operator.
In this paper, both demand and processing times are considered as probabilistic. A non-linear stochastic mathematical model is developed for CMS design. Experimentation is carried out with deterministic model given by Suer et al. (2010) and the proposed stochastic mathematical model. Subsequently, the results are simulated with Arena software to validate the mathematical model and observe performance of CMS in terms of cell utilization, WIP, waiting times and total number of units waiting.

The remainder of the paper is organized as follows. In Section 2, literature is reviewed. In Section 3, the manufacturing system studied is explained. In Section 4, the calculation of similarity coefficients and capacity requirements and the deterministic capacitated P-median model are discussed. In Section 5, stochastic CMS design is introduced, where bottleneck machine identification and probabilistic capacity requirements are discussed and the proposed non-linear stochastic mathematical model is described. An example problem is solved and provided in Section 6. In Section 7, the validation of models and performance analysis are explained. The experimentation and the results are presented in Section 8. Finally, the concluding remarks are made and the future work is given in Section 9.

2. Literature review

Many different optimizing procedures have been used to solve CMS design problem. The majority of the procedures are based on mathematical optimization. In addition to mathematical models, simulation, heuristics and meta-heuristics are also used in some of these works. While most of the studies in the literature have addressed the deterministic CMS design problem, less attention is paid to the problems which include uncertainty in such parameters as demand and processing times. The literature is classified into two sections as works that consider deterministic behavior of the problem and also stochastic behavior of the problem.

2.1. Deterministic CMS design

Mathematical modeling is employed in numerous studies in the literature. Purcheck (1974) studied the group technology problem and applied linear programming. Kusiak (1987) compared matrix and integer programming (the p–median) models and showed the improvement in the quality of cell formation. Shhtub (1989) proposed the Generalized Assignment Problem (GAP) as the equivalence of simple cell-formation problem. Rajamani, Singh, and Aneja (1990) developed three integer programming models for CMS design and analyzed the impact of alternative process plans. Wei and Gaither (1990) used an optimal 0–1 integer programming model to provide an analysis for the CMS design problem to minimize the cost of manufacturing exceptional products outside the cellular system, subject to machine capacity constraints.


Sofianopoulou (1999) developed a mathematical model and a two-phased simulated annealing algorithm to solve the problem of grouping machines into cells. A unique product process plan for each product is selected, where machine duplication, operation sequence constraints and several design requirements exist. Akturk and Turkcan (2000) proposed an integrated algorithm to solve the product-family and machine-cell formation problem by simultaneously considering the within–cell layout problem. The efficiency of both individual cells and the overall system in monetary terms are considered and a local search heuristic is provided. Albadawi, Bashir, and Chen (2005) proposed a two-phase mathematical model that includes the identification of machine cells by applying factor analysis and the assignment of products to the machine cells. They also applied their model to a case study.

Some of the works used simulation in addition to mathematical optimization in cell formation problems. Kamrani et al. (1995) proposed a simulation-based methodology considering design and manufacturing attributes to form manufacturing cells. A three phased-hierarchical methodology is used, namely: (1) part dissimilarity-based cell formation, (2) grouping of machines into manufacturing cells and (3) simulation. The proposed approach is beneficial in terms of designing, analyzing, optimizing and justifying the cellular manufacturing system considered. Hachicha, Masmoudi, and Haddar (2007) provided a simulation-based improvement approach considering exceptional elements and the effect of intercellular movement. The initial configuration of machines and parts in cells are known and the intercell transfer is allowed since some process routes require parts visit different cells. Based on the cluster given, the designed system is improved via simulation such that the impact of remaining exceptional elements on system performance is minimized. Mean transfer, machining, waiting and flow times are considered as performance measure along with the cost of intercell movement. For more examples of works include simulation, see Shang and Tadikamalla (1998), Habchi and Berchet (2003), and Masmoudi (2006).

In the last decade, metaheuristics such as Genetic Algorithms (GA), Simulated Annealing (SA), Tabu Search (TS) have been widely applied to the cell formation problem. See Moon, Gen, and Suer (1999), Asokan, Prabhakaran, and Satheesh (2001), Suer, Pena, and Vazques (2003), Cao and Chen (2004), Jayaswal and Adil (2004), and Solimanpur, Vrat, and Shankar (2004).

2.2. Stochastic CMS design

Although there are several studies, which considers the deterministic CMS design problem in literature, only a handful consider the stochastic parameters such as demand, processing time, and capacity requirements. Seifoddini (1990) focused on the uncertainty of the product mix for a single period and proposed a probabilistic model for machine cell formation to minimize the expected inter-cell material handling costs of the system. Wicks and Reasor (1999) used forecasting techniques and solved the multi-period cell formation problem with GA (Wicks & Reasor, 1999). Saad (2003) focused on the reconfiguration issues in manufacturing systems and provided several sub-modules, namely: configuration and reconfiguration module, loading module, and simulation-based scheduling module (Saad, 2003). Saidi-Mehrabad and Ghezavati (2009) applied queuing theory to the CMS design problem associated with uncertainty issues (Saidi-Mehrabad & Ghezavati, 2009). Each machine is considered as server and each product is assumed as customer. The objective is to minimize the idleness costs for servers, the total cost of sub-contracting for exceptional elements and the cost of resource underutilization. Suer et al. (2010) proposed an alternative CMS design approach, “layered manufacturing system”, to deal with the uncertainty in product demand and production rates. The proposed layered CMS consists of three
types of cells, namely: dedicated, shared and remainder cells. In their study, the generalized p-median model by Kusiak (1987) is modified to one which considers the utilization of cells as well as the product similarities. In this paper, the CMS design based on deterministic capacitated p-median model is used for comparison.

In this study, a stochastic non-linear capacitated p-median mathematical model is developed to deal with the probabilistic demand, production rates and thus capacity requirements. The proposed model is compared with Süer et al.’s (2010) deterministic capacitated p-median model. The obtained cell configurations from both deterministic and proposed stochastic mathematical models are simulated with Arena Simulation Software to validate the designed CMS and assess the overall performance based on selected metrics, namely, cell utilization, the number of machines, work-in-process inventory, average queue length and average waiting time.

3. The manufacturing system studied

The data is obtained from a jewelry company however this problem has been also observed in pharmaceutical, food and medical device industries. There are thirty products and eighteen machines. Each product has to be processed on several machines depending on the operational route and machine duplication in the same cell is not allowed. However, there might be copies of the same machine in different cells depending on the product’s route. Since each product’s route represents a unidirectional flow, the cell configuration is flow shop. The machine with the maximum expected processing time among all machines on a route is considered to be the bottleneck machine. In this study, the bottleneck machine is the 18th machine for all products. The incidence matrix, which includes the expected processing times, annual expected demand ($\mu$) and standard deviation of demand ($\sigma$) is shown in Table 1.

Cells are considered to be independent from each other (dedicated to one product family) and inter-cell transfer of products is not allowed. Independent cell configuration is necessary in certain systems such as pharmaceutical, food and medical device industries. Independent cell configuration is inevitable since the allowance of inter-cell movement may cause serious health and hygiene problems and long set-up between products (cleaning the production area can also be considered within setup). On the other hand, in some systems, inter-cell movement can also create problems in terms of manageability of cells. In a food manufacturing plant, each manufacturing cell is assigned to a person and that person was in charge for the most of the operational and tactical decisions. In addition to setup and cleaning costs, intercell movement may cause managerial problems as well.

In this paper, each dedicated cell is responsible for the production of only one family. Hence, each product can only be assigned to one family. The single-piece flow is adapted for moving products in cells. Annual capacity is taken as 2000 h (50 weeks/yr x 40 h/week). The annual demand and processing time for each product are assumed as normally distributed. The standard deviation of processing time for each operation is assumed to be 10% of the mean for each operation. The objective is to design a cellular manufacturing system with the minimum number of cells while maximizing similarity among products in families.

4. Capacitated CMS design in a deterministic environment

In a deterministic case, annual demand, processing times and therefore capacity requirements are known exactly. Süer et al. (2010) proposed the capacitated P-Median model to address the cell utilization and similarity trade-off for the deterministic CMS design. The hierarchical framework used includes the identification of similarities, determination of capacity requirements and solving the deterministic capacitated p-median model.

<table>
<thead>
<tr>
<th>Parts</th>
<th>Machines</th>
<th>Annual demand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\mu$</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>16</td>
</tr>
</tbody>
</table>
4.1. Identification of similarities and determination of capacity requirements

The similarity matrix is constructed based on the similarity coefficients. Suer et al. (2010) modified McAuley’s (1972) similarity coefficient definition to find the similarities among products. The similarity coefficients are calculated via the suggested equation by Suer et al. (2010) as shown in Eq. (1). The machine similarity (MS) between products \(i\) and \(j\) is the ratio of the common number of machines to the total number of machines to produce both.

\[
MS_{ij} = \frac{\text{Number of machines processing parts } i \text{ and } j}{\text{Number of machines processing parts either } i \text{ or } j}
\]  

Once similarities are calculated, the capacity requirements are obtained via Eq. (2), where \(P_i\) is processing time of bottleneck machine for part \(i\) in minutes, \(D_i\) is annual demand for part \(i\) in units and \(CR_i\) is annual capacity requirement for product \(i\) in hours.

\[
CR_i = \frac{D_i \times P_i}{60}
\]  

4.2. Deterministic capacitated P-median model

The deterministic approach only considers the mean capacity requirements and the machine similarities. Product families and cell formations are determined considering cell utilization and similarity coefficients. The indices, parameters and decision variables for the model are listed as follows:

<table>
<thead>
<tr>
<th>Indices:</th>
<th>Parameters:</th>
<th>Decision variables:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) (j)</td>
<td>(S_{ij}) Similarity coefficient between product (i) and (j)</td>
<td>(X_{ij}) 1, if product (i) is assigned to family (j); 0, otherwise</td>
</tr>
<tr>
<td>Product index</td>
<td>Product index and family/cell index</td>
<td></td>
</tr>
<tr>
<td>(CR_i) Capacity requirement for product (i) in hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n) Number of products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(TU) Upper limit for cell capacity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The objective function is shown in Eq. (3). It maximizes the total similarity among products that are formed as families and are produced in dedicated cells, while minimizing the number of cells. Eq. (4) limits the cell utilization up to the cell capacity. Eq. (5) forces each product to be assigned to a cell. Eq. (6) guarantees the assignment of each product to only one of the cells that are opened. Eq. (7) determines whether product \(i\) can is assigned to cell \(j\) or not.

**Objective function:**

\[
\max Z = \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} \times X_{ij} - \sum_{j=1}^{n} X_{jj}
\]  

**Subject to:**

\[
\sum_{i=1}^{n} CR_i \times X_{ij} \leq TU \quad j = 1, \ldots, n
\]  

\[
\sum_{j=1}^{n} X_{ij} = 1 \quad i = 1, \ldots, n
\]  

\[
X_{ij} \leq X_{jj} \quad j = 1, \ldots, n \text{ and } i = 1, \ldots, n
\]  

\[
X_{ij} \in [0, 1]; \quad X_{ij} \text{ integer}
\]

5. The proposed solution methodology: stochastic CMS design

The model discussed in the previous section assumes that processing times and demand and hence capacity requirements are all deterministic. In this paper, we relax this assumption and modify the model considering normally distributed demand and processing times. The similarity coefficients used in the deterministic model are kept the same in the proposed stochastic approach. This section consists of three parts: first part explains how the probabilistic capacity requirements are determined; proposed stochastic non-linear mathematical model is explained in the second part and an example problem is provided in the third part.

5.1. Capacity requirements in the presence of stochastic demand and processing times

In the deterministic case, the capacity requirement of a product is calculated via multiplying its demand and processing time (Eq. (2)). However, in the stochastic case, since both demand and processing time are probabilistic, the product of these two random variables becomes probabilistic and requires statistical analysis to find the probability density functions (pdf) of the capacity requirements. To find the fitted distribution (pdf) of the capacity requirement of product \(i\), statistical analysis is performed with Arena Input Analyzer software. The framework of the analysis is shown in Fig. 1.

5.2. Stochastic capacitated P-median model

The proposed solution methodology is based on the mathematical model given in Section 4. The objective function (Eq. (3)), constraints given in Eqs. (5) and (6), and decision variable definitions (Eq. (7)) remain the same. However, the first constraint (Eq. (4)) which limits the allocation of products to the cells in terms of capacity is changed since capacity requirements are now stochastically defined. This section discusses how new constrained is formulated.

For a product to be assigned to a family and thus to a cell, the total capacity requirements after assignment must be less than or equal to the available capacity. A product can be assigned to only one of the opened cells. The probability that capacity requirements will not exceed available capacity when the first product, say product 1 is assigned to a cell is computed as \(P(CR_1 \leq TU)\), where \(CR_1\) is the capacity requirement for product 1 and \(TU\) is the available cell capacity. Similarly, this expression now can be extended to include the second product, product 2 as well. The probability that the total capacity requirements for both products 1 and 2 will not exceed the available capacity when assigned to

**Fig. 1.** The framework of input analysis.
the same cell is given as \(P(CR_1 + CR_2 \leq TU)\). We can specify the minimum probability that capacity requirements will be less than or equal to the available capacity as \((1 - \alpha)\) in our mathematical model and then include this as a constraint for the first product as shown in Eq. (8). In this relation, factor \((\alpha)\) indicates the maximum acceptable risk level that the decision maker would like to take when assigning products to families, i.e., factor \((\alpha)\) indicates the maximum acceptable probability that capacity requirements will exceed the capacity available

\[ p(CR_1 \leq TU) \geq (1 - \alpha) \]  

(8)

Now this constraint type can be re-written to include mean capacity requirements and the decision variable for the first product assigned. For example, the constraint for the first product is described by Eq. (9), where \(\mu_{CR_i}\) is the mean capacity requirements for product 1 and \(X_j\) is the decision variable indicating whether product 1 will be assigned to cell \(j\) or not

\[ p(\mu_{CR_1} + X_{ij} \leq TU) \geq (1 - \alpha) \]  

(9)

Similarly, the assignment of the second product (product 2) is the sum of mean capacity requirements of first two products (if both assigned) and this is illustrated by Eq. (10)

\[ p(\mu_{CR_1} + X_{ij} + \mu_{CR_2} + X_{ij} \leq TU) \geq (1 - \alpha) \]  

(10)

The relations given by Eqs. (9) and (10) can be converted to standard normal distribution equations for products 1 and 2 as shown by Eqs. (11) and (12), respectively

\[ p\left( Z_{ij} \leq \frac{(TU - \mu_{CR_1} - X_{ij})}{\sigma_1^2 + \sigma_{ij}} \right) \geq (1 - \alpha) \]  

(11)

\[ p\left( Z_{ij} \leq \frac{(TU - \mu_{CR_1} + \mu_{CR_2} + X_{ij})}{\sigma_1^2 + \sigma_{ij}} \right) \geq (1 - \alpha) \]  

(12)

The generalized constraint is given in Eq. (13), where \(\mu_{CR_1}\) product is assigned to cell \(j\).

\[ p\left( Z_{ij} \leq \frac{(TU - \sum_{i=1}^{n} \mu_{CR_i} + X_{ij})}{\sqrt{\sum_{i=1}^{n} \sigma_i^2 + \sigma_{ij}}} \right) \geq (1 - \alpha) \]  

(13)

Finally, Eq. (14) replaces the cell capacity constraint given Eq. (4) in the deterministic mathematical model.

\[ p\left( Z_{ij} \leq \frac{(TU - \sum_{i=1}^{n} \mu_{CR_i} + X_{ij})}{\sqrt{\sum_{i=1}^{n} \sigma_i^2 + \sigma_{ij}}} \right) \geq (1 - \alpha) \quad j = 1, 2, \ldots, n \]  

(14)

This constraint brings non-linearity to the mathematical model and it is solved by using Lingo software.

### 5.3. Example problem

In this section, an example problem with 10 products is provided to illustrate the proposed stochastic non-liner mathematical model. The first 10 products of 30 products that are shown in incidence matrix are used in the example problem. The similarity matrix and probabilistic capacity requirements (statistical distributions) of products are calculated via the Eq. (1) and statistical analysis, respectively. The similarity for the 10 products matrix is shown in Table 2. The capacity requirements for each product are given in Table 3.

By using similarity matrix and capacity requirements with the proposed approach, the optimal cell configuration based on the 10% risk level is obtained and shown in Table 4.

Based on the obtained cell formation, the expected and variance of cell utilizations are calculated via Eqs. (15) and (16), where \(i\) and \(m\) are the product and cell indices, \(n_i\) is the number of products assigned to cell \(m\) and \(CR_m\) is the capacity requirement of cell \(m\)

\[ E[CR_m] = \sum_{i=1}^{n_i} E[CR_{im}] \quad for \ m = 1 \ldots M \]  

(15)

\[ Var[CR_m] = \sum_{i=1}^{n_i} Var[CR_{im}] \quad for \ m = 1 \ldots M \]  

(16)

By using the Eqs. (15) and (16), expected and variance of cell requirements are calculated. In this problem, 10% risk level is used. To be able to compare the % risk and actual risk of cell overutilization, Z value is calculated for each cell by using Eq. (17). The related results of calculations are provided in Table 5. According to the results, the greatest probability of exceeding available capacity is obtained as 7.2%, which is less than the risk taken prior to experiment (\(\alpha = 10\%\)).

\[ Z_m = \frac{(2000 - E[CR_m])}{\sqrt{Var[CR_m]}} \quad for \ m = 1 \ldots M \]  

(17)

### 6. Model validation

Model validation is one of the most crucial steps of any model-based methodology. In fact, when demand, processing times and capacity requirements are probabilistic, it is must to validate the proposed approach. According to the Barlas’ classification of models (Barlas, 1996), the modeling (non-linear stochastic mathematical optimization) used in this study is a white-box type, in other words causal descriptive. Therefore, the model must not only reproduce or predict the behavior of system analyzed but also explain how the behavior is generated. To answer the behavior generated, the designed CMS’ performance is assessed with respect to the CMS performance measures obtained from literature. Wemmerlov and Johnson surveyed the reasons why the firms had established manufacturing cells among 46 firms (Wemmerlov & Johnson, 1997). The top five
significant ones are to reduce throughput time, WIP, response time to customer, move distances and move times, and to improve part/product quality.

In this study, simulation is employed for the validation of the proposed approach and performance analysis of CMS with respect to the measures, namely: cell utilization, WIP, waiting time and the number waiting. The results of 30-product dataset are used in validation. A hierarchical framework is developed to validate the proposed approach with simulation. As shown in Fig. 2. The validation procedure consists of three steps: (1) Input analysis, (2) Simulation model development and (3) Simulation experiments and statistical analysis.

6.1. Input analysis

Prior to building the simulation model, inputs are prepared. First of all, annual stochastic product demands are converted to interarrival times (IAT). The same demand data generated for the mathematical model is used in this analysis. System capacity (2000 h) is divided by the demand data of each product and the obtained data is analyzed via the Input Analyzer module of Arena. According to the Kolmogorov–Smirnov and Chi-Square tests’ results, best fitted distribution is determined as the “interarrival time distribution” for each product.

6.2. Simulation model development

After determining the probabilistic distributions for the interarrival times, they are entered into simulation model. Secondly, cell and family formation information obtained from the mathematical model and production routes are defined. Thirdly, performance measures are defined and included in the simulation model.

6.3. Simulation runs and analysis

In this phase, the simulation models are run to let the simulation model reach to the steady state. Ten replications are made for each CMS configuration. After simulation experiments are performed, results are analyzed with ANOVA. One-way ANOVA is employed to compare the cell utilization results obtained from mathematical models and simulation replications. The ANOVA test results for deterministic model and stochastic model with acceptable risk levels of 10% and 50% are shown in Table 6.

There are 11 groups which consist of the mathematical models’ results and the results of 10 replications. According to the results, the test statistics (sigma values) are greater than 0.05 in all CMS configurations. In conclusion, there is no significant difference between each pair of mathematical model and corresponding simulation replications for all models. The patterns of behaviors that of deterministic and stochastic models provided are shown in Fig. 3.

In addition to ANOVA tests, expected cell utilization output obtained from mathematical model and actual cell utilization output from simulation are also visually compared. In this comparison, 95% confidence intervals are established for each cell based on mathematical model results. An example is given in Table 7 to illustrate how the confidence interval is calculated for a cell.

According to the example in Table 7, products 1, 12, 23 and 29 are assigned to cell 1. The expected capacity requirements are given along with variance. The cell capacity is 2000 h annually. The expected capacity requirement of cell (product family) is the sum of expected utilizations of products, 1502 h. The variance is 72575.6 which is the sum of variances of products. Based on the 95% confidence level, z value is 1.64 and the lower and upper bounds are 1060.2 and 1943.8 h, respectively. Then, the average cell utilization obtained from simulation is also included in graphs (Figs. 3–5).

According to the results shown in Fig. 3, all of the utilisations obtained from simulation runs fall in between lower and upper bounds of confidence intervals. Therefore, it is concluded that the mathematical models are valid based on 95% level of confidence in terms of cell utilization. Additionally, in cells which have cell
utilizations closer to 100%, simulation results resulted in 100% utilization due to the impact of variance. The detailed results are provided in the following section.

7. Experimentation and results

In this section, the experiments performed and the results obtained are discussed. This section consists of 4 phases. In the first phase, the data used in the experimentation is explained. Then, warm-up analysis is described. In the third phase, experimentation is expressed. In the final phase, experimentation results are provided and explained in detail.

7.1. Data generation

The part-machine incidence matrix data is obtained from Suer et al.'s (2010) product-18 machine problem. Due to computational limits of Lingo and the computer memory, stochastic non-linear model was able to be experimented with first 30 products. There are three datasets generated and used in experimentation, namely: CMS design for 10, 20 and 30 products. Both annual demand and processing times are normally distributed. The mean of annual demand is generated from uniform distribution (1000, 2000) for each product. The standard deviation of demand is assumed to be 25% of the mean. Means of the processing times are generated via uniform distribution (15, 25). The standard deviations for the processing times are assumed to be 10% of the mean processing times.

7.2. Warm-up analysis

Prior to the simulation experiments, warm-up analysis is conducted. It is performed based on the following performance measures: average Work-in-Process (WIP) inventory, average waiting time and average number in queue. The configuration obtained from stochastic model with 10% risk is used in the warm-up period analysis. Because, neither the deterministic model nor the stochastic model with 50% risk provided a stable system which performs well in terms of corresponding performance measures. In other words, the selected performances had shown a continuous increasing trend due to the high risk taken. The pattern of behavior obtained based on the performance measures are shown in Fig. 4, respectively. The simulation length is 120,000 min (250 days x 8 h x 60 min). Based on the graphs, 200 h (12,000 min) of warm-up time is used since the steady-state behavior is reached.

7.3. Experimentation

Experimentation consists of two steps. Firstly, both deterministic and stochastic models are run with Lingo and the results are obtained. Secondly, the 30-product dataset results obtained from mathematical models are simulated with Arena. The datasets that consist of 10 and 20 products are parts of the dataset which includes 30 products. Mean capacity requirements are used in the experimentation with deterministic model. After three datasets are run with deterministic CMS design model, they are solved with non-linear stochastic CMS design model. Utilizations of cells are calculated according to the cell formations obtained from results. Finally, simulation models are built for 30-product case. The general framework of experimentation is shown in Fig. 5.

7.4. Results

The results are explained in two sections. In the first section, the results obtained from both deterministic and stochastic mathematical models are shown and compared with respect to the number of cells and expected cell utilization. In the second section, the performance analysis of designed CMSs is provided.

7.4.1. Results of deterministic and stochastic CMS design with 10% to 90% risk levels

In this section, the results of Suer et al.'s (2010) deterministic approach and the proposed stochastic approach are discussed. The results of dataset with 10 products are shown in Fig. 6. The deterministic model grouped the products into three cells. On the other hand, stochastic model grouped the products as four cells...
in lower risk levels 10%, 20% and 30% and decreased the number of cells to three for the 40% and greater risk levels.

The stochastic model with 40–70% risk resulted in the same number of cells (3) and expected cell utilizations (92%, 90% and 90%) with the deterministic model. On the other hand, 80% and 90% risk levels resulted in the same number of cells but not even load distribution. Moreover, expected cell utilization values exceeded available capacity in some of the cells under these risk levels. The results show that deterministic solution is at least 40% risky, which is a footprint that how much risk the deterministic model is taking by not considering the variance.

The stochastic model with 40–70% risk resulted in the same number of cells (3) and expected cell utilizations (92%, 90% and 90%) with the deterministic model. On the other hand, 80% and 90% risk levels resulted in the same number of cells but not even load distribution. Moreover, expected cell utilization values exceeded available capacity in some of the cells under these risk levels. The results show that deterministic solution is at least 40% risky, which is a footprint that how much risk the deterministic model is taking by not considering the variance.

The results of experiments with 20-product dataset are shown in Fig. 7. As shown in Fig. 7, similar to previous experimentation, deterministic model formed a CMS with less number of cells than the stochastic approach did with 10% risk level. Since deterministic model only considers the mean capacity requirements, it results in less number of cells and higher cell utilizations. While CMS with seven cells is formed for 10% risk level, the number of cells dropped to six for risk levels of 20%, 30% and 40% and to five for risk levels of 50%, 60%, 70%, 80% and finally to

<table>
<thead>
<tr>
<th>Cell no.</th>
<th>Part no.</th>
<th>Expected capacity requirements</th>
<th>Variance of capacity requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>522</td>
<td>18,769.00</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>482</td>
<td>3721.00</td>
</tr>
<tr>
<td>1</td>
<td>23</td>
<td>258</td>
<td>41,209.00</td>
</tr>
<tr>
<td>1</td>
<td>29</td>
<td>240</td>
<td>8873.64</td>
</tr>
<tr>
<td>1502</td>
<td></td>
<td>Variances of cell capacity requirements</td>
<td>72572.6</td>
</tr>
<tr>
<td>Lower bound</td>
<td>Upper bound</td>
<td>1943.8</td>
<td></td>
</tr>
</tbody>
</table>

Level of confidence = 95%

![Fig. 7. Validation results by cell utilization.](image)

**Table 7**

Example of confidence interval for cell utilization.

The stochastic model with 40–70% risk resulted in the same number of cells (3) and expected cell utilizations (92%, 90% and 90%) with the deterministic model. On the other hand, 80% and 90% risk levels resulted in the same number of cells but not even load distribution. Moreover, expected cell utilization values exceeded available capacity in some of the cells under these risk levels. The results show that deterministic solution is at least 40% risky, which is a footprint that how much risk the deterministic model is taking by not considering the variance.

The results of experiments with 20-product dataset are shown in Fig. 7. As shown in Fig. 7, similar to previous experimentation, deterministic model formed a CMS with less number of cells than the stochastic approach did with 10% risk level. Since deterministic model only considers the mean capacity requirements, it results in less number of cells and higher cell utilizations. While CMS with seven cells is formed for 10% risk level, the number of cells dropped to six for risk levels of 20%, 30% and 40% and to five for risk levels of 50%, 60%, 70%, 80% and finally to
four for risk level 90%. Risk levels 50% or more led to expected cell utilization values greater than 100%.

In 30-product problem category, deterministic approach grouped the products into eight cells as shown in Fig. 8. On the other hand, stochastic approach designed the system with nine cells for 10% and 20% risk levels. Both deterministic model and stochastic approach with 30%, 40% and 50% risk levels designed as 8-cell manufacturing system. Taking 60–80% risk resulted in formation of seven cells. Finally, the stochastic approach with 90% risk grouped the products into 6 manufacturing cells.

In summary, as the risk level increases, the number of cells decreases. In stochastic case, expected cell utilizations exceed cell
capacities as risk levels increase (above 50% risk). This in turn increases the probability of having shortages or force planners to schedule overtime to meet the customer demand. To observe the impact of risk on CMS performance, performance analysis of designed CMSs are also included in this study. This issue is discussed in the next section.

![Fig. 6. The results of the dataset with 10 products.](image)

![Fig. 7. The results of the dataset with 20 products.](image)

![Fig. 8. The results of the dataset with 30 products.](image)
7.4.2. CMS system performance analysis

In this section, stochastic and deterministic approaches are compared according to four performance measures, namely number of machines, total WIP, total waiting time and total number waiting according to the results obtained with 30-product dataset. The simulation results obtained for all CMS configurations are shown in Table 8. The results are consolidated as the number of cells formed, the total number of machines in the system, total WIP in the system, average waiting time of parts and average number of jobs waiting in the system. According to the results of deterministic case, there is a total of 71 machines allocated in 8 manufacturing cells (Table 8). According to the results, as the risk level decreases, there is significant improvement observed in all performance metrics with a slight increase in the number of machines.

The best configuration is obtained with the proposed approach with 10% risk level. The number of machines increases 15.4% comparing to Suer et al.'s (2010) deterministic approach, whereas there are significant improvements observed in all the performance metrics as 68% reduction in WIP, 96% decrease in average waiting time and 94% decrease in average number waiting.

The detailed results of the proposed stochastic model with 10% risk are given in Table 9. According to Table 9, a more stable CMS performance is observed in all performance measures. There are 11 more machines and 1 more cell allocated in stochastic CMS design with 10% risk in comparison to Suer et al.’s (2010) deterministic approach. Therefore, shorter waiting times, lower WIPs and lower total number of jobs in queues are obtained. The CMS is designed with nine cells and 82 machines. In this case, WIP is 46.5, waiting time is 71.6 h and 5.17 jobs waiting in queues obtained by the average of the results of 10 replications.

### Table 8
Performance analysis of CMS configurations.

<table>
<thead>
<tr>
<th>Model</th>
<th># Cells</th>
<th># Machines</th>
<th>Avg. WIP</th>
<th>Avg. waiting time</th>
<th>Avg. number waiting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suer et al. (2010)</td>
<td>8</td>
<td>71</td>
<td>144.64</td>
<td>201.20</td>
<td>87.42</td>
</tr>
<tr>
<td>The proposed Approach with 50% risk level</td>
<td>8</td>
<td>73</td>
<td>122.55</td>
<td>288.80</td>
<td>72.82</td>
</tr>
<tr>
<td>The proposed Approach with 40% risk</td>
<td>8</td>
<td>82</td>
<td>55.13</td>
<td>27.22</td>
<td>13.63</td>
</tr>
<tr>
<td>The proposed Approach with 30% risk</td>
<td>8</td>
<td>82</td>
<td>46.74</td>
<td>10.64</td>
<td>6.34</td>
</tr>
<tr>
<td>The proposed Approach with 20% risk</td>
<td>9</td>
<td>82</td>
<td>46.37</td>
<td>9.20</td>
<td>5.03</td>
</tr>
<tr>
<td>The proposed approach with 10% risk</td>
<td>9</td>
<td>82</td>
<td>46.50</td>
<td>7.96</td>
<td>5.17</td>
</tr>
</tbody>
</table>

### Table 9
Performance analysis of stochastic CMS with 10% risk level.

<table>
<thead>
<tr>
<th>Cell ID</th>
<th># Machines</th>
<th>Avg. WIP</th>
<th>Avg. waiting time</th>
<th>Avg. number waiting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>4.44</td>
<td>7.67</td>
<td>0.43</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>5.20</td>
<td>8.13</td>
<td>0.49</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>6.24</td>
<td>8.05</td>
<td>0.56</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>3.62</td>
<td>4.09</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4.31</td>
<td>6.64</td>
<td>0.33</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>5.82</td>
<td>9.21</td>
<td>0.70</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>6.09</td>
<td>14.35</td>
<td>1.17</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>7.54</td>
<td>9.57</td>
<td>0.88</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>3.24</td>
<td>3.86</td>
<td>0.26</td>
</tr>
<tr>
<td>Average</td>
<td>-</td>
<td>-</td>
<td>7.96</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
<td>46.50</td>
<td>-</td>
<td>5.17</td>
</tr>
</tbody>
</table>

### Table 10
Comparison of capacitated cell formation approaches.

<table>
<thead>
<tr>
<th>Cap Form Method</th>
<th>Suer et al. (2010)</th>
<th>The proposed stochastic approach with 10% risk</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Cells</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td># of Machines</td>
<td>73</td>
<td>82</td>
</tr>
<tr>
<td># of families</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>WIP</td>
<td>(1.23, 0.29)</td>
<td>(1.23, 0.29)</td>
</tr>
<tr>
<td>Waiting Time</td>
<td>(1.24, 0.27)</td>
<td>(1.24, 0.27)</td>
</tr>
<tr>
<td>Number Waiting</td>
<td>(1.24, 0.27)</td>
<td>(1.24, 0.27)</td>
</tr>
</tbody>
</table>
7.4.3. Analysis of cell formation and comparison with previous approaches

Several cell formation approaches have been proposed in literature. However, since majority of them does not consider capacity while grouping products into cells, ones that are in the same classification with the proposed approach are compared. Capacitated single, average and complete linkage cell formation methods (CFMs) proposed by Vega (1999) are used for comparison purposes. The results of cell formation methods are provided in Table 10.

According to the results, the minimum number of machines is obtained with deterministic capacitated P-Median model (Suer et al., 2010). The number of machines ranges between 71 and 82 and the number of cells varies from 8 to 11. Since machine costs are similar for the problem studied, it is sufficient to check the number of machines required as a result of cell formation. The proposed stochastic capacitated P-median model produced comparable results with respect to the number of machines compared to other capacitated clustering procedures, when a risk level of 50% is considered. However, there is a tradeoff between the number of machines thus the number of cells and risk of exceeding available capacity. Deterministic capacitated cell formation methods do not provide any information about this tradeoff. However, the proposed approach identifies the risk of exceeding the available capacity while maximizing the total similarity among products. Therefore, decision maker can vary risk levels and generate alternative designs prior to implementation.

8. Conclusion and future work

In this study, stochastic capacitated cellular manufacturing system has been addressed. Suer et al.’s (2010) deterministic capacitated P-Median model is modified and a non-linear mathematical model is developed to deal with uncertain demand and processing times. Demand and processing times are normally distributed and the statistical distributions for capacity requirements are determined via statistical analyses and non-parametric tests. Independent cells are assumed and inter-cell movement is not allowed. Three datasets are used for experimenting with deterministic and proposed stochastic mathematical models, which consist of 10, 20 and 30 products, respectively.

Cell and family formations obtained from mathematical models are used to build simulation models. The results of mathematical models are validated by comparing with simulation results via ANOVA. According to the ANOVA results, there is no significant difference observed between the results of mathematical models and simulations. Both CMSs designed with deterministic and stochastic mathematical models are compared in terms of CM performance measures suggested by Wemmerlov and Johnson (Wemmerlov & Johnson, 1997).

According to the results of mathematical models, deterministic model resulted in less number of cells in comparison to the stochastic model with lower levels of risks. As the risk level increased, the proposed approach formed CMS with lower number of cells and machines in all datasets and as a result, cell utilizations are increased.

To analyze the performance of designed CMSs by deterministic and proposed stochastic approaches, simulation experiments are performed. A total of six CMS configurations are considered, namely: Suer et al.’s (2010) deterministic approach and the proposed stochastic approach with 10%, 20%, 30%, 40% and 50% risk levels. In this paper, Suer et al.’s (2010) deterministic capacitated P-median model is used as benchmark work. Deterministic model formed the CMS with less number of cells and machines which resulted in higher cell utilizations. However, according to simulation results, the highly utilized cells resulted in significant amount of total WIP, average waiting time and average number of jobs waiting. Results indicate that the proposed stochastic model with 10% risk level performed as the best in all performance measures except the number of machines. If the number of machines is a significant concern to the designer, then 30% risk level is also a viable option as it increases the number of machines by 7% while obtaining significant improvements in all other performance metrics. As the risk level increases, cell utilizations increase which result in longer queue sizes and waiting times and more WIP.

This problem may be extended to include other features of CMS design such as allowance of inter-cell movement, system implementation costs, setup times, worker allocation, different similarity matrix definitions. Another possible direction is to study the bottleneck machine definition in detail and assess its impact on the results. Finally, genetic algorithms or other meta-heuristics can be also considered to solve larger problem sizes.

References


