International Journal of Production Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tprs20

Joint cell loading and scheduling approach to cellular manufacturing systems

M. Selim Akturk [a]
[a] Department of Industrial Engineering , Bilkent University , Bilkent, Ankara, 06800, Turkey
Published online: 31 Jan 2011.

International Journal of Production Research
Vol. 49, No. 21, 1 November 2011, 6321–6341

RESEARCH ARTICLE

Joint cell loading and scheduling approach to cellular manufacturing systems

M. Selim Akturk*

Department of Industrial Engineering, Bilkent University, Bilkent, Ankara, 06800, Turkey

(*Received 12 March 2010; final version received 8 October 2010*)

A hierarchical multi-objective heuristic algorithm and pricing mechanism are developed to first determine the cell loading decisions, and then lot sizes for each item and to obtain a sequence of items comprising the group technology families to be processed at each manufacturing cell that minimise the setup, inventory holding, overtime and tardiness costs simultaneously. The linkage between the different levels is achieved using the proposed pricing mechanism through a set of dual variables associated with the resource and inventory balance constraints, and the feasibility status feedback information is passed between the levels to ensure internally consistent decisions. The computational results indicate that the proposed algorithm is very efficient in finding a compromise solution for a set of randomly generated problems compared with a set of competing algorithms.

Keywords: cell scheduling; production planning; cellular manufacture; linear programming

1. Introduction

Group technology (GT) is an innovative approach for batch-type production that seeks to rationalise small-lot production by capitalising on the similarities that exist among component parts and/or processes. GT tries to bring the benefits of mass production, such as reduced material handling and manufacturing lead time, and simplified planning and scheduling activities, to high variety, medium-to-low volume quantity production. The application of GT to manufacturing is called cellular manufacturing (CM), which is the physical division of the manufacturing facilities into production cells, representing the basis for advanced manufacturing systems such as just-in-time and flexible manufacturing. On the other hand, it has been reported in different studies (e.g., Wemmerlov and Johnson (1997) and Suresh and Gaalman (2000)) that a direct conversion from a functional layout to a cellular layout by itself may not bring about all the stated advantages suggested in the literature. It would appear that a new cellularly divided shop must be controlled with efficient production planning systems so as to fully benefit from the advantages of GT. Therefore, the thrust of this paper is the development of a hierarchical multi-objective approach to solve the cell loading, lot sizing and cell scheduling problems in CM systems.

*Email: akturk@bilkent.edu.tr

ISSN 0020–7543 print/ISSN 1366–588X online

DOI: 10.1080/00207543.2010.532165
http://www.informaworld.com

In most production systems, cell loading, lot-sizing and scheduling decisions are made at different levels of the hierarchy, however there is a strong interaction among these decisions. Furthermore, most of the existing models do not utilise the shop floor conditions in lot-sizing and scheduling decisions, even though such a decision might improve the system performance.

There is a vast amount of literature concerning lot-sizing on the one hand and scheduling on the other, but there are a few studies that contain elements of both lot-sizing and scheduling. Most of the research on lot-sizing has concentrated on a trade-off between setup and inventory holding costs and has ignored analysis of the interface between lot-sizing and scheduling decisions. Biggs (1979) shows that there is a strong relation between these decisions, and under changing shop floor conditions, different lot-sizing and scheduling rule combinations perform better than others. The key issue in many scheduling situations is that sequencing methods are not only a means of controlling time performance; in fact, cell loading and lot-sizing have a major impact on lead times. Using fixed lot sizes and subsequent scheduling efforts make the system blind to these interactions that might be exploited to increase the overall system effectiveness. The following justification seems to be prevalent for not evaluating the cell loading, lot-sizing and scheduling problems jointly. The cell loading subproblem is considered as a planning decision and is assumed to be solved at a higher level in an organisation than the scheduling subproblem, which is usually solved after lot-sizing decisions. As a result, the scheduling subproblem searches for a solution in a limited feasible solution space. In addition, the reduced shop flexibility of CM systems combined with a three-step approach severely restricts the number of alternatives possible at the scheduling level. Since cell loading, lot-sizing and scheduling decisions are interrelated, simultaneous solutions can improve the system performance. The key idea is that cell schedules and lot sizes should change as the cell loading decisions change over time, as opposed to having fixed lot sizes as is widely used in the literature.

The complexity of a joint approach has already been shown by Potts and Wassenhove (1992). Therefore, the spatial decomposition of a CM shop configuration is used to simplify the problem where each GT cell is designed to produce a set of part families. We assume that switching from one GT family to another requires a major change in setup of the workstation, while items within a family can be accommodated with minor adjustments that require only a relatively small amount of time that can be added to the processing time. This assumption reflects the fact that each family includes a set of items with different demand and due-date requirements but similar production and setup costs since the similarities of the items are much closer with respect to size, shape and processing routes within a family than across the families.

There has been increasing interest in scheduling problems with setup times and/or costs. Potts and Kovalyov (2000) surveyed scheduling problems with batching, while Allahverdi *et al.* (1999) reviewed the scheduling literature involving setup considerations. For a more recent review, we refer to Allahverdi *et al.* (2008). Crauwels *et al.* (1998) and Azizoglu and Webster (2003) developed branch-and-bound-based approaches for the weighted flow time problem with family setup times on single and parallel machines, respectively. Baker (1999) considered the case where the setup times are the same for different job families. He proposed and compared several heuristic procedures for the problem. With respect to the cell scheduling problem, Bokhorst *et al.* (2008) investigated the impact of family-based dispatching rules on the mean flow time using a simulation study in small manufacturing cells with and without labour constraints.

Venkataramanaiah (2008) developed a simulated annealing (SA)-based algorithm for scheduling in a flowline-based manufacturing cell with missing operations with the objective of minimising the weighted sum of the makespan, flowtime and idle time, and compared the proposed approach with dispatching rules. In a recent study, Ying *et al.* (2010) presented a computational investigation concerning the performance evaluation of non-permutation versus permutation schedules for the flowline manufacturing cell with sequence-dependent setup times utilising an SA algorithm. For a more detailed discussion on scheduling job families, we refer to Baker and Trietsch (2009). There is a limited number of studies on the cell loading problem. Suer *et al.* (1999) proposed new rules for the cell loading problem, and studied the impact of cell loading and scheduling considering the entire system and reported that no single rule performed better in terms of multiple performance measures. Suer *et al.* (2008) studied a fuzzy bi-objective cell loading problem in labour-intensive cellular environments. Suer *et al.* (2009) proposed heuristic procedures for the cell loading problem in a shoe manufacturing company.

In addition to a spatial decomposition, we also employ a structured product-based aggregation/disaggregation (A/D) scheme based on GT-oriented classification and coding (CC) systems. An A/D scheme is applied to reduce the size of the problem, where the decomposition of the manufacturing system proceeds in three dimensions: by floor space (or resource-based), by product, and by time horizon. In the floor space decomposition, the manufacturing system is divided into a set of GT cells where each cell is designed to produce a GT family or families. In the product-based decomposition, similar items are grouped into GT families. Throughout this research a GT family is defined as a set of items that require similar machinery, tooling, machine operations, jigs and fixtures. Both GT cell formation and the prerequisite product family determinations are assumed to have been done *a priori* to this planning activity, but their impact on the performance of the results are tested in Section 5. In the time scale decomposition, the levels of the decision hierarchy differ by complexity, scope and time horizon in that higher levels deal with longer range and more aggregated issues, and lower levels deal with short-term and more specific issues. The main contributions of this paper are to develop a hierarchical multi-objective approach and to propose new solution techniques to solve cell loading, lot-sizing and scheduling problems as discussed in Sections 2 and 4. The linkage between the different levels is achieved using the proposed pricing mechanism (discussed in Section 3), and the feasibility status feedback information is passed between the levels to ensure internally consistent decisions.

In summary, there are four concepts that could impact the performance of a cellular manufacturing system: (i) cell formation, (ii) cell loading, (iii) lot-sizing, and (iv) scheduling decisions. As stated above, cell formation is not the topic of the paper and it is assumed that cells and product families are already formed. Issues (ii), (iii) and (iv) are jointly tackled in this paper in a two-level approach. The cell loading problem is solved at the first level, whereas the joint lot-sizing (iii) and scheduling (iv) problem is addressed at the second level. The overall decision-making hierarchy of the proposed hierarchical multi-objective approach is outlined in Figure 1. Section 2 describes concept (ii) and its mathematical formulation. In this stage, initial family lot size estimates per cell and per period are required. Section 4 addresses joint lot-sizing (iii) and scheduling (iv). In order to solve the joint problem for each cell in each period at the second level for a given cell loading decision, we also propose a multi-objective joint algorithm (MJA) as discussed in Section 4.

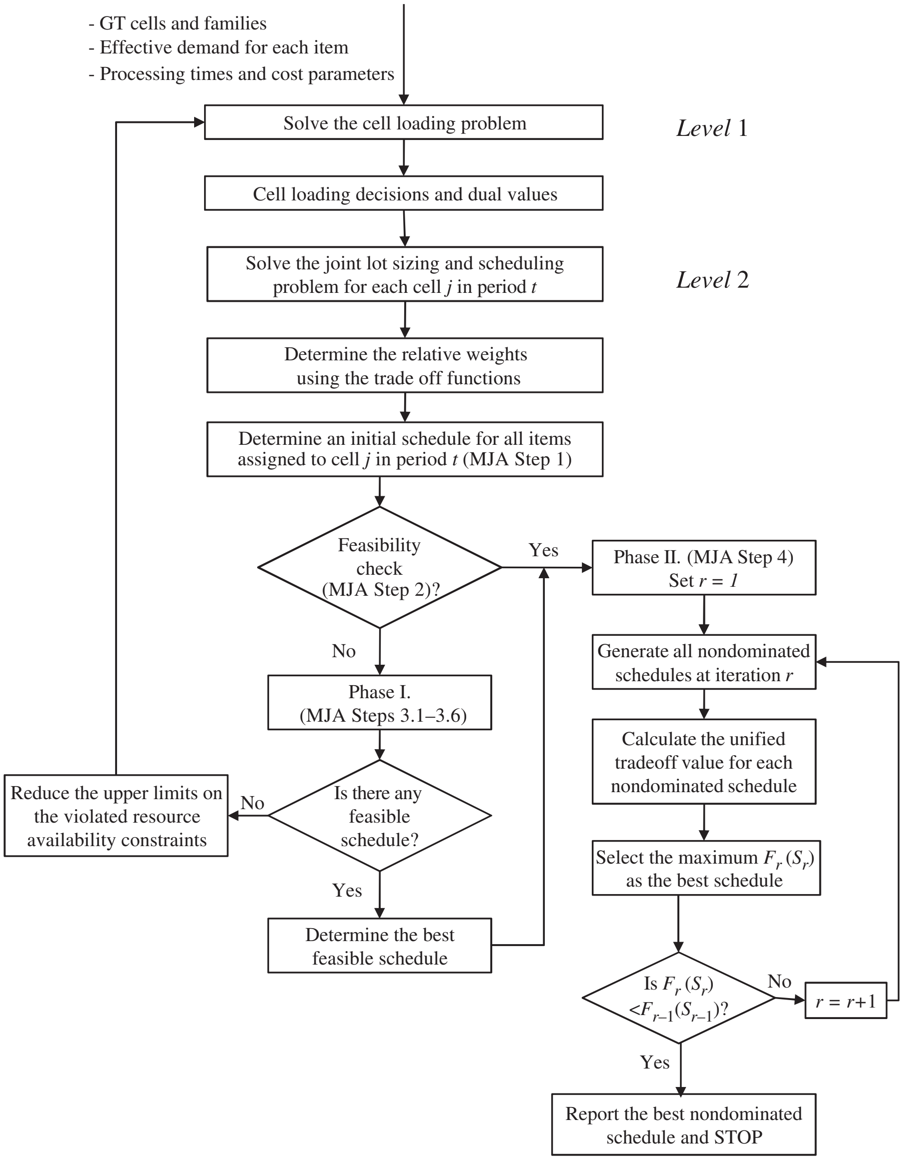

Figure 1. Flow-chart representation of the algorithm's framework.

2. Cell loading

Cell loading, or production planning, in a CM environment is a decision activity that determines the kind of items and the quantities to be produced in each cell in the specified time period, subject to the production capacity and demand forecast. Our aim is to allocate production capacity among GT families and items by means of the proposed aggregate planning model. This can be achieved by solving a multi-period optimisation problem that minimises the summation of production, setup, inventory holding, and regular and overtime capacity costs subject to production and inventory balance

Table 1. Cell loading parameters.

C_{ijt}	Average unit cost for producing one unit of family i by cell j in period t
r_{jt}	Average cost of one regular time unit for cell j during period t
o_{jt}	Average cost of one overtime unit for cell j during period t
h_{it}	Average holding cost for family i in period t
df_{it}	Demand for family i in period t
a_{ij}	Average total time required to produce one unit of family i in cell j
BS_{ij}	Setup cost for family i in secondary cell j
BP_{ij}	Setup cost for family i in primary cell j
bs_{ij}	Setup time for family i in secondary cell j
bp_{ij}	Setup time for family i in primary cell j
Q_{ijt}	Initial estimate for the lot size of family i in cell j in period t
P_j	Set of families whose primary cell is j
S_j	Set of families whose secondary cell is j
$FS(j)$	A feasible set of families assignable to cell j
d_{kt}	Demand for item k in period t
$TI(i)$	Set of items belonging to family i
PR_{kl}	Average total time required to produce one unit of item k using resource l
$LR(j)$	Set of resources belonging to cell j
$NO(kj)$	Number of operations for item k in cell j
$NI(i)$	Number of items belonging to family i
$NR(j)$	Number of resources in cell j
J	Number of cells
N	Number of families
K	Number of items
L	Number of resources
T	Planning horizon

Table 2. Cell loading decision variables.

X_{ijt}	Number of units of family i produced by cell j in period t
IF_{it}	Inventory of family i at the end of period t
O_{jt}	Overtime used by cell j in period t
R_{jt}	Regular time used by cell j in period t
Z_{kjt}	Number of units of item k produced by cell j in period t
I_{kt}	Inventory of item k at the end of period t
OR_{lt}	Overtime used by resource l in period t
RR_{lt}	Regular time used by resource l in period t

constraints for families and items, and capacity feasibility constraints for GT cells and resources over the planning horizon. The objective function corresponds to the minimisation of the variable cost of production. The set of parameters and decision variables are given in Tables 1 and 2, respectively.

A mathematical formulation of the cell loading problem is as follows:

$$\min \sum_{t=1}^{T} \sum_{j=1}^{J} \left[\sum_{i \in FS(j)} C_{ijt} \cdot X_{ijt} + \sum_{i \in S_j} (BS_{ij}/Q_{ijt}) \cdot X_{ijt} + \sum_{i \in P_j} (BP_{ij}/Q_{ijt}) \cdot X_{ijt} + o_{jt} \cdot O_{jt} + r_{jt} \cdot R_{jt} \right] + \sum_{t=1}^{T} \sum_{i=1}^{N} h_{it} \cdot IF_{it}, \tag{1}$$

subject to

$$\sum_{j=1}^{J} X_{ijt} + IF_{i,t-1} - IF_{it} = df_{it}, \quad \text{for } i = 1, \ldots, N,\ t = 1, \ldots, T, \tag{2}$$

$$\sum_{i \in FS(j)} a_{ij}.X_{ijt} + \sum_{i \in S_j} (bs_{ij}/Q_{ijt}) \cdot X_{ijt} + \sum_{i \in P_j} (bp_{ij}/Q_{ijt}) \cdot X_{ijt} - O_{jt} = R_{jt},$$
$$\text{for } j = 1, \ldots, J,\ t = 1, \ldots, T, \tag{3}$$

$$0 \le O_{jt} \le (\text{upper limit}), \quad \forall j, t, \tag{4}$$

$$0 \le R_{jt} \le (\text{upper limit}), \quad \forall j, t, \tag{5}$$

$$\sum_{j=1}^{J} Z_{kjt} + I_{k,t-1} - I_{kt} = d_{kt}, \quad \forall k \in TI(i), t, \tag{6}$$

$$\sum_{k \in TI(i)} I_{kt} - IF_{it} = 0, \quad \forall i, t, \tag{7}$$

$$\sum_{i \in FS(j)} \sum_{k \in TI(i)} (PR_{kl} \cdot Z_{kjt}) - OR_{lt} = RR_{lt}, \quad \forall l \in LR(j), j, t, \tag{8}$$

$$0 \le OR_{lt} \le (\text{upper limit}), \quad \forall l, t, \tag{9}$$

$$0 \le RR_{lt} \le (\text{upper limit}), \quad \forall l, t, \tag{10}$$

$$\sum_{l \in LR(j)} OR_{lt} - O_{jt} = 0, \quad \forall j, t, \tag{11}$$

$$\sum_{l \in LR(j)} RR_{lt} - R_{jt} = 0, \quad \forall j, t, \tag{12}$$

$$X_{ijt}, IF_{it}, O_{jt}, R_{jt}, Z_{kjt}, I_{kt}, OR_{lt}, RR_{lt} \ge 0, \quad \forall i, j, k, l, t. \tag{13}$$

The constraint sets (2) and (6) are the inventory balance constraints for families and items, respectively, in which both the amount of inventory left in stock at the end of each period and the demand in each period are supplied by the amount of production in each period and the amount of inventory carried over from the previous period. No backordering is allowed. Moreover, a deterministic, but time-varying, demand for every item in every time period is assumed. Constraint (7), which represents the inventory consistency equations, links the item inventories to the inventory of the associated family. This constraint requires that the inventory for a family is equal to the sum of the inventories of the items contained in the family. As a result, individual items are mapped into their corresponding families. Given that $\sum_{k \in TI(i)} d_{kt} = d_{it}$, it can be shown that the constraint set, which includes all the resource, production and inventory constraints, implies that $\sum_{k \in TI(i)} Z_{kjt} = X_{ijt}$ for all i, j and t; that is, for each time period, the total family production equals the sum of the production quantities for its items. Constraints (3) and (8) are the capacity feasibility constraints for GT cells and resources. Upper limits on regular and overtime usages are also defined by constraints (4), (5), (9) and (10). Constraints (11) and (12) link the available time for each GT cell to the resources comprising that cell.

Implicit in constraint (3) is the possibility that each GT family can have more than one feasible cell for its production. A feasible cell is defined as a cell in which a family can be processed entirely within that cell considering feasibility requirements. A more detailed discussion on the formation of primary and secondary cells can be found in Akturk and Yayla (2005). It is assumed that the primary cell of a family is capable of producing the family at the lowest possible cost. Secondary cells are those in which the manufacture of the family is possible at a higher cost, due to both increased setup and material handling costs, assuming that all cells are initially tooled for their primary families. An additional cost is incurred when families other than the primary families need to be produced at that cell. Therefore, the setup costs and setup times for the families in their secondary cells are assumed to be greater than the setup costs and setup times in the primary cells. The parameter Q_{ijt} is an initial estimate for the lot size allowing the cell resource constraints to approximately account for the total setup time, which is directly proportional to the number of setups required to meet the desired production quantities at each cell. The definition of primary and secondary cells for each family also allows the production management system to react to the variations in the families' total demand. For example, during a very low demand period, one cell may be completely shut down because of maintenance and that cell's families are assigned to some other cell.

In this study we present a linear programming (LP) formulation of the cell loading problem that can be solved by any commercial LP solver, such that an optimal solution is found by using the IBM ILOG CPLEX optimisation package. Linear programming is a convenient type of model to use at this level because of the wide availability of LP codes. LP also permits sensitivity and parametric analysis to be performed quite easily and information on dual values can be derived at little additional computational cost. Both the cell loading decisions as well as the dual values will be passed to the second level in order to solve the joint lot sizing and scheduling problem for each cell j in period t as depicted in Figure 1.

3. Trade-off functions: pricing mechanism

Since we are trying to solve cell loading, lot-sizing and scheduling problems simultaneously, we have multiple and usually conflicting criteria representing the various objectives of these particular problems at different levels. In practice, the weights of cost- and due-date-related objectives for the decision maker may vary over time. For example, if the workload is heavy, scheduling related objectives become more important. One of the most important steps in multi-objective optimisation is to estimate the weights associated with each criterion which reflect the relative importance of each objective. If indifference curves could be developed to relate various objective functions and transform different units of measure into a common unit, then they would serve as a true arbiter when comparing several non-dominated solutions. One possible approach to unify multiple local objective functions into a single global objective function would be to use a trade-off function to define the relationship among several criteria. Due to the size of these problems, and since the loading and scheduling decisions are being made on a regular, perhaps frequent, basis, interactive multiple criteria decision making (MCDM) approaches may not be practical. This creates the need for a MCDM with an *a priori* objective weighting scheme.

The following function is used to serve as a trade-off function for the setup cost criterion:

$$y_1(S, t_1) = -c_1(1 - e^{a_1 t_1}), \tag{14}$$

where t_1 is the total setup time in the cell during the planning period, y_1 (S, t_1) is the 'relative' dollar value of the aggregated total setup time on a particular cell, c_1 and a_1 are constants, $c_1 > 0$ and $a_1 > 0$.

An exponential trade-off function indicates that the setup time is considered as a surrogate for the cost of violation of capacity constraints, and also incorporates the concept of diminishing returns on the value of the setup time adjustments. If the capacity utilisation on a resource is near or above its upper limit, then that resource is limiting the throughput of the system, and the capacity takes on some economic value, e.g. a dual value. This might be, for example, the opportunity cost of not producing additional units of output per unit time, relative to using another more expensive alternative (e.g., overtime) or relative to lost revenue due to not meeting due dates. Furthermore, the value of a dual variable is a monotonically non-increasing function of the right-hand side value of a capacity constraint. When the right-hand side becomes less restrictive, then the dual value moves closer to zero. When the capacity constraint is not active, then the dual variable is equal to zero. The relationship between a dual variable and the decreasing right-hand side of a capacity constraint resembles an exponential trade-off function for the setup cost criterion. Therefore, the parameters of the proposed trade-off function, c_1 and a_1, can be derived using a set of dual variables from the cell loading problem formulation, which was solved at the first level of the proposed hierarchy.

The instantaneous rate of change of $y_1(S, t_1)$ with respect to t_1 can be written as $\partial y_1(S, t_1)/\partial t_1 = c_1 a_1 e^{a_1 t_1}$, where $\partial y_1(S,\ t_1)/\partial t_1$ measures the rate of cost increase at the available capacity (consequently, the rate of cost increase of the overall objective function) due to increasing the total setup time. Furthermore, (∂y_1) is an approximation of the cost increase per unit total setup time increase. In the mathematical formulation of the cell loading problem, the capacity restrictions for each cell and for each resource are explicitly considered. If the dual variable associated with the cell constraint (j, t) is equal to β^1_{jt}, the objective function value in the cell loading level will be decreased by β^1_{jt}, where $\beta^1_{jt} \geq 0$, if the available capacity of the resource is increased by one time unit (assuming no basis change results). Therefore, the marginal value of the total setup time, when the right-hand side value of the constraint (j, t) is equal to b_{jt}, may be written as

$$\beta^1_{jt} = c_1 a_1 e^{a_1 b_{jt}}. \tag{15}$$

In addition, a sensitivity analysis gives the amount by which the right-hand side can be changed before the current basis becomes infeasible (all other LP parameters remain constant). An allowable decrease for each right-hand side, say γ_{jt}, gives the maximum amount by which the right-hand side of a constraint can be decreased with the current basis remaining optimal (all other LP parameters remain constant). Therefore, if the dual variable associated with the cell constraint (j, t) is equal to β^2_{jt} when the current right-hand side value, b_{jt}, is decreased by $(\gamma_{jt}+\sigma)$, then the following relationship gives another point on the trade-off function:

$$\beta_{jt}^2 = c_1 a_1 e^{a_1(b_{jt}+(\gamma_{jt}+\sigma))}. \tag{16}$$

Consequently, the parameters c_1 and a_1 of the trade-off function can be determined by equating the partial derivatives of the trade-off functions to the dual variables at the cell loading level. Since this is a minimisation problem, both β_{jt}^1 and β_{jt}^2 are non-negative and $\beta_{jt}^2 > \beta_{jt}^1 \geq 0$. In addition, $(b_{jt}+(\gamma_{jt}+\sigma)) > b_{jt} > 0$. Therefore,

$$\frac{\beta_{jt}^2}{\beta_{jt}^1} = \frac{c_1 a_1 e^{a_1(b_{jt}+(\gamma_{jt}+\sigma))}}{c_1 a_1 e^{a_1 b_{jt}}} \Rightarrow \ln\left(\frac{\beta_{jt}^2}{\beta_{jt}^1}\right) + a_1 b_{jt} = a_1(b_{jt} + (\gamma_{jt}+\sigma)). \tag{17}$$

Consequently,

$$a_1 = \frac{\ln(\beta_{jt}^2/\beta_{jt}^1)}{(\gamma_{jt}+\sigma)}, \quad c_1 = \frac{\beta_{jt}^1}{a_1 e^{a_1 b_{jt}}}, \tag{18}$$

indicating that $a_1 > 0$ and $c_1 > 0$.

By applying similar logic for the tardiness criterion, a set of trade-off functions for each family allocated to a particular cell can be defined as follows:

$$y_2(S, t_2) = -c_2(1 - e^{a_2 t_2}), \tag{19}$$

where t_2 is the total tardiness for all items in each family during the planning period, $y_2(S, t_2)$ is the 'relative' dollar value of the total tardiness, t_2, c_2 and a_2 are constants, $c_2 > 0$ and $a_2 > 0$.

This exponential trade-off function indicates that, if total tardiness is high, we are willing to spend more money to reduce tardiness by one time unit than if it is low. The rate of change of $y_2(S,\ t_2)$ with respect to t_2 can be written as $\partial y_2(S, t_2)/\partial t_2 = c_2 a_2 e^{a_2 t_2}$, where $\partial y_2(S,\ t_2)/\partial t_2$ has an approximate economic interpretation as the cost increase in dollars (∂y_2) incurred by increasing the total tardiness by (∂t_2). In the constraint set of the cell loading problem, there are production–inventory balance equations for each family and for each item. The dual variable associated with the balance equation (i, t) is defined by λ_{it}^1. The objective function value in the cell loading level will be decreased by λ_{it}^1 if the effective demand, df_{it}, for family i in period t is reduced by one unit. If the total tardiness for family i in a particular cell, given by the initial schedule, is equal to T_i, and $T_i \geq 0$, then we can define the following relationship:

$$\lambda_{it}^1 = c_2 a_2 e^{a_2 T_i}. \tag{20}$$

A sensitivity analysis on an allowable decrease, χ_{it}, of df_{it} provides a new dual value λ_{it}^2, if χ_{it} is made slightly larger. If the weighted average across all items within the family for the total time required to produce one unit of family i is equal to r_i^1, then the following approximation gives another point on the trade-off function:

$$\lambda_{it}^2 = c_2 a_2 e^{a_2(T_i - (\chi_{it}+\delta) r_i^1)}. \tag{21}$$

Consequently, the parameters c_2 and a_2 of the trade-off function can be determined by equating the partial derivatives of the trade-off functions to the dual variables already found at the cell loading level, λ_{it}^1, and alternative ones, λ_{it}^2, that are easily derived similar to the development of the trade-off function for the setup cost criterion, using the weighted average of producing one unit of family i. Furthermore, $\lambda_{it}^{1} > \lambda_{it}^{2} > 0$ and $T_i > (T_i - (\chi_{it} + \delta) r_i^1)$. Therefore,

$$a_2 = \frac{\ln(\lambda_{it}^{1}/\lambda_{it}^{2})}{((\chi_{it} + \delta) r_i^1)}, \quad c_2 = \frac{\lambda_{it}^{1}}{a_2 e^{a_2 r_i^1}}, \tag{22}$$

thus $a_2 > 0$ and $c_2 > 0$.

In summary, the parameters of the trade-off functions, a_1, c_1, a_2, and c_2, are derived from the pricing mechanism which is based on the duality information from the cell loading problem solution. This information is based on optimised economic decisions and not 'expert judgements'. This allows us to consider multiple criteria at the next level, and a schedule evaluation will be based on a combination of both cost and performance criteria. Furthermore, we could dynamically adjust the lot sizes and cell schedules as the cell loading decisions change over time as opposed to having fixed lot sizes as is widely used in other models in the literature.

4. Joint lot sizing and scheduling

The second level of the decision hierarchy addresses the economic lot scheduling problem (ELSP) and the relationship between lot sizes, inventory levels and meeting due dates. As already mentioned in the literature review, most of the research on ELSP has concentrated on a trade-off between setup and inventory holding costs and has failed to consider adequately the interface between ELSP and sequencing decisions in either the GT context or when there are non-commensurable objectives. The key issue is that, in many scheduling situations, sequencing methods are not the only means of controlling time performance; in fact, cell loading and lot sizing have a major impact on performance measures such as tardiness and flow time. This leads to an optimisation problem that balances the need to minimise setup and inventory holding costs with the need to minimise tardiness due to not meeting due dates. As a result of this observation, the proposed approach to the joint ELSP and scheduling problem, assuming GT family groups, minimised the lot size and due-date-related costs for each cell j in each time period t for a given set of resources and due-date requirements. Although items within each family are similar with respect to design and manufacturing attributes, they might have different demand and due-date requirements. Furthermore, the total sum of item requirements in a given period, Z_{kjt}, was one of the decision variables in the cell loading level. It is important to note that Z_{kjt} is actually a collection of different customer orders, each with a distinct due date. Consequently, scheduling all of the items within the same family consecutively may not be the best scheduling rule when we consider the due-date-related scheduling criteria. Therefore, we propose a new multi-objective scheduling algorithm at the second level to determine the lot sizes for each item and to obtain a sequence of items comprising the group technology families to be processed at each manufacturing cell as outlined in algorithm 1 (denoted MJA).

Algorithm 1 : Multi-objective Joint Algorithm (MJA)

Step 1: *Initial schedule generation*: For all items (e.g., L_{ikf}, for customer order i of item k in family f) that are assigned to cell j in period t, find an initial schedule, S_0, to minimise the average tardiness.

Step 2: *Feasibility check*: For a given schedule S_0, find the MS_{hf} values, where MS_{hf} is defined as a 0–1 binary variable that takes a value of 1 if family h immediately precedes family f, and zero otherwise. Let the initial total setup time, $TS_0 = \sum_h \sum_f S_{hf} \cdot MS_{hf}$, where S_{hf} is the required setup time switching from family h to family f, i.e. $S_{ff} = 0$. For each resource l in cell j ($\forall l \in LR(j)$) in period t, check for feasibility using the following resource availability constraint: $\sum_{i \in FS(j)} \sum_{k \in TI(i)} (PR_{kl} \cdot Z_{kjt}) + \sum_h \sum_f S_{hf} \cdot MS_{hf} \leq RR_{lt} + OR_{lt}$. If the given schedule is feasible, then it will always be feasible since the proposed algorithm only reduces the total setup time at each iteration. Therefore, we do not have to check for feasibility again; go to Step 4, Phase II. If schedule S is found to be infeasible, let the maximum amount of infeasibility on cell j in period t be $\Upsilon = \max_{l \in LR(j)} [\sum_{i \in FS(j)} \sum_{k \in TI(i)} (PR_{kl} \cdot Z_{kjt}) + TS_0] - (RR_{lt} + OR_{lt})$. Then, go to Step 3, Phase I.

Step 3: *Phase I*: Since our aim is to achieve feasibility in Phase I, at each iteration r we will only consider the forward or backward shifts that could make the corresponding MS_{hf} value equal to zero. Set $r = 0$ and $\delta_0 = 0$.

Step 3.1. $TS_r = TS_{r-1} -$ (decrease in total setup time, β_r). We will only consider the alternative schedules for which $\beta_r > 0$, and denote this set as $\mathcal{S}$. If $\mathcal{S}$ is an empty set, then no feasible solution exists; thus, stop and go back to the cell loading level and reduce the upper limits on the resource availability constraints (i.e. constraint sets (4) and (9)) for the bottleneck resource (or resources).

Step 3.2. For each alternative schedule $S_r \in \mathcal{S}$, calculate (total tardiness)$_r$ = (total tardiness)$_{r-1} + \Delta_r$. The value of Δ_r will be found by considering both the forward and backward shifts.

Step 3.3. *Forward shift*: If an order, L_{nkf}, is shifted forward, then let set $A: \{m \mid \text{in } S_{r-1}, m < n; \text{ in } S_r, n < m\}$, and $B: \{m \mid n < m \text{ in both } S_{r-1} \text{ and } S_r\}$, and $m < n$ means that order m precedes order n (not necessarily directly) in the given sequence. $\Delta_r^f = \sum_{m \in A} \max\{(C_{mkf}(r-1) + t_{nkf} - dd_{mkf}), 0\} - \sum_{m \in B} \max\{(C_{mkf}(r-1) - \beta_r - dd_{mkf}), 0\}$ $-(\max\{(C_{nkf}(r-1) - dd_{nkf}), 0\} - \max\{(C_{nkf}(r) - dd_{nkf}), 0\})$, where dd_{mkf} and t_{mkf} are the due date and total processing time of order m of item k of family f, respectively, whereas $C_{mkf}(r-1)$ and $C_{mkf}(r)$ are the completion times of the same order in iterations $r-1$ and r, respectively.

Step 3.4. *Backward shift*: If an order, L_{nkf}, is shifted backward, then let set $C: \{m \mid \text{in } S_{r-1}, n < m; \text{ in } S_r, m < n\}$. $\Delta_r^b = (\max\{(C_{nkf}(r) - dd_{nkf}), 0\} - \max\{(C_{nkf}(r-1) - dd_{nkf}), 0\}) - \sum_{m \in C} \max\{(C_{mkf}(r-1) - t_{nkf} - \beta_r - dd_{mkf}), 0\} - \sum_{m \in B} \max\{(C_{mkf}(r-1) - \beta_r - dd_{mkf}), 0\}$.

Step 3.5. Let $\Delta_r = \min\{\Delta_r^f, \Delta_r^b\}$. Among all the alternative schedules in $\mathcal{S}$, identify the non-dominated schedules.

Step 3.6. For each non-dominated schedule, calculate the unified trade-off value as $F_r(S_r) = y_1(S_r, \beta_r) - y_2(S_r, \Delta_r)$, and select the one that gives the maximum $F_r(S_r)$ value. Set $\delta_r = \delta_{r-1} + \beta_r$. If $\delta_r \geq \Upsilon$, go to Step 4. Otherwise, set $r = r + 1$ and go to Step 3.1.

Step 4: *Phase II*: Use the same search heuristic explained in Steps 3.1 to 3.6 to find the best schedule, S^*, and the non-dominated schedule found at the end of Step 3, Phase I, becomes a starting solution. Phase II maintains the feasibility by searching within the feasible region, and the only difference is the stopping condition. We generate a sequence of schedules until $F_r(S_r^*) < F_{r-1}(S_{r-1}^*)$.

In the proposed algorithm, we first find an initial schedule to minimise the average tardiness and check its feasibility in Step 2. The two-phase approach in Steps 3 and 4 is analogous to the two-phase simplex method. In Phase I of the proposed heuristic, the objective is to achieve feasibility with as little deviation from the current schedule as possible. If a feasible schedule exists, then Phase II retains feasibility, and searches for the best solution while staying in the feasible region. The proposed multi-objective search heuristic first identifies a list of candidate schedules in Step 3.1 for each bottleneck resource leading to a set of non-dominated schedules at each iteration. Each candidate schedule corresponds to a single change of the order of items which results in a decrease in the amount of total setup time and the number of setup occurrences, although it might increase total tardiness. By merging two items from the same family that are apart from each other, one setup occurrence can be saved. Therefore, a list of candidate schedules is identified that can decrease the total setup time. The following example describes the forward and backward shifts.

Numerical Example 4.1: During the search procedure, we only consider the forward or backward shifts that could make the corresponding MS_{hf} value equal to zero. For example, let S_{r-1} be $\{\ldots, L_{111}, L_{233}, L_{132}, L_{211}, L_{142}, \ldots\}$, and customer order L_{211} is shifted forward and appended to L_{111}. The total setup time in the new schedule S_r, $\{\ldots, L_{111}, L_{211}, L_{233}, L_{132}, L_{142}, \ldots\}$ would become $TS_r = TS_{r-1} - (S_{12} + S_{21})$ (or $\beta_r = S_{12} + S_{21}$). In this particular schedule, the lot size of item 1 is increased from L_{111} to $L_{111} + L_{211}$. In the forward shift, an order is appended to the end of a group of the same family, and in the backward shift an order is appended to the beginning of a group of the same family. This shift combination possesses primal–dual characteristic, where the current schedule is the best one for the tardiness criterion, and we are searching for a better schedule for the setup cost criterion by reducing the number of setups, which also reduces the average flow time and makespan, while staying as close as possible to the current schedule.

The proposed search heuristic can be defined as a map $\alpha: R \rightarrow X$, where X is a list of candidate schedules of R. The map α is referred to as a generating mechanism for the heuristic. That is, $\alpha(s) \subseteq R$ for each $s \in R$, and is called a neighbourhood of R. Furthermore, $y \in \alpha(s)$ is a member of R if it is a result of either a forward shift,

$$y = 1, 2, \ldots, i-1,\ i,\ \mathbf{j}, i+1, \ldots, j-1, j+1, \ldots, n, \quad \text{for } 1 \leq i < j \leq n, \tag{23}$$

for which items i and j belong to the same family but not item $i+1$, or a backward shift,

$$y = 1, 2, \ldots, i-1,\ i+1, \ldots, j-1,\ \mathbf{i}, j, j+1, \ldots, n, \quad \text{for } 1 \leq i < j \leq n, \tag{24}$$

for which items i and j belong to the same family but not item $j-1$. The proposed search heuristic begins with an initial schedule S_0 and selects from the set $\alpha(S_0)$ a list of schedules for which $\beta_1 > 0$. In Phase II we utilise the proposed pricing mechanism to handle multiple criteria, and select the best schedule such that $F_1(S_1) > F_0(S_0)$, if one exists. In this fashion, the heuristic generates a sequence of schedules $\{S_r\}$, where $S_{r+1} \in \alpha(S_r)$ and $F_{r+1}(S_{r+1}) > F_r(S_r)$, terminating only when the best S^* is found. That is, S^* is a schedule such that $F(S^*) \geq F(S)$ for all $S \in \alpha(S^*)$.

In summary, during the search procedure, a set of candidate schedules is first defined at each iteration. Each candidate schedule corresponds to a change of the order of two lots which results in a possible increase in the amount t_2 of total tardiness, and a decrease in the amount t_1 of the total setup time; as a result the total number of setups is also reduced. A set of non-dominated schedules from the candidates schedules is defined in Step 3.5 based on the following conditions.

Condition 1: Given a current schedule Q, if there is a candidate schedule S_r such that $t_2(S_r) \leq 0$ and $t_1(S_r) > 0$, then the current schedule Q is a dominated schedule.

Condition 2: Given a set of candidate schedules $S_r \in \mathcal{S}$, if there is a candidate schedule S_v such that $t_1(S_v) < t_1(S_z)$ and $t_2(S_v) \geq t_2(S_z)$ for any $S_z \in \mathcal{S}$, then the candidate schedule S_v is a dominated schedule.

Condition 3: Given a set of candidate schedules X, the set of non-dominated schedules $\mathcal{N}$ is defined as $\mathcal{N} = \{x : x \in X$, there exists no other $x' \in X$ such that either $t_1(x') > t_1(x)$ and $t_2(x') = t_2(x)$, or $t_2(x') < t_2(x)$ and $t_1(x') = t_1(x)\}$.

Briefly, Condition 1 considers the current schedule and shows that it may be dominated by one of the candidate schedules. Similarly, Condition 2 considers the set of candidate schedules and recognises the schedules that will yield new solutions that are 'not as good' as can be found by choosing another candidate schedule. Condition 3, on the other hand, allows us to consider all the candidate schedules and defines the non-dominated ones. The best schedule from the set of non-dominated schedules is found by simultaneously considering the multiple performance criteria with an *a priori* objective weighting scheme based on a trade-off function technique as discussed in the previous section. The following example outlines how the proposed pricing mechanism and multi-objective search algorithm work.

Numerical Example 4.2: Let the total setup time at iteration $r-1$ for the incumbent schedule be equal to 60, and the total tardiness be equal to 200. At iteration r, the following four candidate schedules are found as a result of either a forward or a backward shift. Let {Candidate schedule ‖ Total setup time, Total tardiness}: $\{S_1 \| 55, 215\}$, $\{S_2 \| 56, 210\}$, $\{S_3 \| 54, 214\}$, and $\{S_4 \| 58, 212\}$. The non-dominated schedules are S_2 and S_3, because schedule S_1 is dominated by schedule S_3 (i.e. $55 > 54$ and $215 > 214$), and schedule S_4 is dominated by schedule S_2 (i.e. $58 > 56$ and $212 > 210$). For this particular iteration, S_2 is better than S_3 with respect to the total setup time criterion, whereas S_3 is better in terms of the total tardiness criterion. In order to compare and correspondingly select the best schedule S_r at iteration r, we utilise the trade-off functions to unify the multiple objectives into a single value as discussed in Section 2. In these functions, the associated parameters c_1, a_1, c_2 and a_2 are found by using the dual values associated with the resource and inventory balance constraints in the optimal solution of the cell loading problem solved at the first level. For the non-dominated schedule $S_2, F_r(S_2) = (-c_1(1 - e^{a_1(56)})) - (-c_2(1 - e^{a_2(210)}))$. For the non-dominated schedule S_3, $F_r(S_3) = (-c_1(1 - e^{a_1(54)})) - (-c_2(1 - e^{a_2(214)}))$. Consequently, we select the one with the maximum unified trade-off value.

As a main contribution, the proposed approach allows more accurate portrayal of the operation of CM systems by using the capacity constraints to assess the impact of the cell loading decisions on the lot sizing and cell scheduling problems through a set of dual variables associated with the resource and inventory balance constraints. Coordination between the decision levels was achieved using the proposed pricing mechanism instead of using a top-down constrained approach as is typically done. Pricing information from dual variables gives global information to more localised lot sizing and cell scheduling level decision making. The pricing mechanism calculates the economics of the CM system which becomes a portion of the scheduling cost at the lower level by varying parameters in the corresponding objective function. As a result, lower level searches focus in areas most likely to contain good solutions. In addition, the number of non-dominated schedules in the objective space is reduced, which shortens the search process significantly, since the best schedule is chosen from the set of non-dominated schedules.

5. Computational results

A hierarchical, multi-objective modelling approach and solution technique developed in this paper to solve cell loading and scheduling problems in cellular manufacturing systems is evaluated empirically in this section to determine its quality. We wish to investigate to what degree the proposed approach is robust in the face of uncertainty and how sensitive it is to the assumptions we have made throughout this research with regard to machine-component groupings, GT cells and families, the inclusion of new products in the existing families, and resource availability. There are a large number of variables that could have an effect on our results. Within the conceptual framework of the hierarchical procedure, an experimental design is developed with two objectives in mind. The first objective is to generate a set of test problems to compare the results of the proposed approach with other scheduling rules. The second objective is to explain the relationships between the dependent variables (such as tardiness, flow time, number of tardy orders, earliness, and makespan) and the system parameters.

There are six experimental factors that can affect the efficiency of the proposed approach, which are listed in Table 3. The initial estimation of lot sizes depends on the direct setup cost for each item, and plays an important role within the context of the trade-offs between inventory holding and overtime costs since it determines the number of setups required. The direct setup cost, to make the results of the research meaningful, must be compared with the inventory holding cost as a ratio, S/I, as suggested by Maes and Van Wassenhowe (1986). The S/I ratios, factor A, are used to find the initial lot size of family i in cell j in period t as $Q_{ijt} = \sqrt{2 \cdot \text{S/I ratio} \cdot df_{it}}$ for every $i \in FS(j)$. In order to assign a due date to each customer order, we utilise the total work content rule, which assigns due dates proportional to the product of the total processing time of an item, t_{ikf}, and parameter F, the flow allowance factor. The flow allowance factor (factor B) essentially controls the tightness of the due dates, such as the due date of order i of item k of family f, $dd_{ikf} = F \cdot t_{ikf}$. Factor B is set at two levels to generate tight and loose due-date requirements.

Table 3. Experimental factors.

Factor	Definition	Low	High
A	S/I ratio	0.75	1.25
B	Flow allowance factor	20	24
C	Number of families	15	35
D	Upper limits on resource availability	No idle time	10% idle time
E	Number of GT cells	5	10
F	Set of items within each family	Low variability	High variability

The representative ranges for factor C, the number of families, and factor E, the number of cells, are based on the study of Hyer and Wemmerlöv (1989) on current GT practices seen in industry. In addition, an assignment of the items to the GT families is one of the objectives of the part-family and machine-cell formation problem. These assignments are done depending upon the similarities that exist between the items, and similarity coefficients are calculated using several criteria, as discussed by Offodile *et al.* (1994). Factor F, the number of items in each family, reflects the fact that the variability within each family could be different depending upon the threshold values used to form the GT families. A high threshold value means a low feature variability and a high similarity among the items in each family. Since the total number of items is a fixed parameter for all runs, factor F is used to measure the impact of variability in each family by varying the size of each family, and the processing time for each item k on resource l, PR_{kl}, is based upon the degree of similarity in each family, as can be seen in Table 4. The levels of factor D specify the upper limits on resource availability, where the low level corresponds to a congested shop floor, while the high level represents 10% idle time. Since there are six factors and two levels, our experiment is a 2^6 full-factorial design, which corresponds to 64 treatment combinations.

Other variables in the system are treated as fixed parameters and are summarised in Table 4, where UN $\sim [a, b]$ represents a uniformly distributed random variable in interval $[a, b]$. All of the parameter values are constant throughout the planning horizon, except for the inventory holding cost for family i in period t, h_{it}, which increases over time to approximately account for factors such as inflation and the time value of money. Furthermore, there are other parameters, such as df_{it} and a_{ij}, that assume fixed parameter values. The effective demand for each item in each period, d_{kt}, is fixed. Therefore, the demand for each family in a particular period should be calculated by summing the demands of all of the items belonging to that family corresponding to that period, i.e. $df_{it} = \sum_{k \in TI(i)} d_{kt}$, $\forall i, t$. The processing time of each item at each feasible resource, PR_{kl}, and the number of operations for each item are fixed. Therefore, the average total time

Table 4. Fixed parameters.

Parameter	Set of values
Total number of items, K	250
Total number of resources, L	50
Number of periods, T	24
Cost of production, C_{ijt}	UN $\sim [0.75, 1.25]$ if $i \in P_j$ UN $\sim [1.5, 2.0]$ if $i \in S_j$
Cost of regular time, r_{jt}	UN $\sim [1.25, 2.0]$
Cost of overtime, o_{jt}	$2 * r_{jt}$
Inventory holding cost, h_{it}	$(1 + 0.003(t-1)) *$ UN $\sim [1.5, 2.5]$
Setup cost for the families, BS_{ij} and BP_{ij}	$0.03 * C_{ijt} * d_{it}$
Processing times, PR_{kl}	
(i) Low variability	UN $\sim [0.25, 0.35]$ if $i \in P_j$ UN $\sim [0.45, 0.55]$ if $i \in S_j$
(ii) High variability	UN $\sim [0.2, 0.4]$ if $i \in P_j$ UN $\sim [0.4, 0.6]$ if $i \in S_j$
Setup times between families	UN $\sim [2, 3]$
Setup times on each resource	$(0.1 * a_{ij})$ if $i \in FS(j)$
Number of operations per item	UN $\sim [3, 5]$
Effective demand, d_{kt}	UN $\sim [6, 15]$

required to produce one unit of family i at cell j, a_{ij}, should be the product of the average processing time of an item belonging to family i at cell j and the average number of operations required for item k in family i. The mathematical expression is

$$a_{ij} = \frac{(\sum_{k \in TI(i)} \sum_{l \in LR(j)} PR_{kl})(\sum_{k \in TI(i)} NO(kj))}{NI(i)^2 \cdot NR(j)}, \quad \forall i \in FS(j). \tag{25}$$

The proposed algorithm (denoted HMA) was compared with the shortest weighted processing time (SWPT) method, which sequences the customer orders in the non-decreasing order of their weighted lot processing time of (Setup time/(lot size) + (Total processing time)), the earliest due-date (EDD) rule, which sequences the customer orders in non-decreasing order of the due dates, dd_{ikf}, the apparent tardiness cost (ATC) rule, and the initial schedule found in Step 1 of Algorithm 1 of the Wilkerson and Irwin heuristic (INIT). Under the apparent tardiness cost (ATC) rule, jobs are scheduled one at a time, i.e. every time the machine becomes free a ranking index is computed for each remaining job. The job with the highest ranking index is then selected to be processed next. The ranking index is a function of the time $tnow$ at which the machine became free, as well as t_{ikf} and dd_{ikf} of the remaining jobs as well as the required setup time, S_{hf}, switching from family h to family f, i.e. $S_{ff} = 0$, and the average processing time of the remaining jobs, p_{avg}, at the current time, $tnow$. For each unscheduled job at time $tnow$, we calculate the ATC priorities as follows:

$$\pi_{ikf}(tnow) = \frac{1}{S_{hf} + t_{ikf}} \exp(-\max(0, dd_{ikf} - tnow - (S_{hf} + t_{ikf}))/p_{\text{avg}}). \tag{26}$$

The reason for choosing the above heuristics for the computational analysis is that it has been shown that each heuristic works well for one or more performance measures. The SWPT rule is known to be effective with respect to minimising the average flow time and minimisation of the number of tardy items in highly congested shops. The EDD rule is commonly used in practice and performs well for the due-date related-criteria. The Wilkerson and Irwin (W&I) heuristic is based on the pairwise interchange search idea and provides good solutions, on average, for the single machine tardiness problem. The ATC sequencing rule has been shown to be superior to other sequencing rules for the $1 \| \sum w_i T_i$ problem.

In most production environments, schedule evaluation is based on a mixture of several performance criteria, since an optimal schedule for one performance measure is not necessarily optimal for other performance measures. Therefore, the comparison of different rules was based on multiple performance criteria such as minimising the average tardiness, $\bar{T}$, the average flow time, $\bar{F}$, the number of tardy items, NT, the average earliness, $\bar{E}$, and the makespan, $C_{\max}$. For each run, the heuristic from the five mentioned above that yielded the best schedule for a certain performance measure is called the best heuristic for that measure. Thus, for each performance measure in each individual run, the best heuristic could be different.

Since the performance measures are not expressed in commensurable terms, a scaling function, θ_{ip}, for each heuristic p is defined to ensure the same range for each performance measure i. This range corresponds to the interval (0, 1), where 0 indicates the best value and 1 indicates the worst value. This scaling is accomplished by defining the scaling function as

$$\theta_{ip} = \frac{Z_{ip} - Z_i^*}{Z_i^{**} - Z_i^*}, \quad \text{for each } p, \tag{27}$$

Table 5. Performance measures and scaled values: best scenario for HMA.

Performance measure	Heuristic				
	SWPT	ATC	EDD	INIT	HMA
$\bar{T}$	340.3	120.3	109.0	101.0	61.7
	(1)	(0.21)	(0.17)	(0.14)	(0)
$\bar{F}$	1092.4	1240.6	1229.5	1221.4	1129.7
	(0)	(1)	(0.93)	(0.87)	(0.25)
NT	72	101	100	99	56
	(0.36)	(1)	(0.98)	(0.96)	(0)
$\bar{E}$	392.2	24.0	23.8	23.8	76.3
	(1)	(0)	(0)	(0)	(0.17)
$C_{\max}$	2411.8	2774.7	2614.2	2608.8	2404.1
	(0.04)	(0.81)	(1)	(0.97)	(0)

where Z_{ip} is the objective function value of heuristic p for performance measure i, Z_i^* is the best value for the performance measure $i = \min_p\{Z_{ip}\}$, and Z_i^{**} is the worst value for the performance measure $i = \max_p\{Z_{ip}\}$.

As an example, we summarise the best factor combination (among all the randomly generated runs in our computational experiments) for the proposed HMA in Table 5. For this particular factor combination, factors A, B and C are set at their high levels, and factors D, E and F are set at their low levels. The best case corresponds to the case where the number of families is equal to the maximum level, and the number of cells to the minimum; as a result, several families were allocated to each cell, which means that the size of the scheduling problem at each cell is greater. Consequently, the setup time between families becomes a critical issue. In addition, due to the factor combinations in this run, there is a significant setup cost, and the upper limit on resource availability is set at the low level (no idle time), which increases the problem complexity. Since the primary advantage of the HMA is increased solution flexibility, it can compare several trade-offs and vary lot sizes and sequences to find the 'best' schedule for various performance criteria. Therefore, it can handle the scheduling complexity more effectively than other heuristics by creating more alternatives to be considered.

Table 6 summarises the ranges for the average scaled deviations for each heuristic for each individual measure along with the overall averages and ranges for each heuristic. The proposed HMA performed significantly better than the others in all runs and increased its superiority when the scheduling complexity was increased as discussed above. On the average scaled deviation function, θ_{ip}, the HMA is 3.3 times better than the SPT rule, 2.8 times better than the ATC, and 3.2 times better than the EDD rule. In addition, the initial schedule given by the W&I heuristic is improved by three times on the average scaled deviation. The ranges of the average scaled deviations for each heuristic indicate that even the worst case result for the HMA is better than the others, which indicates its robustness to widely varying conditions.

Obviously, the scaling function, θ_{ip}, and, especially, its denominator, $(Z_i^{**} - Z_i^*)$, have a significant impact, which may be attributed to the average scaled deviations for each heuristic. If the ranges for the actual values, $(Z_i^{**} - Z_i^*)$, for the individual performance measures are wide enough, then the average scaled deviations can be considered

Table 6. Averages and ranges of scaled deviations.

Performance measure	Heuristic				
	SWPT	ATC	EDD	INIT	HMA
$\bar{T}$	1	[0, 0.38]	[0, 0.61]	[0, 0.38]	[0, 0.04]
$\bar{F}$	0	[0.64, 1]	[0.91, 1]	[0.78, 1]	[0.17, 0.8]
NT	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 0.33]
$\bar{E}$	1	[0, 0.02]	[0, 0.04]	[0, 0.04]	[0.03, 0.33]
$C_{\max}$	[0, 0.16]	[0.27, 1]	[0.91, 1]	[0.87, 1]	[0, 0.31]
Overall average	0.5366	0.4580	0.5203	0.4993	0.1643
Overall range	[0.4, 0.632]	[0.4, 0.6]	[0.372, 0.71]	[0.372, 0.638]	[0.078, 0.258]

Table 7. Averages and ranges of scaled deviations.

Performance measure	Heuristic					
	SWPT	ATC	EDD	INIT	HMA	$Z_i^{**} - Z_i^*$
$\bar{T}$	[278.5, **374.1**]	[23.2, 195.1]	[23.1, 264.3]	[23.1, 205.9]	[**5.4**, 124]	368.7
$\bar{F}$	[**1036**, 1092.4]	[1173.6, 1335.8]	[1174.9, **1365.1**]	[1175.1, 1344.5]	[1113.4, 1174.8]	329.1
NT	[60, 125]	[43, **226**]	[41, 225]	[41, 223]	[**29**, 99]	197
$\bar{E}$	[363.3, **491.2**]	[**3.7**, 67.1]	[**3.7**, 67.1]	[4, 67.1]	[28.8, 149.4]	487.5
$C_{\max}$	[**2339.5**, 2420]	[2440.6, 2687.6]	[2501, **2825.6**]	[2501, 2803.9]	[2365, 2466.5]	486.1

representative of significant relative differences between each heuristic. Table 7 summarises the best and worst case results for each heuristic for each performance measure, and the objective function differences, $(Z_i^{**} - Z_i^*)$. As can be seen, the differences are large enough to allow the average scaled deviations to indicate significant differences among the heuristics. Furthermore, each heuristic is a non-dominated one for a certain local performance measure; e.g., SPT is the best for the average flow time criterion, ATC and EDD for the average earliness criterion, and HMA for both the average tardiness and the number of tardy items criteria. But, for the overall multiple criteria objective function, HMA performs significantly better than the others.

The final question is why does the HMA perform significantly better than the other heuristics. There could be several reasons. (i) In this paper, instead of solving cell loading, lot sizing and cell scheduling problems independently, they all become levels in the proposed decision hierarchy. The interactions between these problems are utilised to increase flexibility by adding more attractive alternatives to be considered. The proposed pricing mechanism coordinates the different levels by varying their parameters in their objective functions, and provides 'soft' constraints to the higher level problems through feedback-based adjustments in resource availability, which focuses lower level searches in areas most likely to contain good solutions. For example, in a highly congested system, the dual values for the bottleneck resources will be relatively high, and, as a result, the search heuristic seeks to minimise the total setup time on the bottleneck resources, consequently reducing the average flow time and makespan. When the dual values for resource

Table 8. Analysis of variance results.

Performance measure	Factor A	B	C	D	E	F
$\bar{T}$	a	a	a	a	a	b
$\bar{F}$	a	–	a	c	a	a
NT	a	a	a	a	a	–
$\bar{E}$	–	a	a	a	a	–
$C_{\max}$	a	–	a	c	a	a

Significant at the (a) 0.5% level, (b) the 2.5% level, and (c) the 25% level.

constraints are small, due-date-related criteria, i.e. the average tardiness and number of tardy items, become relatively more important than the setup cost criterion, so due dates can be easily met. As a result, the number of non-dominated schedules in the objective space is reduced, which shortens the search process significantly, given that the best schedule is chosen from the set of non-dominated schedules. (ii) Another key issue is that lot sizes and job sequences should change as the cell loading decisions change over time to compensate for the scheduling restrictions, so that due dates can be more easily met, as opposed to having fixed lot sizes as is widely used in other models in the literature. (iii) Another advantage is that the HMA makes an effort to consider all of the performance criteria instead of a single criterion. As a result, the 'best' solution is sought for the set of criteria instead of seeking a best solution for a local goal.

The second objective of the computational study was to explain the relationship between the multiple performance criteria and the six factors considered for the experimental design. We performed an analysis of variance (ANOVA) test to understand the sensitivity of the proposed approach as summarised in Table 8. Factor A, S/I ratio, represents the relationship between the setup cost and the inventory holding cost. Therefore, it affects the lot sizes for each item, and consequently has a significant effect on the average flow time and the makespan criteria. Factor A also has a significant effect on the average tardiness and the number of tardy items criteria, which shows again that there is a significant interaction between lot sizes and completion times for each item, validating the importance of the inclusion of the second level, joint lot sizing and scheduling, in the proposed production planning hierarchy. Factor B, the flow allowance factor, is set at high and low levels to generate loose and tight due-date requirements, respectively, for each order. Therefore, factor B has a significant effect on the due-date-related criteria such as the average tardiness, the number of tardy items, and the average earliness. The number of GT families, factor C, and the number of GT cells, factor E, are the most important outputs of the GT machine-cell formation algorithms. Our results indicate that both factors have a significant effect on all of the performance measures considered. Therefore, the interface between the design and planning and scheduling decisions becomes a critical issue. Unfortunately, most of the GT machine-cell formation algorithms in the current literature do not consider parameters associated with the production planning and scheduling activities. A similar conclusion is also presented in a recent paper of Wang *et al.* (2010), who address the joint decision problem of cell formation and parts scheduling in a cellular manufacturing system.

The coordination mechanism between the levels, based on the set of dual values of the resource constraints at the cell loading level, influences the second level problem by varying parameters in its objective function. In addition, the upper limits on resource availabilities at the cell loading level were adjusted depending upon the feasibility status feedback from the lower level. Since lot sizes and scheduling decisions were based on the dual values, upper limits on resource availability, factor D, had an effect on the average flow time and the makespan criteria at only the 25% level of significance. This was due to the fact that the feasibility status feedback mechanism and pricing adjustments of the resources at the scheduling level softened the effect. However, it does have a significant effect on all of the due-date-related criteria at the 0.5% level of significance, since it affects the amount of time available on each resource. The number of items in each family, factor F, was used to represent the variability of the items in each family. If there is variability in the characteristics of the items in each family, more inventory is the natural consequence, which is not desirable. ANOVA tables show that the variability had the most significant effect on the average flow time, which is directly proportional to the in-process inventories. Therefore, if the variability of the items in each family is reduced, then the in-process inventory level will be reduced.

6. Conclusion

In this study, instead of solving the cell loading, lot sizing, and cell scheduling problems independently, they all became the levels of the proposed decision hierarchy. We can state the contributions of this paper as follows. First, the interactions between these problems were exploited to increase flexibility by adding more alternatives to be considered and directing the search for a solution in a more favourable area of the solution space. The proposed pricing mechanism coordinates the different levels by varying the parameters in their objective functions, and provides 'soft' constraints to the higher level problem through a feedback mechanism. Another key issue is that lot sizes and cell schedules change as the cell loading decisions change over time to compensate for scheduling restrictions; therefore, due dates can be more easily met. An advantage arises from simultaneously considering multiple performance criteria with an *a priori* objective weighting scheme based on a trade-off function technique instead of optimising a local performance measure at each level. The proposed multi-objective search heuristic initially identified a list of candidate schedules leading to a non-dominated schedule at each iteration. The best schedule from the set of non-dominated schedules was found based on the trade-off functions, which transformed all of the performance measures into a common metric (monetary terms). The parameters of the proposed trade-off functions are derived from a pricing mechanism that calculates the economics of the CM system resources through a set of dual variables. Consequently, a unified total cost function, which captures multiple and usually conflicting objectives at different levels, is minimised instead of seeking a best solution for a local goal. As future research, the most logical extension would be the inclusion of the cell formation problem in the proposed decision-making hierarchy.

References

Akturk, M.S. and Yayla, H.M., 2005. Management of product variety in cellular manufacturing systems. *International Journal of Flexible Manufacturing Systems*, 17 (2), 93–117.

Allahverdi, A., Gupta, J.N.D., and Aldowaisan, T., 1999. A review of scheduling research involving setup considerations. *OMEGA*, 27 (2), 219–239.

Allahverdi, A., *et al*., 2008. A survey of scheduling problems with setup times or costs. *European Journal of Operational Research*, 187 (3), 985–1032.

Azizoglu, M. and Webster, S., 2003. Scheduling parallel machines to minimize weighted flowtime with family setup times. *International Journal of Production Research*, 41 (6), 1199–1215.

Baker, K.R., 1999. Heuristic procedures for scheduling job families with setups and due dates. *Naval Research Logistics*, 46 (8), 978–991.

Baker, K.R. and Trietsch, D., 2009. *Principles of sequencing and scheduling*. New York: Wiley.

Biggs, J.R., 1979. Heuristic lot-sizing and sequencing rules in a multistage production-inventory system. *Decision Sciences*, 10 (1), 96–115.

Bokhorst, J.A.C., Nomden, G., and Slomp, J., 2008. Performance evaluation of family-based dispatching in small manufacturing cells. *International Journal of Production Research*, 46 (22), 6305–6321.

Crauwels, H.A.J., *et al*., 1998. Branch and bound algorithms for single-machine scheduling with batch set-up times to minimize total weighted completion time. *Annals of Operations Research*, 83, 59–76.

Hyer, N.L. and Wemmerlöv, U., 1989. Group technology in the U.S. manufacturing industry: A survey of current practices. *International Journal of Production Research*, 27 (2), 1287–1304.

Maes, J. and Van Wassenhowe, L.N., 1986. Multi item single level capacitated dynamic lotsizing heuristics: A computational comparison. *IIE Transactions*, 18 (2), 124–129.

Offodile, O.F., Mehrez, A., and Grznar, J., 1994. Cellular manufacturing: A taxonomic review framework. *Journal of Manufacturing Systems*, 13 (3), 196–220.

Potts, C.N. and Van Wassenhove, L.N., 1992. Integrating scheduling with batching and lot-sizing: A review of algorithms and complexity. *Journal of the Operational Research Society*, 43 (5), 395–406.

Potts, C.N. and Kovalyov, M.Y., 2000. Scheduling with batching: A review. *European Journal of Operational Research*, 120 (2), 228–249.

Suer, G.A., Saiz, M., and Gonzalez, W., 1999. Evaluation of manufacturing cell loading rules for independent cells. *International Journal of Production Research*, 37 (15), 3445–3468.

Suer, G.A., Arikan, F., and Babayigit, C., 2008. Bi-objective cell loading problem with non-zero setup times with fuzzy aspiration levels in labour intensive manufacturing cells. *International Journal of Production Research*, 46 (2), 371–404.

Suer, G.A., Subramanian, A., and Huang, J., 2009. Heuristic procedures and mathematical models for cell loading and scheduling in a shoe manufacturing company. *Computers & Industrial Engineering*, 56 (2), 462–475.

Suresh, N.C. and Gaalman, G.J.C., 2000. Performance evaluation of cellular layouts: Extension to DRC system contexts. *International Journal of Production Research*, 38 (17), 4393–4402.

Venkataramanaiah, S., 2008. Scheduling in cellular manufacturing systems: A heuristic approach. *International Journal of Production Research*, 46 (2), 429–449.

Wang, X., Tang, J., and Yung, K.L., 2010. A scatter search approach with dispatching rules for a joint decision of cell formation and parts scheduling in batches. *International Journal of Production Research*, 48 (12), 3513–3534.

Wemmerlöv, U. and Johnson, D.J., 1997. Cellular manufacturing at 46 user plants: Implementation experiences and performance improvements. *International Journal of Production Research*, 35 (1), 29–49.

Wilkerson, L.J. and Irwin, J.D., 1971. An improved algorithm for scheduling independent tasks. *AIIE Transactions*, 3 (3), 239–245.

Ying, K.C., *et al*., 2010. Permutation and non-permutation schedules for the flowline manufacturing cell with sequence dependent family setups. *International Journal of Production Research*, 48 (8), 2169–2184.