ORIGINAL ARTICLE

Flow time analyses of a simulated flexible job shop by considering jockeying

Adil Baykasoğlu & Zeynep D. U. Durmuşoğlu

Received: 2 August 2010 /Accepted: 15 February 2011 / Published online: 14 June 2011 \circ Springer-Verlag London Limited 2011

Abstract It has been essential to include flexibility in manufacturing policy making since variability in demand and products are considerably increasing. However, it is important to know and to monitor the proper level and type of flexibility that is required to obtain full benefits from it. This paper analyses the effects of flexibility on flow time performance of a simulated job shop. For that purpose, several scenarios are developed under four flexibility levels with two different machine selection rule and three types of dispatching rules. Furthermore, effect of jockeying as a queuing policy on the flow time performance is also investigated through simulation modeling. Results indicated that full flexibility is a preferable state for most of the cases. However, in some cases, chain configurations perform similar results since it combines the benefits of pooling and specialization. In addition, it is observed that a queue control mechanism like jockeying is an effective way to improve performance even though it may increase complexity of controlling policy.

Keywords Flexibility Job shop manufacturing . Simulation .Jockeying

1 Introduction

In today's highly dynamic and turbulent environment with unprecedented and unpredicted events, managers of organ-

Z. D. U. Durmuşoğlu Department of Industrial Engineering, University of Gaziantep, Gaziantep, Turkey

A. Baykasoğlu (⊠) Faculty of Engineering, Department of Industrial Engineering, Dokuz Eylul University, Tinaztepe Campus, Buca, 35160 Izmir, Turkey e-mail: adil.baykasoglu@deu.edu.tr

izations are forced to continuously redefine their activities and strategies in order to produce "right products" at the "right place", at the "right time", in the "right quantity". Baines et al. [\[1](#page-13-0)] indicated that the manufacturing capabilities establishes the competitiveness of a company's product and can be measured in terms of such variables as service, quality, and cost. Unfortunately, traditional sources of competitive advantages, such as high quality and low costs have became insufficient and firms began to seek for new skills to develop their core competency. One well-known and efficient strategy of developing core competency has been flexibility.

As being multi-disciplinary, concept flexibility has been defined in various forms by the researchers from different disciplines like engineering, architecture, biology, and economics, etc. Flexibility can be interpreted as the ability to respond and adapt to environmental changes easily and rapidly. This change can be chaotic or stochastic as well. However, current implementations show that increasing the amount of flexibility usually does not equally increase the "ability to cope with the change [\[2\]](#page-13-0)". It is discussed by several authors [[3](#page-13-0)–[5](#page-13-0)] that there are some certain trades-offs among the flexibility and some other performance criteria. As mentioned by Baykasoglu [\[6\]](#page-13-0), increasing the variety in order to increase flexibility can increase the system complexity. Increased complexity can create controlling difficulty. Therefore, in many cases, fully flexible systems may generate uncompetitive performance which is mainly due to controlling difficulty or insufficient control. As an example, consider the case of scheduling a flexible manufacturing system with a conventional scheduler which is not able to effectively utilize alternative routes or alternative machine assignments. In such a case, most of the potentially available flexibility will be hidden to the controlling system. This will generate unexpected or uncompetitive performance in most of the time. The reason for undesired performance is not mainly due to high level of flexibility it is mainly due to ineffective control. We should note here that in order to utilize flexibility for

performance improvement, we must improve and advance the control system. As also noted by Baykasoglu [\[6](#page-13-0)], we should also keep in mind that in some cases additional flexibility may not be necessary for an existing production scenario at all. In such cases, flexibility (which is more than needed) might also be counterproductive. Here, we can say that design purpose of the system is very important (similar to [\[3\]](#page-13-0) in Fig. 1, we try to depict the behavior of well-designed systems under increased flexibility and control complexity conceptually). Therefore, we should consider flexibility as a parameter of the system design and closely relate it to the needs in order to obtain the optimal level of system performance. In this respect, flexibility needs to be adjusted considering all of the tradeoffs. In other words, the level and the type of flexibility are required to be adjusted and to be monitored for getting the expected benefits fully. On the other hand, the debates on the variety of definitions and classification of flexibility makes it difficult to decide on the "type of flexibility" to implement. At this instant, a fundamental research question arises about the adjustment of the flexibility. What kind/type/amount of flexibility makes a manufacturing system to perform better? These questions inspired us to analyze the effects of different flexibility levels on the flow time performance of job shops. We choose the average flow time as a performance indicator since the information about the shop status directly effects the flow time and the estimations made about it [[7](#page-13-0), [8\]](#page-13-0). Therefore, average flow time is the most sensitive and crucial performance indicator that can be affected from various flexibility scenarios.

In addition to the above discussions, the effect of jockeying as a queuing policy on flow time performance is also critical and requires to be investigated especially in manufacturing systems. Jockeying can be considered as an effective way to improve performance in some cases even though it may increase the complexity of the controlling policy. We should mention here that, this paper actually

presents one of the first attempts to explicitly model and investigate effects of jockeying on job shop manufacturing.

In the present study, a simulation model of a hypothetical job shop is developed in SIMAN as a test bed, since simulation is a useful methodology for understanding the high-level dynamics of a complex manufacturing system. By making use of the simulation model, scenarios are developed under four level of flexibility with two different machine selection rules and three types of dispatching rules.

2 A brief review on flexibility and its effects on manufacturing systems

Due to increasing interest and applications of flexibility, there has not been a consensus on the definition and classification of it. Efforts for defining "flexibility" and classifying its varieties in a proper way is still a research subject. One of the earliest definitions of flexibility was made by Jan [[9\]](#page-13-0). He defined flexibility as "a buffer for the effects of a changing environment". On the other hand in a manufacturing context, Tsubone and Horikawa [[10\]](#page-13-0) defined flexibility as the "ability of a system to adapt quickly to any changes in relevant factors such as product, process, workload, or machine failure"

In parallel to the efforts on defining flexibility, there have been also several attempts to classify it. Buzacott [\[11](#page-13-0)] classified flexibility in two main groups as "job flexibility", which is the ability of the system to cope with changes in the jobs to be processed by the system, and "machine flexibility", which is the ability of the system to cope with changes and disturbances at the machine and workstations. One another classification has been performed by Browne et al. [[12\]](#page-13-0). Slack [[13](#page-13-0)] considered "*range and response*" concepts on the classification of flexibility. Another classification which is based on the classification of Browne et al. [[12\]](#page-13-0) was proposed by Sethi and Sethi [\[14](#page-13-0)]. They added three types

Fig. 1 Well-designed system's behavior under increased flexibility and complexity

Ideal or good system behaviour under effective, well designed control

of flexibility, which were "material handling flexibility", "program flexibility", and "market flexibility", to the classification of Browne et al. [\[12\]](#page-13-0). Chen et al. [[15](#page-13-0)] also analyzed flexibility using a more general perspective. They classified flexibility into three main classes as "*product flexibility*", "marketing flexibility", and "infrastructural flexibility". In their study, machine, process, routing, material handling, and programming flexibility grouped in the first class, product, volume, mix, and expansion flexibility grouped in the second class and flexibility of an organization took place in the last class. Similar to Chen et al. [\[15\]](#page-13-0), Benjaafar and Ramakrishn [\[16\]](#page-13-0) classified flexibility into two groups as follows: "*product* related flexibility" and "process related flexibility".

Apart from the aforementioned classification schemes, some authors have categorized flexibility by means of their hierarchy. Taymaz [[17\]](#page-13-0) analyzed flexibility by considering three levels; "component level flexibility", "operations level flexibility" and "system level flexibility". In a similar manner, some other authors [\[3,](#page-13-0) [18](#page-13-0), [19\]](#page-14-0) classified flexibility as being strategic (long-term flexibility), tactical (mid-term), and operational (short-term). "Strategic flexibility" is a firm's ability to respond to uncertainties by adjusting its objectives with the support of its superior knowledge and capabilities [\[20\]](#page-14-0). "*Tactical flexibility*" is about the specific performance features of manufacturing such as machine utilization, workin-process inventory, and efficiency. "Operational flexibility" is concerned with unpredictable problems like machine breakdowns, unreliable raw material supplies encountered during day-to-day manufacturing operations. Toni and Tonchia [[21\]](#page-14-0) scanned the literature on flexibility and grouped the existing classifications considering their ontology. They grouped classifications via their phases, hierarchies, temporality, and by the object of variation.

Based on the literature review, it seems that unfortunately there is not a consensus about the classification as well as definition of flexibility. Some researchers like Oke [\[22\]](#page-14-0) investigated the sources of this confusion. According to Oke [\[22](#page-14-0)], the fundamental reason of this confusion has been entitling of the same type of flexibility with different names. Another reason has been use of "flexibility" term as it is equivalent of "capability". The third reason of this confusion has been the relativity of flexibility since the meaning of flexibility differs according to the market, manufacturing system level, and the actual tools and techniques used to deliver it.

Similar to the case for the defining and classifying flexibility, many studies have also been issued for measurement and evaluation of flexibility [\[6,](#page-13-0) [23](#page-14-0)–[26](#page-14-0)]. There is also no universally accepted method for measuring flexibility yet. This variety of proposals on the measurement is explained by Gupta [\[23](#page-14-0)] with the nature of flexibility, its usage, and its variety to particular to situations. Similar to Gupta [[23\]](#page-14-0), Narain et al. [[27\]](#page-14-0) investigated the causes of difficulty on measuring flexibility and indicated that this difficulty origins

from the properties of flexibility concept such as being a measure of the potential rather than performance.

It is a very well-known fact that flexibility is a multidimensional concept with many definitions and interpretations [\[28](#page-14-0)]. As also noted by Sharma and Sushil [\[29](#page-14-0)] flexibility has many different connotations in different contexts and situations which make it very difficult to define, classify, and measure.

There have been considerable amount of studies on modeling manufacturing facilities and investigating the "manufacturing flexibility" and "manufacturing performance" for different manufacturing environments. However, findings have not been generally consistent with each other. Some researchers such as Gerwin [[30\]](#page-14-0) and Olhager [\[31](#page-14-0)] claimed that manufacturing flexibility is a critical sign to manufacturing performance. However, others, for instance Bengtson [[32](#page-14-0)] and Das and Nagendra [\[33\]](#page-14-0), presented that the relation between "manufacturing flexibility" is not directly proportional with "manufacturing performance". Since this work is intended to find effects of flexibility in job shop manufacturing, it is worthy to state how flexibility can be adjusted for job shop manufacturing environment.

In a typical job shop manufacturing environment, processing requirements for each job consist of a group of ordered tasks where each of them to be performed on a different machine. Ordered tasks, determine the route of a job in the shop. Since there might be lots of jobs processed by the shop, there will be lots of routings which are not particularly grouped. These routes are created by the jobs received from the other machines; or similarly jobs may be routed to next work centers or they may be disposed from the shop if completed. This variety of routings makes "work flow" a very sophisticated issue. Therefore, production control gets harder and it is often structured based on "queue management system". Therefore, a typical job shop can be defined as the combination of machines, routings, and the queues. For that reason, strategies for flexibility design in a job shop can be established through routing flexibility, machine flexibility, and some distinctive strategies for machine sharing, queue management, etc.

2.1 Machine sharing in job shops

Machine sharing can be defined as the capability of a group of machines in a manufacturing system to share the production of a set of different part types so that a part type in the set can be indiscriminately allocated to any machine in the group [\[34](#page-14-0)]. Benjaafar [[34\]](#page-14-0) also explained that machine sharing requires flexible part routing capabilities, allowing parts to be directed to any one of the shared machines, and sharing and/or duplication of some of the auxiliary resource. In the related literature, previous studies focused on two scenarios which are pooling and specialization.

2.1.1 Pooling as a flexibility strategy

In pooling, machines are set into functional pools, where each pool is associated with a set of operations, whereas in specialization, each machine is dedicated to a single operation type. Pooling is frequently employed in flexible manufacturing systems (FMS), cellular manufacturing systems, and job shops. Specialization is more common in dedicated flow lines where a sufficiently high volume of a single part type is being manufactured [[35\]](#page-14-0).

Some researchers [[36,](#page-14-0) [37\]](#page-14-0) investigated the subject of resource pooling in closed queuing network models of FMS and revealed that resource pooling increases the system throughput. In open queue network models, some authors [\[34](#page-14-0), [38](#page-14-0), [39\]](#page-14-0) found that part waiting times and work in progress (WIP) inventory can be reduced significantly by increasing pooling. However, Smith and Whitt [[39](#page-14-0)] presented that in a system where rare customers have long service times, a pooling can be made arbitrary worse than specialization. Benjaafar [[34\]](#page-14-0) claimed that machine sharing can be beneficial, the benefits are conditioned and limited by a number of system-operating parameters such as setup times, part mix variety, system utilization, and batch sizes. He added that for systems with small setups, the reduction in flow times and WIP can be significant. On the other hand, for systems with high setup times, machine sharing can lead to lower production rates and higher congestion levels. Similar conclusion has also been reported by Buzacott [[40\]](#page-14-0) who considered a variety of pooling scenarios. The author explained that when the service times of different customer classes are not identical in distribution, pooling can lead to longer queuing delays. He also added that sensitive to service time variability, the size of demand from each class, and the routing policy are the factors which cause difference in performance between different pooling scenarios.

2.1.2 Chaining as a flexibility strategy

It is well known that sharing resources is often performed with additional investment (e.g., flexible material handling, programmable machines, and flexible workforce) or higher operating costs (e.g., more frequent setups, increased material handling, increased control complexity), therefore increased resource sharing could result in a trade-off between performance and cost. Sheikzadeh et al. [[41](#page-14-0)] focused on a sharing strategy that combines the benefits of both pooling and specialization. This strategy is referred as machine chaining. Sheikzadeh et al. [[41\]](#page-14-0) presented that chain systems, with homogenous demand and service times, achieve most of the benefits of total pooling. The concept of a chain was introduced by Jordan and Graves [[35\]](#page-14-0). They employed chaining in

the context of process flexibility for a single-stage manufacturing system with random demand. Graves and Tomlin [\[42](#page-14-0)] extended the system of Jordan and Graves [\[35\]](#page-14-0) to multistage manufacturing systems. Another study employing chaining was performed by Hopp et al. [[43](#page-14-0)] who investigated serial production systems operating under a CONWIP protocol with queuing models of flexible workers. They revealed two-skill chaining structure (that allows every worker to directly or indirectly redirect two customers from the base station to any other station) possesses strong capacity balancing and variability buffering properties.

In a chained configuration, each operation (or customer) can be routed to one of two neighboring machines (or servers) and each machine can process operations from two neighboring classes. Chaining may capture the benefits of full flexibility without requiring every machine to process every party type. Because the number of part types that are processed on any resource is limited to two in a changing configuration. This characteristic is important because additional flexibility often comes at the cost of either additional investment or high operating overhead. Sheikzadeh et al. [\[41](#page-14-0)] indicated that by increasing the scope of manufacturing resource, the routing flexibility of products is also increased.

2.2 Queuing strategies for flexibility in job shop

Manufacturing systems can be represented by classes of queuing networks. Therefore, managing these networks and developing strategies for queues (i.e., machine queues) can be essential for manufacturing performance. One well-known strategy for queue management has been the use of dispatching rules. Dispatching rules determine which job will be selected from a machine queue when the machine becomes idle. The main role of a dispatching rule is to assign priority to the entities (such as operations) waiting in a queue. The entity having the highest priority is started to process when the machine becomes idle. However, when the idle machine queue is empty, nothing is performed and machine remains idle. Especially in the co-existence of several identical machines, an entity waiting in the queue is not shifted to idle machine. Therefore, using strict dispatching rules for controlling machine queues may be insufficient.

One of the ways for reducing the insufficiency may be to employ "jockeying" strategy in machine queues. In service systems, jockeying is defined as "the movements of customers who have the option of switching from one line to another when several servers, each having a separate and distinct waiting line, are available" [[44\]](#page-14-0).

Jockeying has been quite common in many queue structures of service systems. Most of the people used to jockey for years (i.e., switching lines in auto license offices and supermarkets) when faced with long queues in order to reduce total time spent in the system. Up to now, few researches performed detailed research on jockeying. Some researchers [[45](#page-14-0)–[47\]](#page-14-0) studied the jockeying problem by considering models with Markovian inputs. Some authors [[48](#page-14-0), [49\]](#page-14-0) compared the results for jockeying and nonjockeying models, and they revealed that significant improvement of the system performance has been achieved for the jockeying model. No matter which type of system is studied, jockeying can only occur when the customer/operation/job have more than one alternative queues. But to our best knowledge, researchers studying on the manufacturing systems assumed that jockeying among machine queues is not allowed. Even though, it may be employed in practical manufacturing circumstances. This assumption may be mainly due to the fact that jockeying can increase controlling complexity. In addition, allowing jockeying increases transportation time of jobs. In contrast to this undesired situation, allowing jockeying decreases waiting time of the entities on their queues extensively. Eventually, as desired to be proved in this paper, decrease in the waiting times on the queues can create much more positive effect on the system performance. In this study, an attempt is made to see the effect of jockeying in a job shop with different flexibility levels.

3 Research methodology

In this study, effects of flexibility on "Average Flow Time" (AFT) in a simulated job shop are investigated under various flexibility levels, shop utilization levels, and machine selection (routing) rules. For that purpose, some simulation models are developed and experimented under nine different scenarios which include jockeying. In this section, details of these factors and scenarios are presented.

3.1 Performance measure

AFT is considered as the main performance measure in this study. We choose the AFT as a performance indicator since the information about the shop status directly effects the flow time and the estimations made about it [[7,](#page-13-0) [8](#page-13-0)]. AFT (or equivalently mean flow time) is an average over the flow time for each of the measured flows. It is a widely used measure for determining the responsiveness of the system. The equation used to compute AFT is given as follows:

$$
AFT = \sum_{n=1}^{N} FT_n/N
$$
 (1)

where, FT_n is the flow time of job *n* and *N* is the number jobs completed. Flow time is the average time an order spends in the system. Equation 2 is used to estimate the flow time:

$$
FT_i = \sum\nolimits_{j \in S\{i\}} \left(s_{ij} + m_{ij} + w_{ij} + p_{ij} \right) \tag{2}
$$

where,

- j machine number
- s_{ij} set-up time for job i at machine j (set-up times are assumed to be negligible in this work)
- m_{ii} transportation time necessary for job *i* moving from machine i to next machine on its route
- w_{ij} queue waiting time for job *i* in machine *j*
- p_{ii} actual processing time of job *i* in machine *j*
- $S(i)$ the set of machines which are placed on job i's route

Some job shop factors, like machine utilization, queue waiting times are varying instantaneously and this variety cause flow time fluctuations. Beside this, increase in transportation time due to jockeying can be considered by employing the average flow time as a performance measure.

3.2 System considerations

- Flexibility levels. In this study, flexibility levels are arranged with respect to number of different operations performed by a machine. Four flexibility levels are considered and named as FL1, FL2, FL3, and full flexibility (FL4). Structures for these configurations are schematically presented in Fig. [2.](#page-6-0)
	- 1. FL1. Each machine can perform two different operations.
	- 2. FL2. Each machine can perform three different operations.
	- 3. FL3 (chained configuration). It is assumed that each operation type "i" can be processed by machine " i " and machine " $i-1$ ", as being the alternative machine for operation type " i ". For an operation type 1 (where $i=1$), machine 6 is the alternative machine (since $i-1=0$).
	- 4. FL4. All six machines perform all types of operations.

The flexibility presented in this study addresses the "routing flexibility" and "machine flexibility" as presented in Section [2.](#page-1-0) Specialization is the case where each operation have only one candidate machine (no alternative machine) and it is accepted as the base case for the comparison of AFT through the defined scenarios. In other words, in base case, each type of operation can be processed only on the machine having the same machine number. For example, operation type 1 can only be processed at machine 1, operation type 2 can only be processed at machine 2…etc.

- Shop utilization levels. Utilization is defined as the percentage of time a resource is busy servicing user requests. Two levels of shop utilization (70% which represent low shop utilization level and 90% which represent high shop utilization level) are configured considering the base case. These two shop utilization levels are configured by adjusting mean of interarrival times.
- Machine selection (routing) rules (MSR). Two different measures are computed in order to decide which machine is going to be selected. These measures are named as search 1 and search 2, respectively. One of these measures (both are experimented for each scenario) is calculated for each alternative machine and the machine giving the minimum value is selected.
- 1. Search 1=(number of jobs waited in machine queue \times average processing time of last 10 operations performed at this machine)+remaining processing time of operation working at this machine+transportation time between previous machine and this machine
- 2. Search 2= (number of potential operations waited in machine queue \times average processing time of last 10 operations performed at this machine)+ remaining processing time of operation working at this machine+transportation time between previous machine and this machine

Number of potential operations waited in the machine queue given in search 2, is calculated by using the following routine.

Job number =1 CONTROL1: IF Job number < = $NQ(M)$ Potential operation = potential operation +1 CONTROL2: IF operation order +1 <= total operation IF type (operation order + 1) ϵ M

Potential operation =potential

operation +1

Operation order = operation order +1

GO TO CONTROL2

ELSE

Job number = Job number +1

GO TO CONTROL1

ENDIF

ELSE

Job number = Job number +1

GO TO CONTROL1

END IF

END IF

Since the purpose is to find possible workload of a machine, total number of "potential operations" is summed up starting from the first job waiting in the queue to the last job in the queue. It is remarkable to state here that, "potential operation" refers the total number of operations of a job that can be performed at that machine without leaving the machine. Therefore, potential operations should be consecutive operations not to be routed to another machine. NQ(M) indicates the total number jobs waiting in

the queue of the Mth machine. Thus, starting from the first job in the queue (where "job number" is 1), to the last job in the queue (where "job number"= $Q(M)$) potential operations are added up.

For checking the potential operations of a job, "operation order" and "total operations" attributes are used. "Operation order" refers to the order of the current operation of a job waiting in the queue. If a job is already in the queue of a machine, it means its current operation is a potential

operation to be performed at that machine (where potential operation=potential operation+1). Subsequently, the next operation is checked if it exists or not. If it exists its type (type (operation+1)) is checked whether it is a potential operation or not. If it is potential, then potential operation is increased by 1 if not consecutive job is checked.

3.3 Experimental conditions and assumptions

Performance or the impact of new configurations and control algorithms on manufacturing systems is generally unknown in advance. Simulation models have proved to be useful for examining the performance of different system configurations and/or alternative operating procedures for complex logistic or manufacturing systems. The applicability of simulation in the general area of manufacturing systems analysis and design is well known since it can handle complex stochastic systems in arbitrary detail. Simulation is a highly flexible tool that can be used effectively for analyzing complex systems [\[50](#page-14-0)]. Different configurations can be tested by using various system parameters. On the other hand, the specification of the "parameter set" to run these simulations may not as easy as expected. Since it is difficult to reflect all variability which exists in a real system, some theoretical assumptions are defined prior to experiment. If there is an assumption which is entirely and exclusively utilized during a procedural aspect, then it is should be indicated before the models are

presented. There are also some assumptions for the models which are developed in this study. The assumptions considered for the simulation models which are presented in this paper are as follows:

- Labor and machine are assumed to be always available (i.e., without any shifts or machine breakdowns) and waiting time occurs only when a machine is busy.
- There are six different types of machines and there is only one unit of each machine type. Machines posses the capability of operating more than one type of operation. Types of operation performed by a machine depend on the flexibility level.
- Each machine maintains a separated queue for all flexibility levels. Each machine can handle at most one operation at a time.
- Jobs are strictly ordered sequences of operations without assembly. There are 10 different types of jobs. Probabilities for the arrival of job types to the system are equal (10%) .
- Job orders arrive to the system randomly with interarrival times that are exponentially distributed with a mean 2.5 and 3.25 h for two different shop utilization levels.
- All jobs are accepted, there is no job acceptance/ rejection stage. There is no "entrance gate". Jobs enter the first machines without any time delay.
- & A job leaves the system after completing its all operations from the same exit gate. Job leave the system with time delay solely depends on the distance between its last machine and the exit gate.
- There are six different types of operations. Any operation type can exist on a job. Number of operations for completing a job varies between four and eight. The job routings in terms of operations are fixed. However, the machines to process these operations are not unique [[51](#page-14-0)].
- Table 1 presents the corresponding operations of job types and routes.

 $-1 - 5 - 4 - 6 - 1 - 3 - 4$

14 1–2–1–5 24 2–3–6–2

- Operation preemption is not allowed.
- Setup time is ignored. However, processing time of an operation increases if the operation is performed in the alternative machine(s).Processing time of each operation is randomly generated from uniform distribution. The parameters of uniform distribution (minimum and maximum value) of each processing time are given in Table 2.
- Transportation time between machines is deterministic and only depends on the distance between machines (8 min between two neighborhood machines). The position of the machines in the job shop is as presented in Fig. [2.](#page-6-0)
- Machine selection rule and dispatching rule are the two components of the control policy. Control policies are applied at each decision point (machine selection and job selection). Arrival of a job to the shop or completion of an operation by a machine is the two events that trigger the decision point.
- WIP buffers have unlimited capacities.
- & Inserted idle time is not allowed. A machine will not remain idle for waiting for a particular operation of a routed job (a job delivered to a machine but not arrived yet) to arrive. If there are other operations already available and waiting in a machine's queue, it is processed without considering the routed job. This assumption is released in scenarios 7, 8, and 9.

Scenarios are developed for of the each cases and each level of the factors under three dispatching rules. These are;

- 1. First-in-system-first served (FSFS). When the machine becomes idle, the job in queue with the earliest shop arrival date is selected first.
- 2. First come first served (FCFS). When the machine becomes idle, the operation of a job in queue with the earliest queue arrival date is selected first. The term operation is used instead of job because when an operation of a job is performed at the same machine, it does not change the current machine. The early coming job does not lose its priority and consecutive operations of that job are performed at the same machine.

3. Minimum average processing time (MAPT). When the machine becomes idle, the job (operation) in queue with the minimum average processing is selected first. It is equivalent of shortest processing time rule.

For the second and the third dispatching rules, FSFS is used as a tie-breaker.

The statistical analysis of the dynamic job shop requires that the job shop be in the steady state [[51\]](#page-14-0). It is known that a brief warm-up period is required for the simulations to reach a steady state. In this paper, models are terminated after 18,000 jobs are completed. Therefore; the duration required for completing 1,800 jobs (10% of all jobs) are not included in the statistics of the results since it has been treated as warm-up periods for each model. All models are run using 25 replications in order to minimize variability in results. In addition, common random numbers are used to provide the same experimental condition across the runs for each scenario.

3.4 Scenarios

The selection of the machine for the first operation of each job is the same for all scenarios whereas MSR changes. The selection process is given as follows: when a job enters the system, it searches among the alternative machines of the first operation by using one of MSR. Then a machine is selected.

- & IF the selected machine is idle, then the operation is processed as soon as it enters the system.
- IF the selected machine is not idle (i.e., there is an operation processing at this machine), then the operation enters the queue of this machine.

Conditions for the scenarios are given in Table [3](#page-9-0). Details of these scenarios are defined as follows:

Scenarios 1, 2, and 3 When an operation of a job is finished, first of all it is checked whether it is the last operation of this job or not.

- 1. IF it is the last operation, then this job leaves the system.
- 2. IF it is not the last operation, then it is checked whether it can be processed at the current machine or not.
	- (a) IF it can be processed at the current machine, then it is controlled whether there is any job that is waiting for this machine having higher priority, with respect to the employed dispatching rule, than the current one. (For scenario 3, consecutive operation/operations of the same job is/are allowed to be processed at the same machine

without interruption. In other words, there is no need to search the current machine's queue.)

- (b) IF there is any job whose priority is higher than the current job, then the job having the highest priority is processed. Current job, for processing the next operation, searches among the alternative machines by using one of MSR. It selects either a different machine or the current machine.
	- (b.1) IF current machine is selected again, then next operation of this job passes the same machine's queue.
	- (b.2) IF different machine is selected, then this job is routed to the selected machine in order to process the next operation.
- (c) IF there is not any job having higher priority than the current operation, next operation of the current job is processed at the same machine
- 3. IF it cannot be processed, then current job searches among the alternative machines by using one of MSR. It is routed to the selected machine in order to process the next operation.
	- (a) IF there is any job waiting for current machine, then the job having the highest priority is processed at current machine.
	- (b) IF there is not any job waiting for this machine, nothing is performed.

Whenever an operation ends all machines queues are searched sequentially.

- 1. IF the searched machine is idle and its queue is not empty (for an instant where the time does not progress), then the job having the highest priority is selected for processing on this machine.
- 2. IF the searched machine is not idle or the searched machine is idle but there is not a job waiting on its queue, nothing is performed.

This search process continues until the last machine is controlled.

Scenarios 4, 5, and 6 Scenario 4 is similar to scenario 1. scenario 5 is similar to scenario 2, and scenario 6 is similar to scenario 3. Differences between these scenarios are allowance of jockeying between alternative machine's queues. When an operation of a job is finished, the same procedure which is employed in scenarios 1, 2, and 3 is repeated. Whenever an operation ends, queue of all machines are searched sequentially.

1. IF the searched machine is idle and its queue is not empty, then the job having the highest priority is selected for processing on this machine.

Table 3 Conditions for scenarios

- 2. IF the searched machine is not idle, nothing is performed.
- 3. IF the searched machine is idle and its queue is empty, then the nearest machine queue which the relevant machine can perform the same operation is sought. Seeking process continues until a machine queue having any job for processing is found or all the alternatives are dead.
	- (a) IF any job is found among the alternatives, then the job having the highest priority is routed to the idle machine.
	- (b) IF there is more than one machine having the same distance to the idle machine, the job having the highest priority among these machines is routed to the idle machine.
	- (c) IF alternatives are dead, nothing is performed.

This search process continues until last machine is controlled.

Scenarios 7, 8, 9 Scenario 7 is similar to scenario 4, scenario 8 is similar to scenario 5, and scenario 9 is similar to scenario 6. Differences between these scenarios are as follows: in scenarios 4, 5, and 6, jockeying is allowed when the selected machine is idle and its queue is empty; however, in scenarios 7, 8, and 9, in addition to these conditions, there should be no routed job to the idle machine.

All these scenarios are programmed under four flexibility levels, two shop utilization levels, and two machine selection rules. In other words, each of these scenarios is programmed for 16 $(4 \times 2 \times 2)$ different factors. Besides, scenarios 1 and 2 are programmed for specialized job shop (base case) conditions (since there is no alternative machines, MSR is not employed) under two shop utilization levels. Scenario 3 is not programmed for the base case because for this case, each machine is allowed to perform only one type of operation with the same expected (average) duration of processing time and FSFS is employed as the tie breaker. Therefore, results do not change. In conclusion totally 148 ($(16\times9)+4$) programs are run.

4 Results and discussion

In this section, the results of the experiments are discussed for all nine scenarios. The AFT values for the first three scenarios (without jockeying) are shown in Table [4](#page-10-0).

As it can be seen from Table [4,](#page-10-0) FL4 always gives the minimum AFT in 70% shop utilization level. In addition, in each scenario, search 2 gives better result than search 1. FL2 is also the second best in each case and search condition.

For 90% shop utilization, except scenario 1 with FCFS, FL4 always gives the minimum AFT. On the other hand, dissimilar to 70% shop utilization, FL3 gives the second best result. In addition, it can be easily seen from Table [4,](#page-10-0) for the first two scenarios, AFT value does not change for FL4 for both of the shop utilization levels. In scenarios 1 and 2, jobs do not leave their machines until they have completed all of their operations. According to the defined dispatching rules, machines can only capture the operations with minimum system arrival time (FSFS, scenario 1) and with minimum machine arrival time (FCFS, scenario 1) and thereby when an operation is being performed in a machine, another job having an earlier arrival time cannot logically be in the queue. However in the other levels of flexibility (FL1–FL2–FL3), all operations of a job may not be completed in the same machine and there may be some machine switches. Thereby, the order of operations with respect to machine arrival times is not the same with the order of the jobs with respect to the system arrival times. Consequently, an early comer to the system looses it advantage and may have longer average flow times when compared to scenario 1.

Table 4 The average flow times (hour) for scenarios 1, 2, and 3

Simulation results for the next three scenarios (with jockeying) are shown in Table 5. Remember that scenarios 4, 5, and 6 are the same with scenarios 1, 2, and 3, respectively. However, scenarios 4, 5, and 6 allow for jockeying. Therefore, all commends for the first three scenarios are valid for scenarios 4, 5, and 6.

As it can be seen from Table 5, allowance of jockeying has created some improvements for AFT. AFT values decreases or remained same for all of the states. Amount of the improvement becomes apparent for FL4. Since all operations can be performed by all of the machines, a job waiting at queue can be routed to another machine with a higher probability for both shop utilization level. Another improvement with jockeying is the reduction of differences of AFT values obtained for search 1 and search 2 rules. In scenario 4, the differences of AFT values for searches 1 and 2 gets smaller relative to scenario 1 and dissimilar to scenario 1, FL4 gives the minimum value with the search 1. This state can be explained with the precision of search rules. Search 1 rule, is structured on the jobs waiting at the queue and search 2 is based on the total operations waiting at the queue. Inherently, search 1 rule is less precise for determining the job loads of machines. Therefore, in scenarios 1, 2, and 3, differences of AFT values are relatively large, especially for 90% shop utilization. However, scenarios 4, 5, and 6 allow jockeying and the jobs can switch their queues and as a consequence, the precision handicap for search 1 is released.

AFT values for the next three scenarios (with jockeying) are given in Table [6.](#page-11-0) As mentioned before, scenarios 7, 8, and 9 also includes jockeying; however, conditions required for jockeying is a bit different. In these scenarios, a job does not leave its queue if there is an already routed job to the candidate idle machine. Therefore, this system is proactive about the state of the idle machine. Therefore, results are consistently better or equal to the results given in scenarios 4, 5, and 6.

As it can be seen from Table [6](#page-11-0), AFT values are better but they do not create significant improvements. The graphical representation of the AFT values of scenarios for 70% and 90% shop utilization are given in Figs. [3](#page-11-0) and [4,](#page-11-0) respectively. In these figures, it can be easily seen that the range of AFT values for 70% shop utilization is considerable smaller than the 90% shop utilization. Therefore, it can

Table 6 The average flow times (hour) for scenarios 7, 8, and 9

be safely said that flexibility management is more critical under highly utilized job shops.

Analysis of variance (ANOVA) provides a significance test for the main effect of each variable in the design. Using simulation results, ANOVA–F test has been carried out for each factor to determine whether the means are significantly different from each other. The analysis of variance for the AFT is performed to verify the effective factors, and the results are summarized in Table [7.](#page-12-0) In all cases, since the P value of the F test is less than 0.01, there is a statistically significant difference between the flow times of different configurations of the system at 99% confidence level. Results indicate that all of the considered factors (job utilization, search rules, and flexibility level jockeying) have significant effect on the flow time performance.

The main effect of an independent variable is the effect of the variable averaging over all levels of other variables in the experiment. Main effect analysis shows the general tendency of variables to change at interested dependent

Fig. 3 AFT values with respect to scenarios for 70% shop utilization rate

variable. For this purpose, "main effect plot" has been generated by using Minitab software (see Fig. [5\)](#page-12-0).

Based on the obtained results, as also depicted in the previous tables and figures, some general observation and statements can be made as follows:

Observation 1. In 70%, shop utilization level system is not as crowded as 90% shop utilization level and therefore system is relatively more relaxed and queue lengths are naturally small. As a consequence, effects of improvements obtained with flexibility are relatively small when compared with 90% shop utilization.

Statement 1. Effect of flexibility increases as the shop utilization level increases.

Observation 2. It should be noticed that, in both FL1 and FL3 levels, each machine is capable of performing two different operations; however, FL3 is capable of doing much more then FL1 does for 70% shop utilization level. Moreover, FL3 sometimes (in two of

Fig. 4 AFT values with respect to scenarios for 90% shop utilization rate

Table 7 Results of ANOVA

Factor	Degrees of freedom	F value	P value
Job utilization		642.63	
Search rules	2	7.62	0.001
Flexibility rate	4	7.63	θ
Jockeying		4.63	0.003

over 80 rankings) performed as well as FL2 performs. In 90% shop utilization level, it performs much better and it sometimes even competes with full flexibility. The potential of FL3 (chain flexibility) is much more than expected since it can shift the excess workload to idle machines.

Statement 2. Effect of chain configuration is much more than any other two machine flexibility.

Observation 3. Experimentations indicate that, in full flexibility (without jockeying conditions), AFT value does not change for FL4 for FCFS and FSFS for both of the shop utilization levels. As mentioned before, this case is the result of the equivalence of arrival time and machine arrival times. However, in the other levels of flexibility FSFS performs better than FCFS.

Statement 3. FSFS have the highest positive effect for both shop utilization levels.

Observation 4. Experimentations indicate that AFT values for base case of FSFS and MAPT does not change since each machine is allowed to perform only one type of operation with the same expected (average) duration of processing time. As indicated in the system description section, if conditions are equal, FSFS is employed as the tie breaker. Therefore, results do not change.

Statement 4. If the system is not flexible, it does not matter whether MAPT or FSFS is employed as a dispatching rule.

Observation 5. Experimentations show that the scenarios where jockeying is allowed, have better or at worst the same AFT values. In these scenarios, jockeying reduced differences of AFT values among the searches 1 and 2 rules.

Statement 5. A queue control mechanism like jockeying may improve the performance of a system.

Observation 6. Observations 4 and 5 presented that the proper use of a control mechanism may convert potential flexibility into better performance. However, there is more to imply. Improvements obtained with the indicated control mechanisms are more significant for 90% shop utilization level. This means that crowded systems have much more opportunity to utilize flexibility for improving performance.

Statement 6. As the shop utilization level increases, importance (the benefits obtained) of control mechanisms (such as dispatching rules, queue management policies, etc.) increase.

5 Conclusions and future work

It is well known that several attempts were made in the literature to understand AFT performance behavior of job shops. AFT is a basic measure of a shop's performance at turning around orders, and it is therefore often used as an

Fig. 5 Main effects plots

indicator of success in responding quickly to customers [\[52](#page-14-0)]. Authors frequently mentioned and experimented on minimization of the AFT by making proper adjustments on the components flow time. As mentioned before, flow time is formulated as the sum of setup times, transportation time, waiting times in the machine queues, and the processing times. Therefore; efforts presented in the literature are mostly based on these main components. If any improvement can be obtained on one of these components, AFT values is expected to be positively affected from these changes. Waiting times in the machine queues, is one of the most important component building up AFT. Therefore, efforts mostly focused on decreasing the waiting times. Certainly, there has been a significant body of knowledge that has produced guidelines on how to decrease queue lengths such as alternate routings for jobs and shifts of operations between the machines. Flexibility, thereby resource/machine sharing, has been key elements to decrease queue lengths and waiting times at queues. Previous studies focused on two scenarios which are pooling and specialization. In pooling, machines are set into functional pools, where each pool is associated with a set of operations, whereas in specialization each machine is dedicated to a single operation type. However, in the literature, some authors mentioned that pooling may not be the best strategy. For example, Buzacott [\[40](#page-14-0)] studied a variety of pooling scenarios and reveled that when different customer classes have different service times, pooling increases the queue waiting times. Gurumurthi and Benjaar [\[53](#page-14-0)] claimed that higher flexibility does not always improve throughput. Majority of these studies was performed in service systems under no or limited flexibility and they all assumed that mechanism like jockeying is not allowed. Different from the previous studies, in this work we have observed that if effective control mechanisms like jockeying are utilized increased flexibility is a preferable state for improving flow time performance especially under highly utilized job shops.

In this study, a limited number of scenarios and configurations have been considered. However, as a future study, these scenarios can be extended with the addition of new control mechanisms. These control mechanisms may include different dispatching rules, conditional search mechanisms, and different versions of queue jockeying (such as controlling queue in predetermined intervals and jockeying according to work load balances). We should point here that much more studies are needed in order to fully understand the effects of jockeying in manufacturing systems. It is clear that jockeying can increase complexity. However, in order to utilize system's flexibility for improved performance, the level of complexity should be regulated effectively. Much more work is needed here.

In addition, different chaining configurations can also be tried and compared with full flexibility especially for the high shop utilization levels. Moreover, transportation times are assumed relatively shorter in this study, as a future work, new trials can be performed with longer transportation times. Furthermore, several other performance measures especially due date-related criteria can also be analyzed.

Acknowledgments The first author would like to thank to "Turkish Academy of Sciences" for supporting his academic studies.

References

- 1. Baines TS, Harrison DK, Kay JM, Hamblin DJ (1998) A consideration of modelling techniques that can be used to evaluate manufacturing strategies. Int J Adv Manuf Technol 14(5):369–375
- 2. Correa HL (1994) Linking flexibility, uncertainty and variability in manufacturing systems: managing un-planned change in the automotive industry. Avebury, Aldershot
- 3. Gustavsson S (1984) Flexibility and productivity in complex production processes. Int J Prod Res 22(5):801–808
- 4. Wadhwa S, Rao K, Chan F (2006) Comparative influence of three flexibility types on manufacturing lead-time performance. Int J Adv Manuf Technol 29(9):1002–1007
- 5. Chan FTS, Bhagwat R, Wadhwa S (2006) Increase in flexibility: productive or counterproductive? A study on the physical and operating characteristics of a flexible manufacturing system. Int J Prod Res 44(7):1431–1445
- 6. Baykasoğlu A (2009) Quantifying machine flexibility. Int J Prod Res 47(15):4109–4123
- 7. Eilon S, Chowdhury IG (1976) Due dates in job shop scheduling. Int J Prod Res 14:223–237
- 8. Weeks JK (1979) A simulation study of predictable due-dates. Man Science 25(4):363–373
- 9. Jan Eppink D (1978) Planning for strategic flexibility. Long Range Plann 11(4):9–15
- 10. Tsubone H, Horikawa M (1999) A comparison between machine flexibility and routing flexibility. Int J Flex Manuf Sys 11(1):83–101
- 11. Buzacott JA (1982) The fundamental principles of flexibility in manufacturing systems. In: The fundamental principles of flexibility in manufacturing systems. New York: Proceedings Book, pp 13–22
- 12. Browne J, Dubois D, Rathmill K, Sethi SP, Stecke KE (1984) Classification of flexible manufacturing systems. The FMS Magazine, April, pp 114–117
- 13. Slack N (2005) The flexibility of manufacturing systems. Int J Oper Prod Man 25(12):1190–1200
- 14. Sethi AK, Sethi SP (1990) Flexibility in manufacturing: a survey. Int J Flex Manuf Sys 2(4):289–328
- 15. Chen I, Calantone R, Chung C (1992) The marketing–manufacturing interface and manufacturing flexibility. Omega 20(4):431– 443
- 16. Benjaafar S, Ramakrishnan R (1996) Modelling, measurement and evaluation of sequencing flexibility in manufacturing systems. Int J Prod Res 34(5):1195
- 17. Taymaz E (1989) Types of flexibility in a single-machine production system. Int J Prod Res 27(11):1891
- 18. Das SK (1991) Design and measurement of flexibility in manufacturing systems. Working paper, (No. 91–06). New Jersey Institute of Technology
- 19. Gerwin D, Kolodny H (1992) Management of advanced manufacturing technology: strategy, organization, and innovation. Wiley, New York
- 20. Lau RSM (1996) Strategic flexibility: a new reality for worldclass manufacturing. SAM Adv Man J (07497075) 61(2):11
- 21. Toni AD, Tonchia S (1998) Manufacturing flexibility: a literature review. Int J Prod Res 36:1587–1617
- 22. Oke A (2005) A framework for analysing manufacturing flexibility. Int J Oper Prod Man 25(10):973–996
- 23. Gupta D (1993) On measurement and valuation of manufacturing flexibility. Int J Prod Res 31(12):2947–2958
- 24. Brill PH, Mandelbaum M (1990) Measurement of adaptivity and flexibility in production systems. Eur J Oper Res 49(3):25–332
- 25. Das SK (1996) The measurement of flexibility in manufacturing systems. Int J Flex Manuf Sys 8:67–93
- 26. Georgoulias K, Papakostas N, Chryssolouris G, Stanev S, Krappe H, Ovtcharova J (2009) Evaluation of flexibility for the effective change management of manufacturing organizations. Robot CIM-Int Manuf 25(6):888–893
- 27. Narain R, Yadav R, Sarkis J, Cordeiro JJ (2000) The strategic implications of flexibility in manufacturing systems. Int J Agile Man Sys 2(3):202–213
- 28. Wadhwa S, Rao K (2003) Flexibility and agility for enterprise synchronization: knowledge and innovation management towards flexagility. Stud Inform Control 12:111–128
- 29. Sharma O, Sushil P (2002) Issues in managing manufacturing flexibility: a review. Global J Flex Sys Man 3:11–29
- 30. Gerwin D (2005) An agenda for research on the flexibility of manufacturing processes. Int J Oper Prod Man 25(12):1171–1182
- 31. Olhager J (1993) Manufacturing flexibility and profitability. Int J Prod Econ 30–31:67–78
- 32. Bengtsson J (2001) Manufacturing flexibility and real options: a review. Int J Prod Econ 74:213–224
- 33. Das SK, Nagendra P (1993) Investigations into the impact of flexibility on manufacturing performance. Int J Prod Res 31 (10):2337–2354
- 34. Benjaafar S (1995) Performance bounds for the effectiveness of pooling in multi-processing systems. Eur J Oper Res 87(2):375– 388
- 35. Jordan WC, Graves SC (1995) Principles on the benefits of manufacturing process flexibility. Manage Sci 41(4):577–594
- 36. Stecke KE, Kim I (1989) Performance evaluation for systems of pooled machines of unequal sizes: unbalancing versus balancing.

Available from: [http://deepblue.lib.umich.edu/handle/2027.42/](http://deepblue.lib.umich.edu/handle/2027.42/36155) [36155](http://deepblue.lib.umich.edu/handle/2027.42/36155). Accessed on: 11 March 2010

- 37. Dallery Y, Stecke KE (1990) On the optimal allocation of servers and workloads in closed queueing networks. Oper Res 38(4):694– 703
- 38. Calabrese JM (1992) Optimal workload allocation in open networks of multiserver queues. Manage Sci 38(12):1792–1802
- 39. Smith D, Whitt W (1981) Resources sharing for efficiency in traffic systems. Bell Syst Tech J 60:39–55
- 40. Buzacott JA (1996) Commonalities in reengineered business processes: models and issues. Manage Sci 42(5):768–782
- 41. Sheikhzadeh M, Benjaafar S, Gupta D (1998) Machine sharing in manufacturing systems: flexibility versus chaining. Int J Flex Manuf Sys 10:351–378
- 42. Graves SC, Tomlin BT (2003) Process flexibility in supply chains. Manage Sci 49(7):907–919
- 43. Hopp WJ, Tekin E, Van Oyen MP (2004) Benefits of skill chaining in serial production lines with cross-trained workers. Manage Sci 50(1):83–98
- 44. Koenigsberg E (1966) On jockeying in queues. Manage Sci 12:412–436
- 45. Disney RL, Mitchell WE (1970) A solution for queues with instantaneous jockeying and other customer selection rules. Nav Res Log 17(3):315–325
- 46. Kao E, Lin C (1990) A matrix-geometric solution of the jockeying problem. Eur J of Oper Res 44(1):67–74
- 47. Zhao Y, Grassmann WK (1990) Shortest queue model with jockeying. Nav Res Log 37:773–787
- 48. Zhao Y, Grassmann WK (1995) Queueing analysis of a jockeying model. Oper Res 43(3):520–529
- 49. Tarabia A (2008) Analysis of two queues in parallel with jockeying and restricted capacities. Appl Math Model 32 (5):802–810
- 50. Chan FTS, Chan HK (2004) A comprehensive survey and future trend of simulation study on FMS scheduling. J of Intell Manuf 15 $(1):87-102$
- 51. Subramaniam V, Lee GK, Ramesh T, Hong GS, Wong YS (2000) Machine selection rules in a dynamic job shop. Int J Adv Manuf Technol 16(12):902–908
- 52. Baker KR (1984) The effects of input control in a simple scheduling model. J of Oper Manag 4(2):99–112
- 53. Gurumurthi S, Benjaafar S (2004) Modeling and analysis of flexible queueing systems. Nav Res Log 51(5):755–782