Received: 17 May 2023 | Revised: 28 August 2023 | Accepted: 30 August 2023

IET Collaborative Intelligent Manufacturing

DOI: 10.1049/cim2.12084

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Shop floor dispatching with variable urgent operations based on Workload Control: An assessment by simulation

Mingze Yuan[1,2,3] | Lin Ma[4] | Ting Qu[2,3,5] | Matthias Thürer[6]

[1]School of Management, Jinan University, Guangzhou, China

[2]Guangdong International Cooperation Base of Science and Technology for GBA Smart Logistics, Jinan University, Zhuhai, P.R. China

[3]Institute of Physical Internet, Jinan University, Zhuhai, China

[4]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China

[5]School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, China

[6]Chair of Factory Planning and Intralogistic, Chemnitz University of Technology, Chemnitz, Germany

Correspondence

Ting Qu, Guangdong International Cooperation Base of Science and Technology for GBA Smart Logistics; School of Intelligent Systems Science and Engineering; Institute of Physical Internet, Jinan University, Zhuhai 519070, P.R. China.
Email: quting@jnu.edu.cn

Funding information

2018 Guangzhou Leading Innovation Team Program, Grant/Award Number: 201909010006; National Key Research and Development Program of China, Grant/Award Number: 2021YFB3301701; 2019 Guangdong Special Support Talent Program – Innovation and Entrepreneurship Leading Team, Grant/Award Number: 2019BT02S593; Science and Technology Development Fund, Grant/Award Number: 0078/2021/A; National Natural Science Foundation of China, Grant/Award Number: 52375498

Abstract

Meeting customer time requirements poses a major challenge in the context of high-variety make-to-order companies. Companies need to reduce the lead time and process urgent jobs in time, while realising high delivery reliability. The key decision stages within Workload Control (WLC) are order release and shop floor dispatching. To the best of our knowledge, recent research has mainly focused on order release stage and inadvertently ignored shop floor dispatching stage. Meanwhile, urgency of job is not only related to its due date, but also affected by the dynamics of shop floor. Specifically, urgency of jobs may decrease at downstream operations in the job's routing, since priority dispatching for urgent jobs accelerates production speed at the upstream operations. And occupying production resources increases the waiting time of non-urgent jobs at workstation. This phenomenon leads to the change of urgency of jobs. Misjudgement of urgent jobs therefore may result in actual urgent jobs not being processed in time. In response, the authors focus on shop floor dispatching stage and consider the transient status of urgent operations in the context of WLC. The urgency of jobs is rejudged at the input buffer of each workstation, which is firstly defined as urgent operations and non-urgent operations. Using simulation, the results show that considering the transient status of urgent operations contributes to speeding up production for actual urgent jobs and meeting delivery performance both in General Flow Shop and Pure Job Shop. In addition, percentage tardy performance is greatly affected by norm levels, especially at the severe urgent level. These have important implications on how urgent operations should be designed and how norm level should be set at shop floor dispatching stage.

KEYWORDS

job shop scheduling, manufacturing industries, manufacturing systems

1 | INTRODUCTION

A major challenge for many small and medium-sized make-to-order companies is to meet the customer demand for product delivery [1, 2]. Dynamics from external factors (e.g. arrival time variability) and internal factors (e.g. the production process on the shop floor) lead to urgent jobs, which further leads to frequent job delays. Generally speaking, if the processing of urgent jobs is not timely, it will lead to deviations in the planned delivery date of the operation and delay in product delivery, which will have a fatal impact on the competitiveness of the companies [3, 4]. Therefore, it has become increasingly crucial for managers to produce urgent jobs efficiently and meet delivery times while minimising any adverse impact on the production process [5–7].

Workload Control (WLC) is a production planning and control concept designed to meet the needs of small and medium sized make-to-order companies [8, 9]. This concept is developed to overcome the 'lead time syndrome' in a high-variety MTO context [10]. WLC consists of customer enquiry management, order release control, and shop floor dispatching to show the process of accepting jobs to production jobs, where order release and shop floor dispatching are the key production planning and control stages to meet delivery performance in production [9, 11]. To be specific, (i) order release, where release decisions and selection decisions are taken in the pre-shop pool (e.g. ref. [12]). Jobs are held back in a pre-shop pool and input to the shop floor is regulated in accordance with workload norms, which buffers the shop floor against variance in the incoming order stream and improves throughput performance [13–15]. (ii) Shop floor dispatching, where sequencing decisions are taken at the input buffer of workstation after job is released to the shop floor (e.g. ref. [16]).

Improving the priority of urgent jobs is beneficial to speed up the process and reduce the percentage tardy, and further improve delivery performance [17]. Existing WLC literature mainly focused on order release stage for urgent jobs [15, 17]. Specifically, the assessment of urgency of jobs is executed in the pre-shop pool. That is the classification of jobs (i.e. urgent jobs and non-urgent jobs) is just determined at order release stage and then improves the release priority. However, existing literature ignored the impact of shop floor dispatching on the production of urgent jobs, which does not maximise the production performance of urgent jobs. Therefore, this study focuses on shop floor dispatching stage and addresses this research gap.

While speeding up the production of urgent jobs by improving their priorities, this may cause a shift in urgency of jobs in real life shops. To be specific, urgent jobs have high priorities, which can be prioritised production and reduce waiting time at the input buffer of workstation [18]. Therefore, delay time may be resolved at the upstream operations in the routing and the urgency of urgent jobs may be continuously reduced. Arrival time is earlier than planned start time for urgent jobs at downstream operations. Therefore, urgent jobs may be converted to non-urgent jobs. On the other hand, allocating production resources to urgent jobs may result in multiple operations backlogs for non-urgent jobs at the upstream operations, which can cause non-urgent jobs to arrive later than planned start time at downstream operations in the job's routing. Consequently, urgency of non-urgent jobs may increase over time, leading them to become urgent jobs. Therefore, it is crucial to rejudge the urgent operations of jobs at each workstation on the shop floor dispatching, according to the transient change of urgency of operations. This contributes to speeding up production for actual urgent jobs and avoids delays. However, WLC studies have not considered the status of jobs' operations in production. Transient change of urgency of operations is a crucial factor in production of jobs, especially in small and medium-sized make-to-order manufacturing companies with complex routings.

In response, considering transient change characteristic of urgency of operations, this study proposes a new paradigm for assessing urgency status of jobs. That is judge the transient urgent status of operations of jobs from the perspective of processing routing. Specifically, the urgency of jobs is rejudged at the input buffer of each workstation, which is defined as urgent operations and non-urgent operations. This study proposes an integrated time-oriented and load-oriented dispatching rule for urgent operations and non-urgent operations. Load-oriented element for urgent operations speeds up the process when multiple jobs become urgent; time-oriented element for non-urgent operations ensures production in time, thereby reducing tardiness. This was proved by Land [19]; Thürer et al. [20]. Using simulation experiment, the performance of variable urgency of operations method in production is assessed. In addition, this study compares the performance differences of the methods for different urgent levels.

The remainder of the paper is organised as follows. Section 2 presents a review of the literature on the method of processing urgent jobs. Section 3 describes the problem and defines the concept of this study, and describes the characteristics of the simulation model before the results of the study are presented and analysed in Section 4. Finally, conclusions are provided in Section 5 together with the limitations and future research directions.

2 | LITERATURE REVIEW

The key control stages within WLC are customer enquiry management, order release, and shop floor dispatching. For the customer enquiry management stage, the reader is referred to Hendry and Kingsman [21]; Thürer et al. [20]. In order to keep the study focused, this section reviews the order release mechanism and shop floor dispatching rules, as described in Section 2.1 and Section 2.2 respectively. A final discussion of the literature is presented in Section 2.3.

2.1 | Order release based on WLC

Once a job has been confirmed by the customer enquiry management, it is considered for releasing to the shop floor. The main objective of order release mechanism is to control the workload on the shop floor to reduce the throughput times [22].

There are many order release methods based on WLC in the literature, for examples, see reviews by Wisner [23]; Fredendall et al. [24]; Bagni et al. [25]. Most recent literature on WLC proved that Lancaster University Management School Corrected Order Release (LUMS COR) method was identified as the best order release solution for WLC in practice [14]. This order release method is designed to achieve the same levelling of workload to capacity achieved in make-to-order companies and further reduce both the WIP buffer and throughput time, which is adopted in this study.

LUMS COR combines periodic time intervals and continuous workload trigger to release jobs, as described in Figure 1. For periodic release procedure, jobs in the pool are considered for release periodically according to planned release dates, which allows the workload to be balanced. The selection decision compares the corrected aggregate workload of each workstation against predetermined workload norms. A job is released if the new workload at each workstation in the job's routing is below its workload norm; otherwise, the job is retained in the pre-shop pool. The full periodic release procedure can be formulated as follows and symbols are given in Table 1:

(i) A priority value is determined for each job in the set of jobs J in the pre-shop pool.
(ii) The job $j \in J$ with the highest priority is considered for release first.
(iii) If job j's processing time (p_{ij}) at the ith operation in its routing together with the current corrected workload (W_s) at workstation s corresponding to operation i fits within the workload norm (N_s) at this workstation, that is:

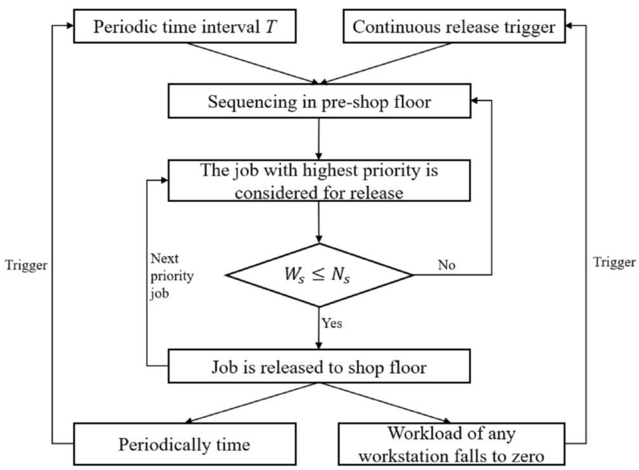

FIGURE 1 The release procedure of Lancaster University Management School Corrected Order Release (LUMS COR).

TABLE 1 Symbols used in Lancaster University Management School Corrected Order Release (LUMS COR).

Symbols	
J	The set of jobs ($J = 1, \ldots, j$)
S	The set of workstations ($S = 1, \ldots, s$)
i	ith operation in job's routing
R_j	Job j's routing
p_{ij}	Job j's processing time (p_{ij}) at the ith operation in its routing
W_s	Workload contribution of workstation s
N_s	Workload norm of workstation s

$$\frac{p_{ij}}{i} + W_s \le N_s, \forall i \in R_j \tag{1}$$

With R_j being the ordered set of operations in the routing of job j, then the job is selected for release, that is removed from J, and its load contribution is included, that is:

$$W_s := W_s + \frac{p_{ij}}{i} \tag{2}$$

Otherwise, the job remains in the pool and its processing time does not contribute to workload of workstation.

(iv) If the set of jobs J in the pre-shop pool contains any jobs that have not yet been considered for release, then return to step 2 and consider the job with the next highest priority. Otherwise, the release procedure is complete and the selected jobs are released to the shop floor.

For the continuous release procedure of LUMS COR, if the workload of any workstation falls to zero, the first job in the pre-shop pool sequence with that workstation as the first in its routing is released from the pre-shop pool irrespective of whether its release would exceed the workload norms of any workstation in its routing. This avoids premature idleness.

To the best of our knowledge, there is only one study rule-based order release WLC that considers urgent jobs and non-urgent jobs at the stage of order release. Thürer et al. [17] classified jobs according to the planned released date in the pre-shop pool: urgent jobs and non-urgent jobs. Urgent jobs are sequenced according to load-oriented sequencing to speed up the process, while non-urgent jobs are sequenced according to time-oriented sequencing. The mix of released jobs can be produced in time, thereby reducing tardiness.

2.2 | Shop floor dispatching based on WLC

To the best of our knowledge, there are many dispatching rules in the WLC literature, which can be divided into two sets: (i) time-oriented dispatching rules; and (ii) load-oriented dispatching rules. The time-oriented sequencing rules are examined:

- First Come First Served (FCFS), which sequences jobs according to their time of arrival in the workstation. This rule was applied, for example, Sabuncuoglu & Karapınar. [26] and Fredendall et al. [24].
- Operation Due Date (ODD), which sequences jobs according to job's due date of workstation. The operation due date of an operation is determined by successively subtracting an allowance for the operation throughput time at each workstation in the routing of a job from the job's due date. This rule was applied, for example, by Mezzogori et al. [27]; Raghu and Rajendran [28].
- Planned Start Time (PST), which sequences jobs according to job's start time of operation at the present workstation. This rule was applied, for example, by Thürer et al. [14]; Yan

et al. [29]. The planned start time δ_{js} of job j at workstation s as given by Equation (1) or Equation (2) in the existing studies:

$$\delta_{js} = d_j - \sum_{i \in R_j}^{n_j} (a_i + p_{ij}) \tag{3}$$

$$\delta_{js} = d_j - \sum_{i \in R_j}^{n_j} b_i \tag{4}$$

where symbols used in Equation (3) and Equation (4) are given in Table 2. And Equation (3) defines δ_{js} as the difference between the due date of job j and the total processing time and waiting time of all unprocessed workstations (e.g. ref. [24]). Equation (4) expresses δ_{js} as the difference between the due date of job j and the estimated throughput time of each workstation in the routing of job (e.g. ref. [30]).

- Modified Operation Due Date (MODD), which sequences jobs according to operation due date and the earliest possible finish time, that is, max (ODD, $t + p_{js}$), where t refers to the current time when dispatching decision is made; p_{js} refers to processing time at workstation s. This rule was applied, for example, by Fernandes et al. [31].
- Load-oriented dispatching rules speed up the process of jobs and reduce throughput time on the shop floor, which is examined:
- Shortest Processing Time (SPT), which sequences jobs according to the job's processing time at the workstation. The shorter the processing time, the higher the priority. This rule was applied, for example, by Akturk et al. [32]; Schultz [33].

Most Workload Control literature assessed the urgency of jobs at the order release stage. The literature set the same dispatching rule for all jobs, and did not emphasise on urgency of jobs at the stage of shop floor dispatching. That is, improving priority for urgent jobs was taken at the order release stage.

2.3 | Discussion of the literature

Prior research suggested that order release control and shop floor dispatching are the upmost important control stages based on WLC in high-variety make-to-order shops [34]. Order release has received much attention in the literature, which has reported that high priority for urgent jobs and using different rules according to classification of jobs can lead to reductions significantly in percentage tardy and achieve short throughput time [17]. However, few studies have focused on the production of urgent jobs at shop floor dispatching. This study therefore focuses on shop floor dispatching to fill this research gap. Due to the transient change characteristic of urgency of operations at shop floor dispatching, misjudgement of urgent jobs may result in true urgent jobs not being processed in time. In response, this study starts by asking the first research question (RQ1):

Can the delivery performance be improved by considering the transient status of operations in the job's routing?

In order to explore performance differences for different urgent levels, this study proposes the second research question (RQ2):

What impact will the performance of method considering the transient status of operations with different urgent levels on shop floor?

TABLE 2 Additional symbols used in Planned Start Time (PST) rule.

Symbols	
d_j	Due date of job j
a_i	Constant for estimated waiting time at the ith operation in the routing of a job
n_j	Length of the routing sequence of job j
b_i	Constant for estimated throughput time at the ith operation in the routing of a job

3 | PROBLEM DESCRIPTION AND SIMULATION MODEL

3.1 | Problem description

There are two directions of change in the urgency of operations in high-variety shops with various job types: (i) urgent operations are converted into non-urgent operations; (ii) non-urgent operations are converted into urgent operations. The first direction means the reduction of job's urgency. For example, job A is regarded as an urgent job in the pre-shop pool, and visits three workstations—1, 2, and 3—in sequence. From Figure 2a, the actual arrival time of the job A is later than planned start time at workstation 1 and workstation 2. The delay time is resolved by increasing priority continuously at upstream operations of job A's routing (i.e. workstation 1 and workstation 2), which causes the actual arrival time of job A to be earlier than the planned start time at workstation 3. Therefore, the urgency of jobs may be continuously reduced and eventually becomes non-urgent jobs. Consequently, the urgency of job A is reduced at workstation 3. That is, workstation 1 and workstation 2 are urgent operations for job A, and workstation 3 is a non-urgent operation. The second direction means the increase of job's urgency. For example, job B is regarded as a non-urgent job in the pre-shop pool, and visits three workstations—1, 2, and 3—in sequence (see Figure 2b). Allocating production resources to urgent jobs may result in multiple operations backlogs for non-urgent job B at workstation 1 and workstation 2, which may cause job B to arrive later than planned start time at downstream operation (i.e. workstation 3) in the job's routing. As a consequence, the urgency of job B is increased at workstation 3. That is, workstation 1 and workstation 2 are non-urgent operations for job B, and workstation 3 is an urgent operation. The transient

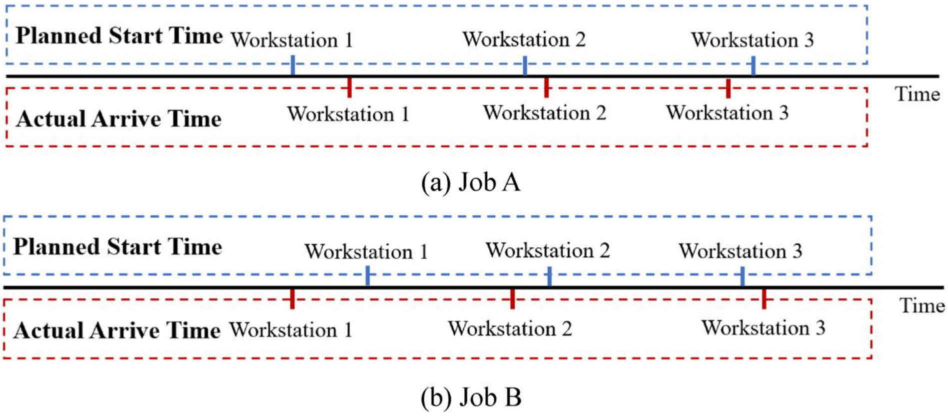

FIGURE 2 Schematic diagram of the Planned Start Time and actual arrive time of job A and job B at each workstation.

change characteristic of urgency of operations at shop floor dispatching implies that transient status of operations should be assessed at each operation, and ensures that actual urgent operations are produced in time to meet delivery performance.

In response, this study proposes a new paradigm for assessing urgency status of jobs. The urgency of jobs (i.e. urgent operations and non-urgent operations) are rejudged at input buffer of each workstation. An integrated time-oriented and load-oriented dispatching rule for urgent operations and non-urgent operations is applied in this study. Load-oriented element for urgent operations speeds up the process when multiple jobs become urgent; time-oriented element for non-urgent operations ensures production in time, thereby reducing tardiness. In addition, this study compares the performance differences of the methods for different urgent levels, using simulation experiments.

3.2 | Simulation model

The shop and job characteristics are first summarised in Section 3.2.1. How this study operationalised the WLC method is then outlined in Section 3.2.2. Section 3.2.3 describes determination methods of job urgency, before the dispatching rules considered are summarised in Section 3.2.4. Finally, a description of experimental design and performance measures is given in Section 3.2.5.

3.2.1 | Shop and job characteristics

This study considers two shop types, a General Flow Shop (GFS) and a Pure Job Shop (PJS). Both shops contain six workstations with equal, constant capacity. The routing length of job varies uniformly from one to six operations. The routing length is first determined before the routing sequence is generated randomly without replacement, that is, re-entrant flows are prohibited; and each workstation has an equal probability of being required in a job routing. This leads to the routing vector for the PJS. For the GFS, the routing vector is sorted such that the routing becomes directed and there are typical upstream and downstream stations. Following previous studies [14, 17, 31], all experiments in the basic design apply a 2-Erlang distribution with a truncated mean of 1 time unit and a maximum of 4 time units for processing times of jobs. The inter-arrival times of jobs follow an exponential distribution with a mean of 0.648 time units. Due dates are established by adding a random exogenous allowance to the entry time of jobs. This study set three time intervals from 30 to 45 time units, 35–50 time units, and 40–55 time units respectively. Three levels of time intervals are treated as different urgent levels.

3.2.2 | WLC order release

As in previous simulation studies on WLC [17, 30], it is assumed that all jobs are accepted, materials are available, and all necessary information regarding shop floor routings, processing times etc. is known. Jobs flow directly into the pre-shop pool at their arrival and await release by the release rules according to Lancaster University Management School Corrected Order Release (LUMS COR) release methods. LUMS COR method is used given its good performance in the literature (e.g. ref. [14]), which uses continuous and periodic release simultaneously. The time interval between releases for the periodic release of LUMS COR is set to 4 time units and sequencing rules are tested with 9 workload norm levels from 4 to 9 time units. As a baseline measure, experiments without controlled order release have also been executed, that is, jobs are released onto the shop floor immediately upon arrival.

3.2.3 | Determination methods of job urgency

In order to prove our research question, this study considers two determination methods of job urgency. The first one does not consider the transient status of operations and does not change the urgency of jobs, namely, *fixed urgent jobs*. Specifically, when the job is just released to the shop floor, this method judges the urgency of job: urgent jobs, that is, jobs with the planned released time that has already passed; and non-urgent jobs. Urgent job will maintain the urgent status until the job is completed. And non-urgent job is considered as non-urgent status at each workstation in the job's routing.

The second one considers the transient status of operations in the job's routing, namely, *variable urgent operations*. Specifically, when dispatching decision is made, this method rejudges the urgency of operations at each input buffer of workstation: urgent operations, that is, jobs with the planned start time at the operation that has already passed; and non-urgent operations. The dispatching procedure is given in Figure 3.

3.2.4 | Shop floor control—priority dispatching

This study applies integrated load-oriented and time-oriented dispatching rules. The load-oriented dispatching rule adopts SPT rule for urgent jobs or urgent operations. The time-oriented dispatching rules adopt ODD rule and developed rule based on ODD for non-urgent jobs or non-urgent operations. The developed rule based on ODD rule is described in Land et al. [16] and Pergher et al. [35], which is as follows:

$$\tau_{js} = t_j^r + i \cdot \frac{\left(d_j - t_j^r\right)}{n_j}, d_j \geq t_j^r \tag{5}$$

$$\tau_{js} = t_j^r, d_j \leq t_j^r \tag{6}$$

where symbols used in Equation (5) and Equation (6) are presented in Table 3. This rule modifies operation due date based on actual released time. For non-urgent jobs/operations in this study, released time is earlier than due date of jobs. Thus, we adopt Equation (5) to modify the operation due date of non-urgent jobs/operations. In addition, urgent jobs/operations will always receive priority over non-urgent jobs/operations.

Therefore, this study considers the following integrated dispatching rules:

(i) Integrated ODD (time-oriented) and SPT (load-oriented), abbreviated as IOS;
(ii) Integrated developed rule based on ODD (time-oriented) and SPT (load-oriented), abbreviated as IDS.

3.2.5 | Experimental design and performance measures

In this study, the experimental factors are: (i) the six workload norms (from 4 to 9 time units); (ii) two determination methods of jobs urgency (fixed urgent jobs and variable urgent operations); (iii) the two dispatching rules (IOS and IDS); (iv) three urgent levels (low, moderate, severe); and (v) two shop types

TABLE 3 Additional symbols used in developed Operation Due Date (ODD) rule.

Symbols	
τ_{js}	The modified operation due date of job j
t_j^r	Job j's release time
b_i	Constant for estimated throughput time at the ith operation in the routing of a job

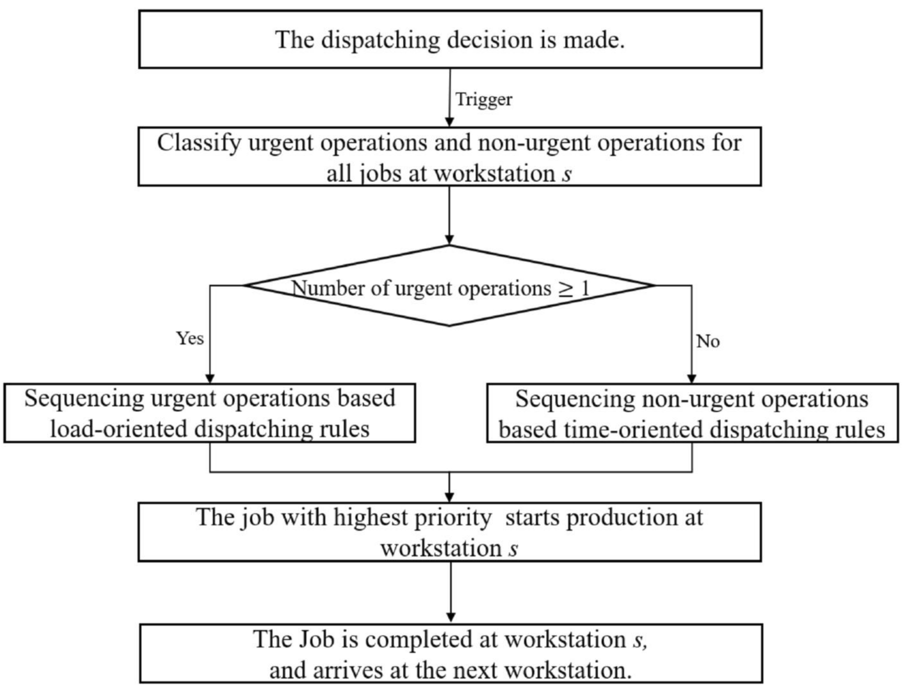

FIGURE 3 The dispatching rule is adopted with considering the transient status of operations in the job's routing.

(GFS and PJS). Using a full factorial design, this results in 144 (6 × 2 × 2 × 3 × 2) scenarios in total. Each scenario is replicated 100 times, and for each replication data is collected for 10,000 time units, being the warm-up period set to 3000 time units. These parameters are in line with those used in previous studies that applied similar shop floor models (e.g. ref. [14]) and allow us to obtain stable results while keeping the simulation run time at a reasonable level.

The four principal performance measures considered in this study are as follows, such as most of the previous literature [31, 36–38]: (i) the *gross throughput time (GTT)*, that is, the completion time of the job minus its entry time; (ii) the *shop floor throughput time (SFTT)*, that is, the gross throughput time minus the queueing time in the pre-shop pool; (iii) the *mean tardiness time (MTT)*, where $T_j = \max\,(0, L_j)$ indicates the tardiness of job j, with L_j being the lateness of job j (i.e. the actual delivery date minus the due date of job j); (iv) the *percentage tardy (PT)*, that is, the percentage of jobs delivered after the customer due date.

TABLE 4 The main effect of ANOVA results.

	Source of variance	*df*	Mean squares	*F*-ratio	*p*-value
GTT[a]	Norm	5	2.36	7.66	0.00
	Classification method	1	119.82	88.67	0.00
	Level of urgency	2	1.55	5.02	0.01
	Shop types	1	0.46	1.48	0.23*
	Dispatching rule	1	0.27	0.09	0.77*
SFTT[b]	Norm	5	133.10	17.63	0.00
	Classification method	1	105.36	88.93	0.00
	Level of urgency	2	1.39	6.45	0.00
	Shop types	1	2.58	11.96	0.00
	Dispatching rule	1	2.39	11.1	0.00
MTT[c]	Norm	5	1.74	5.23	0.00
	Classification method	1	215.93	49.79	0.00
	Level of urgency	2	3.73	11.22	0.00
	Shop types	1	0.32	0.97	0.33*
	Dispatching rule	1	0.12	0.37	0.54*
PT[d]	Norm	5	0.001	3.83	0.00
	Classification method	1	0.058	62.95	0.00
	Level of urgency	2	0.021	29.98	0.00
	Shop types	1	0.002	9.99	0.00
	Dispatching rule	1	0.0001	0.46	0.50*

[a]Gross throughput time.
[b]Shop floor throughput time.
[c]Mean tardiness time.
[d]Percentage tardy.
*Not significant at $\alpha = 0.05$.

4 | RESULTS

Statistical analyses of the results are conducted by using ANOVA to obtain a first indication of the relative impact of the experimental factors. The ANOVA is here based on a block design with the workload level as the blocking factor, that is, the six workload norm levels are treated as different systems. A block design allowed the main effect of the workload norm level and both the main and interaction effects of the dispatching rules, determination methods of urgent jobs, and urgent levels to be captured. Due to space restrictions, we do not present detailed full results and the results of main effect are presented in Table 4. The dispatching rule's main effect is not statistically significant for gross throughput time, mean tardiness time and percentage tardy. Other main effects are shown to be statistically significant except for the shop types in terms of gross throughput time and mean tardiness time. Meanwhile, a small amount of significant three-way interactions in terms of tardiness are observed and there are no significant four-way interactions.

The Scheffé multiple-comparison procedure is used to further prove the significance of the differences between the outcomes of our two determination methods of urgent jobs. In addition to using the results of all experiments, this study also considers subsets—dividing data according to the urgent levels. Test results, as given in Table 5, suggest significant differences in determination methods of job urgency. (i.e. fixed urgent jobs

TABLE 5 Results for the Scheffé multiple comparison procedure: Classification methods of urgent jobs.

			GTT		SFTT		MTT		*s*	
	Method (*x*)[b]	Method (*y*)[b]	Lower[c]	Upper	Lower	Upper	Lower	Upper	Lower	Upper
All[a]	Fixed	Variable	2.66	3.16	2.10	3.60	−3.50	−3.00	−0.06	−0.04
S[a]	Fixed	Variable	1.85	2.59	0.77	3.49	−4.10	−3.29	−0.07	−0.05
M[a]	Fixed	Variable	2.59	3.38	1.61	4.23	−3.59	−2.92	−0.06	−0.04
L[a]	Fixed	Variable	3.14	3.96	2.22	4.78	−3.09	−2.50	−0.05	−0.03

[a]All experiments and three levels of urgent job proportion (i.e. severe, moderate, and low).
[b]Two determination methods of job urgency (i.e. fixed urgency of jobs and variable urgency of operations).
[c]95% confidence interval.
*Not significant at $\alpha = 0.05$.

and variable urgent operations) for all performance measures, focusing on the results obtained for all experiments and different urgent levels. Detailed performance results to further explore shop floor differences are presented in Section 4.1 and urgent level differences are presented in Section 4.2.

4.1 | Performance assessment: Determination methods of job urgency in GFS and PJS

Results for GFS and PJS with low urgent level for different rules are presented in Figure 4, where Figure 4a shows the results of GFS and Figure 4b shows the results of PJS. The dispatching rule is used to create our performance curves. One performance curve represents the result of a dispatching rule. The left-hand starting point of the curves represents the lowest norm level. The norm level increases step-wise by moving from left to right, which each data point representing one norm level.

As the choice of dispatching rule between IOS and IDS does not affect the relative differences between the determination methods of urgent jobs in GFS. And the same holds for PJS. Thus, this study analyses the performance difference from the perspective of determination methods of urgent jobs.

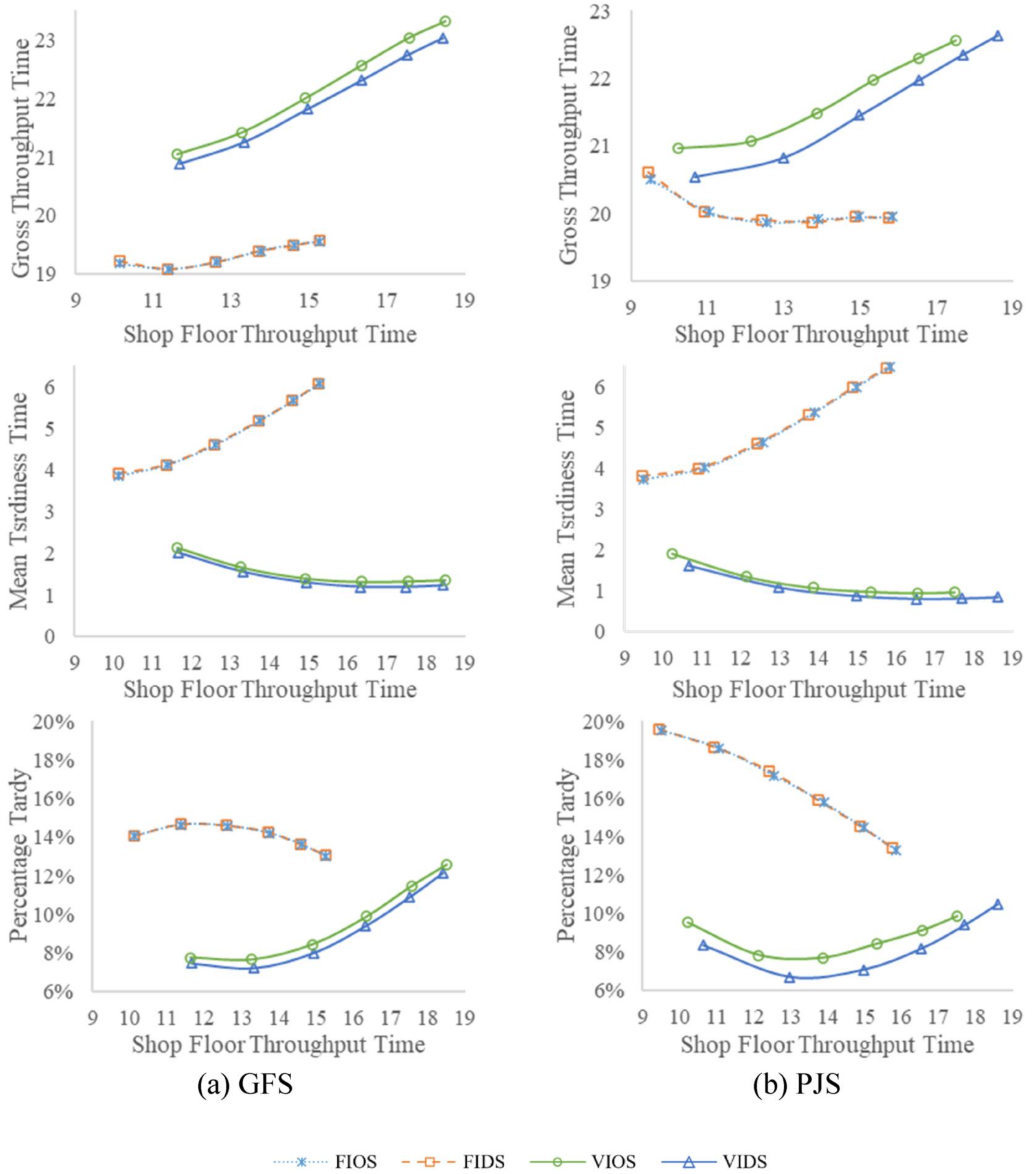

FIGURE 4 The results with the severe urgent level in General Flow Shop (GFS) (a) and Pure Job Shop (PJS) (b), where FIOS and FIDS denote fixed urgent jobs with IOS and IDS dispatching rules respectively; VIOS and VIDS denote variable urgent operations with IOS and IDS dispatching rules.

4.1.1 | Performance assessment in GFS

The results in GFS are illustrated in Figure 4a. In terms of gross throughput time and shop floor throughput time, fixed urgent jobs method outperforms variable urgent operations method, especially during high norm levels. Additionally, the norm levels have more influence on the variable urgent operations method, comparing the fixed urgent jobs method. That is, the gross throughput time of variable urgent operations increases significantly, as the norm levels increase. However, the gross throughput time variation of fixed urgent jobs is not obvious.

In terms of mean tardiness time, variable urgent operations method is superior to fixed urgent jobs method. The tardiness time of variable urgent operations slightly decreases, varied norm level step-wise from four to nine. Slightly improvement is found that to be at the expense of a significantly increase in percentage tardy. The tardiness time of fixed urgent jobs gradually rises, moving from left to right along curves.

In terms of percentage tardy, variable urgent operations method is significantly superior to fixed urgent jobs at the lower norm levels. As the norm levels get looser, the two methods gradually converge. Meanwhile, at higher norm levels, the influence of the norm levels increases on the variable urgent operations method. In other words, moving a data point on the curve during looser norm levels, the variable urgent operations method has a large increase in terms of percentage tardy.

4.1.2 | Performance assessment in PJS

The results in PJS are illustrated in Figure 4b. Comparing the results of PJS and GFS, the fixed urgent jobs method outperforms variable urgent operations method in terms of gross throughput time. For fixed urgent jobs method, the results of GFS are slightly superior to that of PJS in terms of gross throughout time. For variable urgent operations, the results of PJS are slightly superior to that of GFS at the higher norm levels in terms of gross throughput time. Comparing the performance difference between the two determination methods of urgent jobs in GFS and PJS, the differences in terms of gross throughput time reduce in PJS. The curves of fixed urgent jobs method are very close to the curves of variable urgent operations method at the tightest norm level. In addition, the performance differences of dispatching rules in PJS are more obvious than GFS, when we use the variable urgent operations method. In terms of mean tardiness time, we can see that the curves of PJS for mean tardiness time are rather similar to that of GFS. And performance differences across two determination methods of urgent jobs appear to be unaffected by the existence of shop floor type. Therefore, we suggest that the results in terms of mean tardiness time are robust to a change in shop configuration. In terms of percentage tardy, the results in GFS for fixed urgent jobs method outperform that of PJS, especially at the tight norm levels. For variable urgent operations method, the percentage tardy in GFS shows better performance at the tight norm levels. However, results in PJS show better performance at the loose norm levels. Comparing the performance difference between the two determination methods of urgent jobs in GFS and PJS, the performance difference in PJS is larger than GFS at low norm levels, but the results of GFS are larger than that of PJS at high norm levels.

4.2 | Performance assessment: The performance impact of urgent levels

The performance differences in PJS are more obvious than that in GFS. Thus, this study explores and analyses the impact of urgent levels on performance in PJS. The results for different urgent levels in PJS are illustrated in Figure 5, where Figure 5a–c shows the performance at severe, moderate, and low urgent levels respectively. In terms of gross throughput time, the results of variable urgent operations method obviously increase as the norm levels increase, but the results of fixed urgent jobs gradually decrease. The gross throughput time of fixed urgent jobs method increases, as the urgent levels increase. However, the results of variable urgent operations are not affected by urgent levels, which is almost unchanged at the different urgent levels. This same holds for shop floor throughput time.

In terms of mean tardiness time, the results of two determination methods of urgent jobs achieve a reduction by adjusting urgent levels. The performance difference between determination methods of urgent jobs leads to substantial performance improvements, when we reduce the urgent level. From the second row of Figure 5, variable urgent operations method outperforms fix urgent jobs method at all urgent levels. In addition, we can observe that the shape of determination methods curves for mean tardiness time at the severe urgent levels is rather similar to low and moderate urgent levels.

In terms of percentage tardy, the third row of Figure 5 presents that the results of fixed urgent jobs method continuously improve, moving the norm levels from left to right. However, the results of variable urgent operations method firstly reduce and then slightly increase. The percentage tardy of variable urgent operations method has nearly halved, comparing the fixed urgent jobs method, which presents potential improvements. The performance differences between two determination methods of urgent jobs increases at the higher norm levels, reducing the urgent levels.

4.3 | Discussion

Using discrete event simulation it is found that the variable urgent operations method is arguably the better performing method in GFS in our study. The same holds for PJS. From a production perspective, fixed urgent jobs method, which ignores the transient urgent status of operations, results in actual urgent operations are not produced in time due to misjudging urgent jobs, thereby increasing percentage tardy. Based on the above, we recorded, for an arbitrary simulation run, operation

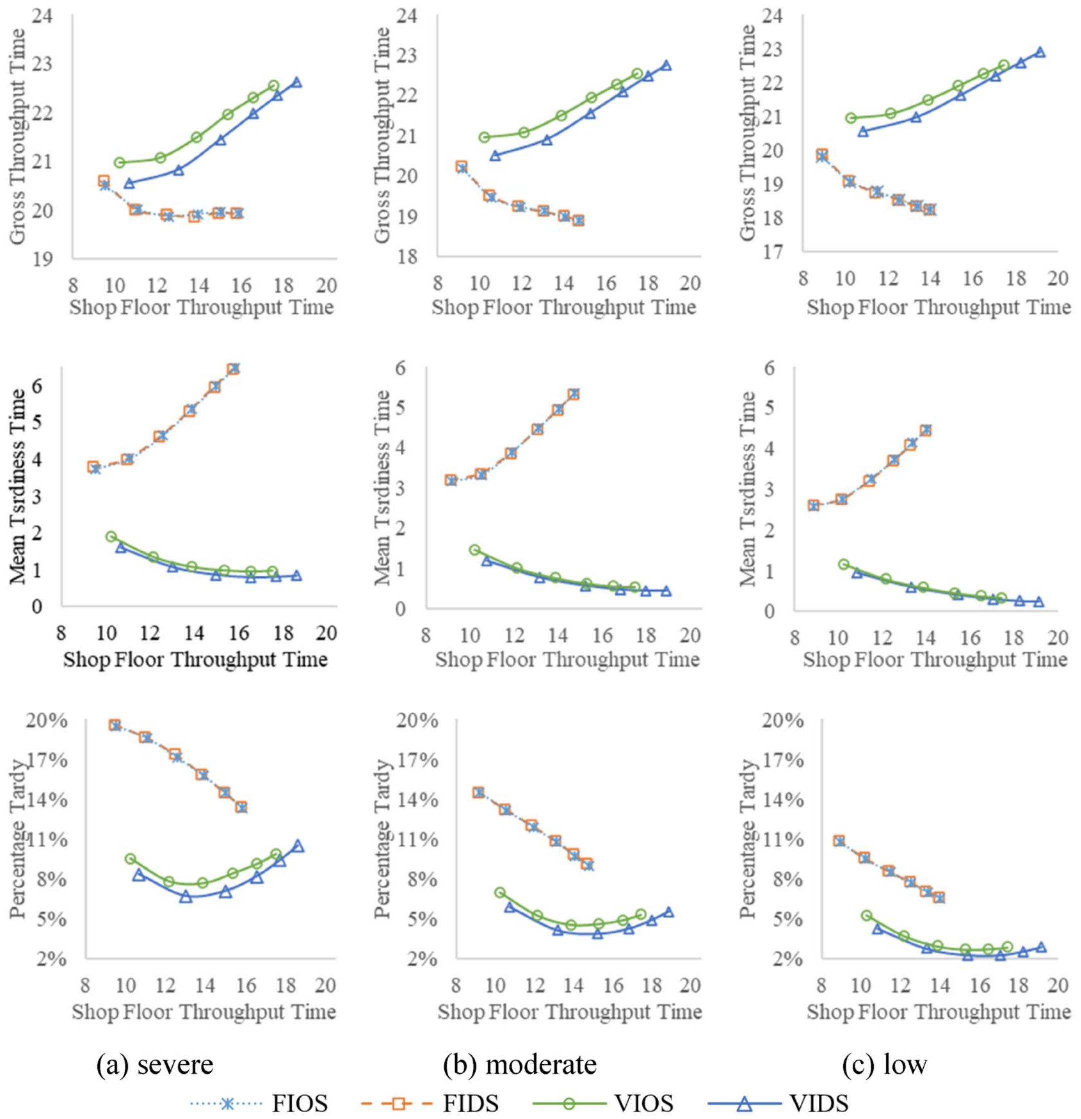

FIGURE 5 The results of Pure Job Shop (PJS) with three urgent levels (i.e. severe (a), moderate (b), and low (c)), where FIOS and FIDS denote fixed urgent jobs with IOS and IDS dispatching rules respectively; VIOS and VIDS denote variable urgent operations with IOS and IDS dispatching rules.

due date of some urgent jobs and non-urgent jobs at the input buffer of workstation, when the fixed urgent jobs method is applied. This is recorded in Table 6. We can see that the operation due date of urgent job is later than that of non-urgent job, since the fixed urgent jobs method maintains the urgency of jobs until jobs are completed. This leads to actual urgent jobs not being processed in time. In addition, we recorded jobs status, arrival time, and completed time at the input buffer of workstation 6 during a time period, as shown in Table 7. We found that continuous arrival of urgent jobs (i.e. urgent jobs B, C, and D) occupied production resources and increased the non-urgent job's waiting time (i.e. non-urgent job A) at the input buffer of workstation, making current time past the operation due date of non-urgent jobs (i.e. urgency of non-urgent jobs may increase over time). The variable urgent operations method, which considers the transient urgent status of operations, rejudges the urgency of operations at each input buffer of workstation (i.e. urgent operations and non-urgent operations). As somewhat expected, this contributes to speeding up production of actual urgent operations, thereby ensuring delivery on time. Meanwhile, the shop floor throughput time are found to be extended. This may be because variable urgent operations method at each operation rejudges the urgency of operations, and the sequencing queue is continuously variable. This leads to severe fluctuation in the waiting time of jobs at the input buffer of workstation. So the shop floor throughput time for variable urgent operations is longer than fixed urgent jobs. At the highest norm level, more jobs are released to the shop floor, which may cause that waiting time increases.

Comparing the results of different urgent levels, it can be found that tardiness time and percentage tardy performance gradually improve as the urgent levels decrease. The tardiness time continuously reduces, moving from left to right along the curves. The percentage tardy decreases and then increases. At the severe urgent level, the performance tardy is greatly affected by the norm levels. This may result from high urgent level and

TABLE 6 Operation due date of some urgent jobs and non-urgent jobs when fixed urgent jobs method is applied.

Workstation	Dispatching decision time	ODD[a] of urgent job *j*	ODD of non-urgent job *j'*
5	33,419	33,445.04	33,434.97
5	33,434	33,465.21	33,453.24
6	33,446	33,467.08	33,456.24
6	33,473	33,490.87	33,487.87
4	33,528	33,552.57	33,534.67
5	33,533	33,550.45	33,543.41
4	33,576	33,596.73	33,573.04
4	33,641	33,662.05	33,650.79
3	33,798	33,828.45	33,812.65
3	33,817	33,833.7	33,828.5

[a]Operation due date.

TABLE 7 The jobs status, arrival time, and completed time at the input buffer of workstation 6.

Job	Urgency of job	Arrival time	Completed time	ODD[a]
A	Non-urgent job	33,443.09	33,500.46	33,466.59
B	Urgent job	33,444.00	33,446.76	33,479.64
C	Urgent job	33,444.00	33,447.37	33,487.45
D	Urgent job	33,467.07	33,444.72	33,448.07

[a]Operation due date.

high norm level. Firstly, the quality of urgent jobs increases at the severe urgent levels. Secondly, when a norm level is increased, more jobs are released to the shop floor, thereby shop floor throughput time. This tends to delay, especially for urgent jobs. That is lead to an increase percentage tardy. Overall, variable urgent operations method improves the performance by half at three urgent levels, comparing fixed urgent jobs method. This presents that the transient status of urgent operations are robust to environmental factors.

5 | CONCLUSION

The production of urgent jobs is a major challenge to meet delivery performance and improve competitiveness for small and medium-sized make-to-order companies. Thus, the production and planning control of urgent jobs becomes an important managerial problem that has received significant practice attention. Existing studies have focused on the order release stage (e.g. ref. [17]), while research on shop floor dispatching has been ignored. The urgent status of operations of jobs may change in the production, due to the fluctuation of shop floor. Therefore, this study proposes a new paradigm for assessing urgency status of jobs. According to the transient urgent status of operations, the status of operations is divided into urgent operations and non-urgent operations at input buffer of each workstation.

According to the transient urgency of jobs, this study uses integrated time-oriented and load-oriented dispatching rules. Load-oriented element for urgent operations speeds up the process when multiple jobs become urgent; time-oriented element for non-urgent operations ensures production in time, thereby reducing tardiness. For our first research question (*Can the delivery performance be improved by considering the transient status of operations in the job's routing?*), the simulation results show that the delivery performance can indeed be improved when the transient status of operations is considered both in GFS and PJS. For our second research question (*What impact will the performance of method considering the transient status of operations be with different urgent levels on shop floor?*), it has been demonstrated that mean tardiness time and percentage tardy decrease significantly with decreasing the urgent levels. Variable urgent operations method is superior to fixed urgent jobs in terms of percentage tardy with different urgent levels. And the percentage tardy performance is improved by half and is greatly affected by norm levels, especially at the severe urgent level. In addition, we can see that shop floor dispatching rules are almost unaffected by other variables, which can be found in ANOVA results.

5.1 | Managerial implications

Simulation experiments results present that considering the transient urgent status of operations contributes to improving mean tardiness time performance and percentage tardy performance. Therefore, the first important managerial implication is that:

In high variety make-to-order shops, managers should consider the transient urgent status of operations and update the urgency of operations to meet delivery performance.

The percentage tardy decreases firstly and then increases, varied norm level step-wise from four to nine. Comparing the percentage tardy performance at different urgent levels, it can be found that variable urgent operations method in terms of percentage tardy performance is greatly affected by norm levels, especially at the severe urgent level. Therefore, the second important managerial implication is that:

Managers should consider both norm levels and urgent levels to ensure delivery performance, especially at the severe urgent level.

5.2 | Limitations and future research

The first main limitation of this study is that we neglected the impact of order release stage. To keep this study focused we mainly considered dispatching rules. Future research could explore how best to combine order release methods with dispatching rules and the impact. A second main limitation of this study is its focus on a manufacturing setting with a classical shop floor layout by simulation. While this allows for a high degree of generalisability, future research should seek to contextualise our findings to real-life job shops, thereby continuing the practice and theory research cycle.

AUTHOR CONTRIBUTIONS

Mingze Yuan: Writing – original draft; formal analysis. **Lin Ma:** Writing – original draft. **Ting Qu:** Project administration; supervision. **Matthias Thürer:** Supervision.

ACKNOWLEDGEMENTS

This paper is financially supported by the National Key Research and Development Program of China (2021YFB3301701), the 2019 Guangdong Special Support Talent Program – Innovation and Entrepreneurship Leading Team (China) (2019BT02S593), the 2018 Guangzhou Leading Innovation Team Program (China) (201909010006), the Science and Technology Development Fund (Macau SAR) (0078/2021/A), and National Natural Science Foundation of China (52375498). We also appreciate the sponsorships from the industry, including but not limited to Carpoly Chemical Group Co., Ltd. GBA and B&R International Joint Research Center for Smart Logistics is a provincial research lab sponsored by the Department of science and technology of Guangdong Province, thanks to which the international collaboration has been effectively conducted.

CONFLICT OF INTEREST STATEMENT

No potential conflict of interest was reported by the authors.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ting Qu https://orcid.org/0000-0003-1012-2856

REFERENCES

1. Ray, S., Jewkes, E.M.: Customer lead time management when both demand and price are lead time sensitive. Eur. J. Oper. Res. 153(3), 769–781 (2004). https://doi.org/10.1016/s0377-2217(02)00655-0
2. Marino, G., Zotteri, G., Montagna, F.: Consumer sensitivity to delivery lead time: a furniture retail case. Int. J. Phys. Distrib. Logist. Manag. 48(6), 610–629 (2018). https://doi.org/10.1108/ijpdlm-01-2017-0030
3. Tersine, R.J., Hummingbird, E.A.: Lead-time reduction: the search for competitive advantage. Int. J. Oper. Prod. Manag. 15(2), 8–18 (1995). https://doi.org/10.1108/01443579510080382
4. Cao, J., Jiang, Z., Wang, K.: Customer demand prediction of service-oriented manufacturing incorporating customer satisfaction. Int. J. Prod. Res. 54(5), 1303–1321 (2016). https://doi.org/10.1080/00207543.2015.1067377
5. Fang, Y., et al.: Digital-twin-based job shop scheduling toward smart manufacturing. IEEE Trans. Ind. Inf. 15(12), 6425–6435 (2019). https://doi.org/10.1109/tii.2019.2938572
6. Huang, T., et al.: An ant colony optimization-based multiobjective service replicas placement strategy for fog computing. IEEE Trans. Cybern. 51(11), 5595–5608 (2020). https://doi.org/10.1109/tcyb.2020.2989309
7. Panzer, M., Bender, B., Gronau, N.: A deep reinforcement learning based hyper-heuristic for modular production control. Int. J. Prod. Res., 1–22 (2023). https://doi.org/10.1080/00207543.2023.2233641
8. Kingsman, B.G., Tatsiopoulos, I.P., Hendry, L.C.: A structural methodology for managing manufacturing lead times in make-to-order companies. Eur. J. Oper. Res. 40(2), 196–209 (1989). https://doi.org/10.1016/0377-2217(89)90330-5
9. Stevenson, M., Hendry, L.C., Kingsman, B.G.: A review of production planning and control: the applicability of key concepts to the make-to-order industry. Int. J. Prod. Res. 43(5), 869–898 (2005). https://doi.org/10.1080/0020754042000298520
10. Mather, H., Plossl, G.W.: Priority Fixation versus Throughput Planning. Mather and Plossl (1977)
11. Missbauer, H., Uzsoy, R.: Order release in production planning and control systems: challenges and opportunities. Int. J. Prod. Res. 60(1), 256–276 (2022). https://doi.org/10.1080/00207543.2021.1994165
12. Thürer, M., Stevenson, M., Qu, T.: Job sequencing and selection within workload control order release: an assessment by simulation. Int. J. Prod. Res. 54(4), 1061–1075 (2016). https://doi.org/10.1080/00207543.2015.1047978
13. Melnyk, S.A., Ragatz, G.L.: Order review/release: research issues and perspectives. Int. J. Prod. Res. 27(7), 1081–1096 (1989). https://doi.org/10.1080/00207548908942609
14. Thürer, M., et al.: Workload control and order release: a lean solution for make-to-order companies. Prod. Oper. Manag. 21(5), 939–953 (2012). https://doi.org/10.1111/j.1937-5956.2011.01307.x
15. Kundu, K., et al.: Order review and release in make-to-order flow shops: analysis and design of new methods. Flex. Serv. Manuf. J. 33(3), 750–782 (2021). https://doi.org/10.1007/s10696-020-09392-6
16. Land, M., Stevenson, M., Thuerer, M.: Integrating load-based order release and priority dispatching. Int. J. Prod. Res. 52(4), 1059–1073 (2014). https://doi.org/10.1080/00207543.2013.836614
17. Thürer, M., et al.: Concerning workload control and order release: the pre-shop pool sequencing decision. Prod. Oper. Manag. 24(7), 1179–1192 (2015). https://doi.org/10.1111/poms.12304
18. Choi, Y., Rhu, M.: Prema: a predictive multi-task scheduling algorithm for preemptible neural processing units. In: 2020 IEEE International Symposium on High Performance Computer Architecture (HPCA), pp. 220–233. IEEE (2020)
19. Land, M.: Parameters and sensitivity in workload control. Int. J. Prod. Econ. 104(2), 625–638 (2006). https://doi.org/10.1016/j.ijpe.2005.03.001
20. Thürer, M., et al.: Lean control for make-to-order companies: integrating customer enquiry management and order release. Prod. Oper. Manag. 23(3), 463–476 (2014). https://doi.org/10.1111/poms.12058
21. Hendry, L.C., Kingsman, B.G.: Customer enquiry management: part of a hierarchical system to control lead times in make-to-order companies. J. Oper. Res. Soc. 44(1), 61–70 (1993). https://doi.org/10.2307/2584435
22. Thürer, M., et al.: Workload control and order release in two-level multi-stage job shops: an assessment by simulation. Int. J. Prod. Res. 51(3), 869–882 (2013). https://doi.org/10.1080/00207543.2012.676685

23. Wisner, J.D.: A review of the order release policy research. Int. J. Oper. Prod. Manag. 15(6), 25–40 (1995). https://doi.org/10.1108/01443579510090318
24. Fredendall, L.D., Ojha, D., Patterson, J.W.: Concerning the theory of workload control. Eur. J. Oper. Res. 201(1), 99–111 (2010). https://doi.org/10.1016/j.ejor.2009.02.003
25. Bagni, G., et al.: Systematic review and discussion of production control systems that emerged between 1999 and 2018. Prod. Plann. Control 32(7), 511–525 (2021). https://doi.org/10.1080/09537287.2020.1742398
26. Sabuncuoglu, I., Karapınar, H.Y.: Analysis of order review/release problems in production systems. Int. J. Prod. Econ. 62(3), 259–279 (1999). https://doi.org/10.1016/s0925-5273(98)00248-5
27. Mezzogori, D., Romagnoli, G., Zammori, F.: Defining accurate delivery dates in make to order job-shops managed by workload control. Flex. Serv. Manuf. J. 33(4), 956–991 (2021). https://doi.org/10.1007/s10696-020-09396-2
28. Raghu, T.S., Rajendran, C.: An efficient dynamic dispatching rule for scheduling in a job shop. Int. J. Prod. Econ. 32(3), 301–313 (1993). https://doi.org/10.1016/0925-5273(93)90044-l
29. Yan, H., et al.: Load-oriented order release (LOOR) revisited: bringing it back to the state of the art. Prod. Plann. Control 27(13), 1078–1091 (2016). https://doi.org/10.1080/09537287.2016.1183831
30. Land, M.J., Gaalman, G.J.C.: The performance of workload control concepts in job shops: improving the release method. Int. J. Prod. Econ. 56, 347–364 (1998). https://doi.org/10.1016/s0925-5273(98)00052-8
31. Fernandes, N.O., et al.: Workload control and optimised order release: an assessment by simulation. Int. J. Prod. Res. 58(10), 3180–3193 (2020). https://doi.org/10.1080/00207543.2019.1630769
32. Akturk, M.S., Ghosh, J.B., Gunes, E.D.: Scheduling with tool changes to minimize total completion time: basic results and SPT performance. Eur. J. Oper. Res. 157(3), 784–790 (2004). https://doi.org/10.1016/s0377-2217(03)00232-7
33. Schultz, C.R.: An expediting heuristic for the shortest processing time dispatching rule. Int. J. Prod. Res. 27(1), 31–41 (1989). https://doi.org/10.1080/00207548908942528
34. Hendry, L.C., Kingsman, B.G., Cheung, P.: The effect of workload control (WLC) on performance in make-to-order companies. J. Oper. Manag. 16(1), 63–75 (1998). https://doi.org/10.1016/s0272-6963(97)00011-9
35. Pergher, I., et al.: Integrating simulation and FITradeoff method for scheduling rules selection in job-shop production systems. Int. J. Prod. Econ. 227, 107669 (2020). https://doi.org/10.1016/j.ijpe.2020.107669
36. Portioli-Staudacher, A., Costa, F., Thürer, M.: The use of labour flexibility for output control in workload controlled flow shops: a simulation analysis. Int. J. Ind. Eng. Comput. 11(3), 429–442 (2020). https://doi.org/10.5267/j.ijiec.2019.11.004
37. Costa, F., Portioli-Staudacher, A.: Labor flexibility integration in workload control in Industry 4.0 era. Operations Management Research 14(3-4), 420–433 (2021). https://doi.org/10.1007/s12063-021-00210-2
38. Mezzogori, D., Romagnoli, G., Zammori, F.: A new perspective on Workload Control by measuring operating performances through an economic valorization. Sci. Rep. 12(1), 14599 (2022). https://doi.org/10.1038/s41598-022-17968-5

How to cite this article: Yuan, M., et al.: Shop floor dispatching with variable urgent operations based on Workload Control: an assessment by simulation. IET Collab. Intell. Manuf. e12084 (2023). https://doi.org/10.1049/cim2.12084