An Assessment of Order Release Models in Hybrid MTO-MTS Flow Shop with Bottleneck

Federica Costa, Kaustav Kundu, Alberto Portioli-Staudacher

An assessment of order release models in Hybrid MTO-MTS flow shop with bottleneck

Federica Costa[1], Kaustav Kundu[2] and Alberto Portioli-Staudacher[1]

Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Milan, Italy[1]
National University of Singapore, Singapore [2]

federica.costa@polimi.it

Abstract
The traditional configurations of market response, make-to-order (MTO) and make-to-stock (MTS), are no more suitable in today's competitive context, forcing companies to a transition towards hybrid solutions. Unfortunately, this has been neglected, in particular at the operational level where the Order Review and Release issue counts only two related articles but restricted to job shops. Therefore, researchers are moving towards flow shop configuration, particularly addressing the issue of bottleneck. Literature links this topic to the Theory of Constraints by Goldratt and Cox [14], from which Drum Buffer Rope (DBR) has been designed as a production planning and control tool. The objective of this paper has been defined as verifying how the decision on the release model changes considering the designed hybrid flow shop model. Simulation results are not clearly in favor of one over the other, since besides severity, there is another important factor to consider: the choice of control at dispatching level, and its connected trade-offs. In particular, implementing this control with bottleneck-based rule is effective for MTO performances, especially in case of high severity. Instead, when control at dispatching is absent, the workload control as release rule is preferred, leading also to the best lead times for low severity.

Keywords: MTO, MTS, Order Review and Release, bottleneck.

1 Introduction

In the current competitive environment, manufacturers deal with mega trends as growing product customization and internationalization of their businesses that ultimately cause shorter product life cycle, the explosion in the number of product families and growing sources of variability [1]. In order to serve the market, firms traditionally choose between make-to-stock (MTS) and make-to-order (MTO) strategies. MTS was typically chosen in low variety context, where products can be standardized, and the focus was on anticipating demand to keep an appropriate level of finished goods stocks. Instead, production within MTO systems is directly triggered by customers' demand, hence the focus is put on assuring short delivery times to clients who can choose among a wide range of families.

However, current context and its business implications force companies to adopt, more and more, a MTO strategy [2, 3, 4, 5]. Actually, for some industries this is not sufficient, making necessary the transition towards a hybrid MTO/MTS production mode [6]. This organizational mode is beneficial for those contexts in which product variability and unpredictability of demand would make inefficient to set up manufacturing activities based on traditional frameworks [7].

Many researches argue the importance of production planning and control mechanism for high-variability MTO context, therefore it is immediate to claim that this statement is valid also for this hybrid scenario, which introduces new and different sources of variability [8]. Despite the relevance of the topic, few are the research effort in the hybrid field, since traditional operations management literature categorizes manufacturing systems as either MTO or MTS [9, 10].

From the previous research, it is revealed that hybrid environment is studied in job shops, not in flow shops. Apart from the load based rules, potential of Drum Buffer Rope (DBR) as production planning and control tool, especially when bottleneck-severity is high is investigated in flow shops [11, 12]. Given these premises, the paper intends to investigate and fill the gap in terms of order release within hybrid flow shop with bottleneck, and adding the recent literature findings on capacity constrained resources.

The article is organized as follows: a theoretical background about the applicability of the Workload Control in hybrid environment followed by the research question is provided in Section 2. Simulation model and methodology are discussed in Section 3. The results and conclusions are depicted in Section 4 along with implications and future research.

2 Literature Review

This section aims at identifying and summarizing the main publications. In hybrid environment, production tries to leverage on the point of strength of both MTS and MTO, assuring shorter delivery lead times compared to pure configurations reaching the required level of product variety. This organizational mode is beneficial for those contexts in which product variability and unpredictability of demand would make inefficient to set up manufacturing activities based on traditional frameworks [7].

Load based rules such as Workload Control (WLC), particularly "Order Review and Release" is proved to be beneficial in MTS and MTO companies but there are very few works in the hybrid MTO-MTS hybrid environment [7, 10]. For details, please refer to Eivazy et al. [10].

Since the reference article by Eivazy et al. [10] explores specifically the semiconductor industry and leaves flow shop without any consideration, the paper intends to enlarge the original study, generalizing the model and addressing the problem in this alternative configuration. The change of perspective in terms of shop floor characteristics makes necessary the deepening of a second stream of literature, specifically addressed towards flow shop with bottlenecks.

Bottlenecks are defined as those portions of a production system which are not able to process any additional job without the use of capacity adjustments. Therefore, they constitute a limit for the output rate, and they have to be carefully considered since companies cannot afford any waste of time for these resources [13]. The management of capacity-constrained resources, in literature, is associated with the Theory of Constraints (TOC) by Goldratt and Cox [14]. In recent years, DBR, a new production planning and control tool adopting the principles of TOC, emerged. The most interesting aspects of recent literature on bottlenecks and DBR are included in four articles by Thürer M., who addressed different thematic: the drum schedule improvement [15], the bottleneck shiftiness [12, 16] and the comparison with other release rule [11].

Constant Work-in-Process (CONWIP) is basically an extension of DBR, since here the control is set at the last station of the shop floor, regardless the position of the bottleneck. The model has been introduced by Spearman et al. [17] as an alternative to Kanban system, which would allow systems to be more robust, flexible and easier to set up. Actually, Framinan et al. [18], in their literature review, recognized that different authors proposed their own conceptualizations, however the most important element was common among all of them: a system which keeps the maximum amount of work-in-process at a constant level.

The implementation of CONWIP is usually associated with the use of cards, which must be attached to a job to be realized in the shop floor without separating them until the end of the last processing stage [19].

Operation management literature recognizes different decision that managers should face when setting up a system based on constant work-in-process [17, 20]: 1. The production quota, 2. The maximum amount of work. 3.The capacity shortage trigger. 4. How to forecast the backlog list. 5. Number of cards operating in the system. 6. How to sequence the jobs in the system.

Researchers clearly specify that these aspects have to be contextualized in a hierarchical planning procedure [21]. In particular, the first five elements of the list could belong to the tactical level, while the last one to the control level.

Given the findings on DBR and the severity factor on one side, and the reference research based on workload on the other, the objective has been defined as verifying how the decision on the release model changes in hybrid MTO-MTS flow shop. Therefore, the research question has been formulated considering observations of the reviewed literature as:

"In the hybrid flow shop model, is Workload Control able to outperform bottleneck-based approach?"

To answer the research question, a simulation model is built and different scenarios are run.

3 Simulation model and methodology

The simulation model has been implemented through Python programming language and some of its dedicated modules. The structure follows the one of the reference study [10] (see Fig. 1), thus it is constituted by a release module and a dispatching module.

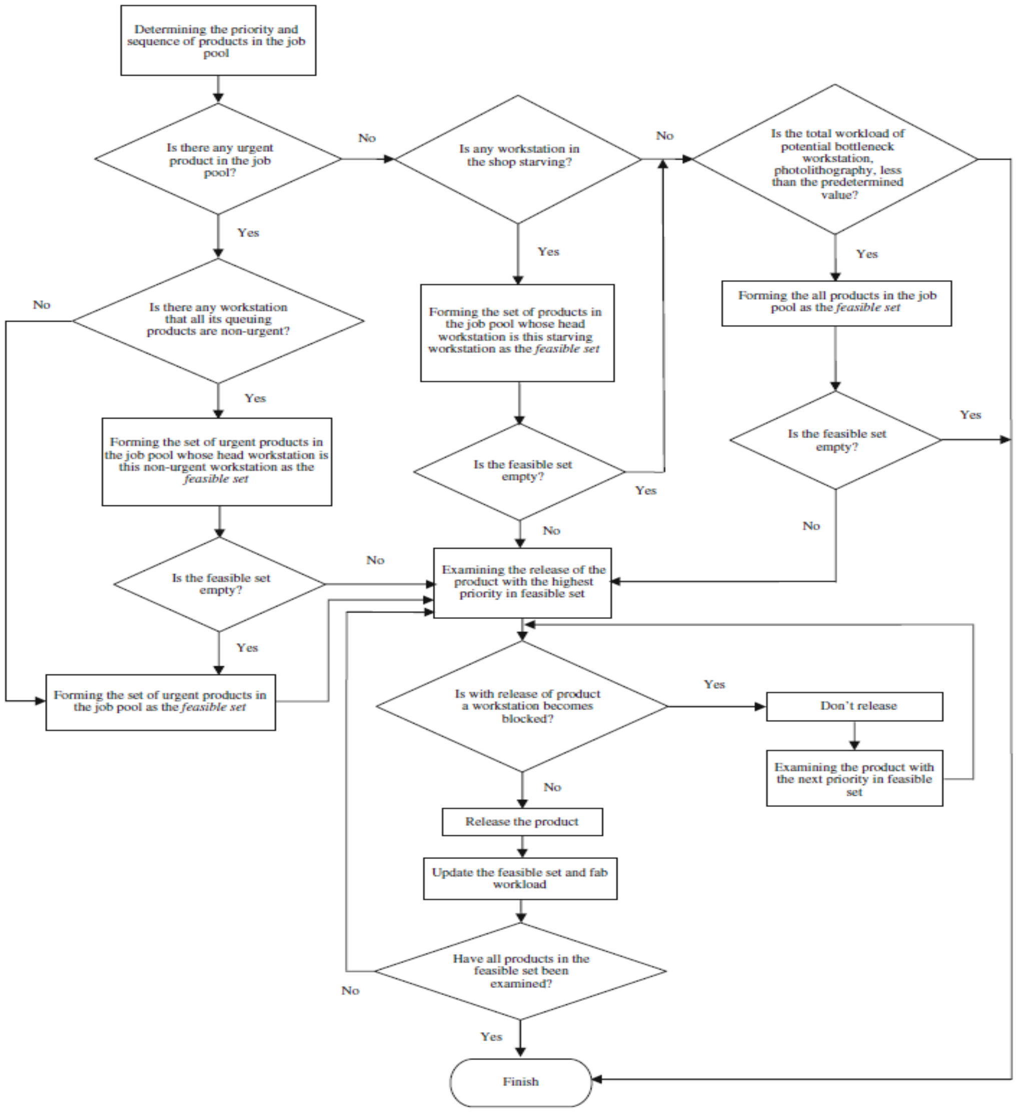

Fig. 1. Release-module algorithm's flowchart, from (Eivazy et al., 2009)

It tests the following release models:

1) DBR: which controls the number of jobs released but not yet completed at the bottleneck station. Orders can enter the shop floor only if the number of jobs are lower than the "buffer limit", and this evaluation is carried out whenever a job is added to the pre-shop pool (PSP) or an operation is completed at the bottleneck station.

2) CONWIP: the structure is similar to DBR, but the controlled station is the last one. Therefore, evaluation of release is based on the arrival at the PSP or when the job leaves the department.

3) Workload Control Continuous (WL-CON): it does not allow the release of a job if its contribution to the workload of stations make one of them to not fit a threshold, called norm. The evaluation of release is activated whenever a job arrives at the PSP or whenever an operation is completed.

4) Immediate Release (IM): it represents the absence of a controlled release.

Among the PSP-sequencing and dispatching rules proposed in the analysed literature, the ones able to identify urgent jobs given a set of orders have been selected. This is necessary to apply the prioritization principle proposed in the reference model: urgent MTO, urgent MTS, non-urgent MTO, and non-urgent MTS. Three alternatives have been selected:

• Critical Ratio (CR): it is the same rule applied by Eivazy et al. [10], hence it already fits the hybrid context. The information on due dates and predicted cycle time are drawn from pre-simulation runs.

• Modified Planned Release Date (MODPRD) and Modified Planned Bottleneck Start Time (MODPBST): they are, respectively, the PSP-sequencing and dispatching rules proposed for the drum schedule improvement. They have been conceived for the MTO context, therefore some modifications were necessary. Concerning the MTO jobs, the measure to establish the urgency remains the same of the literature, instead, for MTS it is applied the one of Critical Ratio, based on WIP and its planned values (taken from pre-simulation runs).

• First in First out (FIFO): which represents the absence of prioritization mechanism.

Models and rules have been tested in a flow shop configuration constituted by 5 stations as in previous studies [4, 22, 23], in which only one could be the bottleneck in the same simulation scenario. Referring to the previous research [15], the first station (Position 1) is considered to be the bottleneck position. Besides the common simulation entities that could be found in similar publications, another one has been added: the warehouse. This object store MTS jobs that have just been processed at the corresponding machine. This is important to realize the logic of MTS: customers withdraw products kept in the Finished Good Warehouse, and this triggers a replenishment mechanism that flow upstream till the PSP.

The processing time setting values are drawn from a statistical distribution taken from the literature, while in the base case a deterministic "total processing time" is established. MTO and MTS jobs differ for the variance of their distribution (lower for MTS, since more standardized), while their means are equal. Two levels of bottleneck severity have been selected, as to account for the following conditions: high and low [11].

Another aspect for the design of the model concerned the identification of the relative weight of MTS and MTO on the overall capacity managed by the system. Therefore, we set 480 minutes as the daily available time and 30 minutes as the average processing time of a job, the maximum daily capacity to 16 jobs. However, the value corresponds to the maximum utilization of machines, thus the value has been lowered to 15 jobs, as to account for a minimum degree of spare capacity. Once these values have been identified, another factor, called "MTS Weight", established the weight of MTS on these 15 jobs. These computations were fundamental for establishing: the arrival rate of MTO and the mean of the normal distribution constituting the demand of MTS products.

Finally, preliminary runs of simulations have been necessary for the setting of due date's parameters, predicted cycle times of both MTO and MTS and values for norms, buffer limit and total WIP. Parameters of the simulation model and the experimental setting are represented in Table 1.

Table 1. Design of Experiment

Shop configuration studied	Pure flow shop
Number of machines	5
Capacity of each stage	480 time unit
Arrival rate distribution	Exponential
Processing time distribution MTO	Truncated Log-normal Mean 30, variance 900 Minimum 0, maximum 360
Processing time distribution MTS	Truncated Log-normal Mean 30, variance 40 Minimum 0, maximum 360
Jobs contractual due date	Uniform [α, β]
ORR model	WL-Continuous, DBR, CONWIP, IM
Sequencing Rule	FIFO, CR, MODPRD
Dispatching Rule	FIFO, CR, MODPBST
Rule levels (Norms, Buffer Limit, Total WIP)	6-levels for each rule WLC-Norms: 72, 90, 120, 150, 420, 540 DBR-Buffer Limit: 2, 4, 6, 10, 15, 35 CONWIP-Total WIP: 10, 15, 20, 30, 60, 70
Bottleneck Position	Position 1
Bottleneck Severity	0.65, 0.80
MTS Weight	30% of 15 jobs

It is important to notice that not all combinations between sequencing and dispatching factors have been included in the experiment. It is more likely that a company adopts a unique logic for these prioritization aspects, therefore the selected levels respect this observation (in all the document, the first acronym stands for sequencing, the second for dispatching):

1. FIFO – FIFO
2. FIFO – CR
3. CR – CR
4. FIFO – MODPBST
5. MODPRD – MODPBST

The response variables monitored are: MTO Gross Throughput Time (GTT), that is the overall lead time, from arrival at the PSP to the completion; the MTO Shop Floor Time (SFT) that is the time from the release to the completion; the tardiness, composed by percentage of tardy jobs, mean tardiness and standard deviation of lateness; and the production achievement, that is the performance of interest for MTS jobs, and it is defined as the average values of the percentage of the daily customer demand that the company is able to satisfy. Each scenario is replicated 10 times and the average value for each performance measure is used for analysis. Paired t-test is performed whenever there is significant difference in results.

4 Results and conclusions

The analysis of results has been conducted following the concept of scenario defined by the levels of: bottleneck position, severity and MTS weight. We firstly reported a preliminary investigation aimed at understanding whether the scenario-factors produce results in line with the literature and with the design phase. Expectations are generally respected. However, CONWIP shows better performances when the bottleneck is positioned at the first station, in contrast to findings in Thürer et al. [18], but in line with previous researches [15]. This deviation is reasonable considering the impact of dispatching rules, since in prior studies it is shown how this may lead to different conclusions; the fact that this simulation uses different parameter setting could affect results.

The core of the analysis focused on the control of the time-performance curves (defined as GTT versus SFT) of the different release rules and on their relative positioning varying scenarios factors. As found also in previous studies, the PSP-sequencing does not add relevant benefit to time performances. Therefore, it is not considered in the main analysis. But it can be observed that to select the proper release rule, the choice of the dispatching is a very significant one to consider, given different levels of severity. In contexts characterized by high severity (0.80), MODPBST is the best performing rule, and especially in combinations with the DBR rule, especially at shorter shop floor time (Fig. 2).

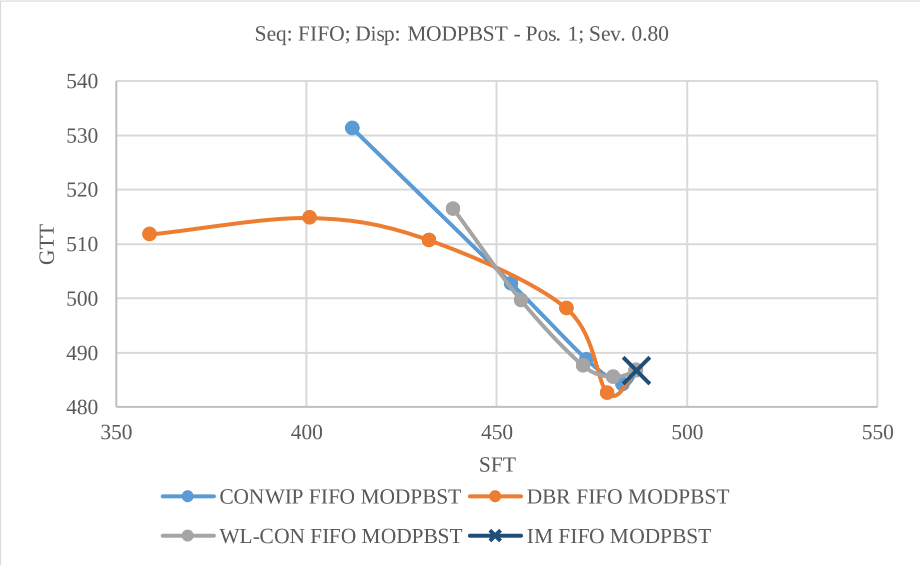

Fig. 2. Performance curve comparison in case of high severity and MODPBST dispatching.

On the other hand, if FIFO rule is applied under high severity, WLC is the most appropriate release rule (Fig. 3). The difference between Fig. 2 and Fig. 3 is explained by the Immediate Release point. Immediate Release is obtained at much lower shop floor time with MODPBST rule than with FIFO rule. This justifies MODPBST performs better than FIFO dispatching rule.

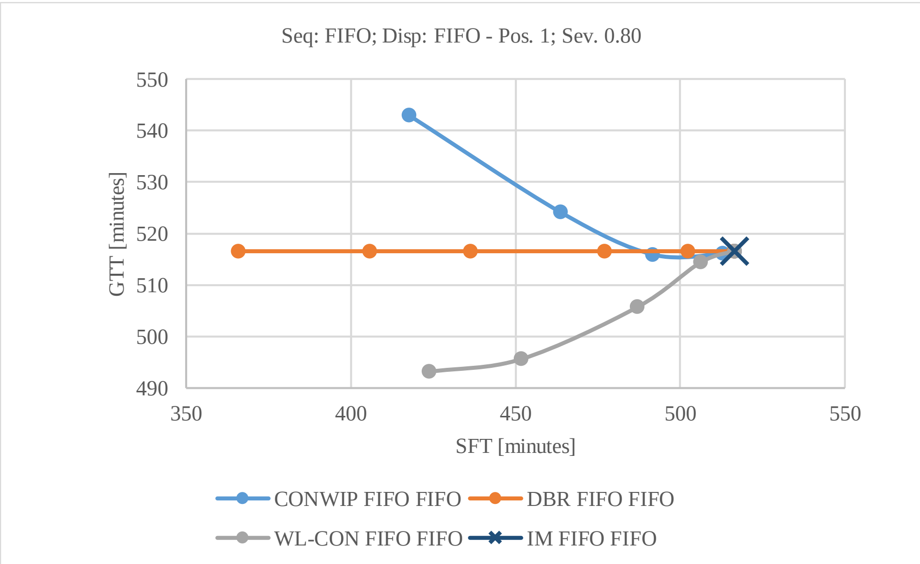

Fig. 3. Performance curve comparison in case of high severity and FIFO dispatching.

At severity level 0.65, WLC is still the best choice compared to the other possibilities with FIFO dispatching. It seems that WLC does not require additional control at dispatching in order to obtain the best performance. A plausible justification comes from the balancing capabilities of WLC that effectively manage the differences between the two typologies of orders. The last rule considered, CR, is completely outperformed by others: probably due to fact that the rule exploits static information on cycle times and also to its inner inefficient effects.

Once the best values for the time performances of MTO (the most important measure) have been found, for sake of completeness, the analysis continues investigating the trade-offs for the other dimension of interest that could be identified choosing between different dispatching rules (i.e. tardiness and production achievement).

It is shown how MODPBST, as in literature, guarantees the best performances in terms of tardiness and percentage tardy, while FIFO worsens these values. Contrarily, looking at production achievement of MTS, FIFO is the most performing rule, since it does not implement an explicit system of prioritization of MTO over MTS.

Given the observations on this part of simulation results, it can be said that the categorical dominance of one compared to the other cannot be stated. Not only the decision of the proper release rule depends on the severity level, as found in the pure MTO context, but also it has to be considered whether a control at dispatching level is implemented. In fact, it influences tradeoffs which decision makers have to carefully analyze in terms of time, tardiness and production achievement.

4.1 Implications and future research

From a practical point of view, the research highlights the main trade-offs that managers have to face in this type of hybrid MTO-MTS production system. In case of high severity, managers have to take into account that choosing MODPBST as dispatching leads to the best performance in terms of MTO lead times and tardiness, which implies higher control cost and lower capability of meeting daily demands of MTS product. Instead, in case of low severity, MODPBST loses its better positioning in terms of MTO lead time. When choosing FIFO as dispatching, thus Workload Control at release level, specular performances for the mentioned indicators are observed at different severity levels.

The results presented in this paper depend on the validation made during the design of the simulation model and its parameters. The research is the first attempt to conceptualize this context through the presented model, therefore future researches are needed to test the worthiness of results. Other possible streams for future efforts may include: different positions of the bottleneck, different values for the MTS weight, different values for the bottleneck severity, the study of environmental factors (e.g. machine failure); testing periodic release models; the deepening of the relation between CONWIP and DBR; and finally, the study of the production achievement performance and the CR dispatching rule.

References

1. Francas, D., Löhndorf, N., Minner, S.: Machine and labor flexibility in manufacturing networks. Int. J. Prod. Econ. 131(1), 165-174 (2011)
2. Kingsman, B., Hendry, L., Mercer, A., De Souza, A.: Responding to customer enquiries in make-to-order companies Problems and solutions. Int. J. Prod. Econ. 46-47, 219-231 (1996)
3. Kingsman, B., Hendry, L.: The relative contributions of input and output controls on the performance of a workload control system in make-to-order companies. Prod. Plan. Control 13(7), 579-590 (2002)

4. Kundu K., Land M.J., Portioli-Staudacher A., Bokhorst, J.A.: Order Review and Release in Make-To-Order Flow Shops. Flex. Serv. Manuf. J., https://doi.org/10.1007/s10696-020-09392-6 (2020)
5. Rossini, M., Audino, F., Costa, F., Cifone, F., Kundu, K., Portioli-Staudacher, A.: Extending lean frontiers: a kaizen case study in an Italian MTO manufacturing company. Int. J. Adv. Manuf. Tech. 104, 1869-1888 (2019)
6. Eivazy, H., Rabbani, M.: A novel production control model for manufacturing systems with pure MTS, pure MTO, and hybrid MTS-MTO products. *2009 International Conference on Computers and Industrial Engineering, CIE 2009*.
7. Fernandes, N.O., Gomes, M., Carmo-Silva, S.: Workload control and order release in combined MTO-MTS production. Rom. Rev. Prec. Mech. Opt. Mechatron. 33-39 (2013)
8. Małachowski, B., Korytkowski, P.: Competence-based performance model of multi-skilled workers. Comput. Ind. Eng. 91, 165-177 (2016)
9. Soman, Chetan Anil, Van Donk, D.P., Gaalman, G.: Combined make-to-order and make-to-stock in a food production system. Int. J. Prod. Econ. 90(2), 223-235 (2004)
10. Eivazy, H., Rabbani, M., Ebadian, M.: A developed production control and scheduling model in the semiconductor manufacturing systems with hybrid make-to-stock/make-to-order products. Int. J. Adv. Manuf. Tech. 45(9), 968-986 (2009)
11. Thürer, M., Stevenson, M., Silva, C., Qu, T.: Drum-buffer-rope and workload control in High-variety flow and job shops with bottlenecks: An assessment by simulation. Int. J. Prod. Econ. 188, 116-127 (2017b)
12. Thürer, M., Stevenson, M.: Bottleneck-oriented order release with shifting bottlenecks: An assessment by simulation. Int. J. Prod. Econ. 197, 275-282 (2018a)
13. Ashcroft, S.H.: Applying the principles of optimized production technology in a small manufacturing company. Engg. Cost. Prod. Econ. 17(1-4), 79-88 (1989)
14. Goldratt, E.M., Cox, J. (1984). The Goal: A Process Of Ongoing Improvement, N*orth River Press, Croton-on-Hudson*, NY.
15. Thürer, M., Stevenson, M.: On the beat of the drum: improving the flow shop performance of the Drum–Buffer–Rope scheduling mechanism. Int. J. Prod. Res. 56(9), 3294-3305 (2018b)
16. Thürer, M., Qu, T., Stevenson, M., Li, C.D., Huang, G.Q.: Deconstructing bottleneck shiftiness: the impact of bottleneck position on order release control in pure flow shops. Prod. Plan. Control 28(15), 1223-1235 (2017a)
17. Spearman, M.L., Woodruff, D.L., Hopp, W.J.: CONWIP: A pull alternative to kanban. Int. J. Prod. Res. 28(5), 879-894 (1990)
18. Framinan, J.M., González, P.L., Ruiz-Usano, R.: The CONWIP production control system: Review and research issues. Prod. Plan. Control 14(3), 255-265 (2003)
19. Hopp, W., Spearman, M.: Basic Factory Dynamics. In Factory Physics: Foundations of Manufacturing Management. (1996)
20. Herer, Y.T., Masin, M.: Mathematical programming formulation of CONWIP based production lines; and relationships to MRP. Int. J. Prod. Res. 35(4), 1067-1076 (1997)
21. Spearman, M.L., Hopp, W.J., Woodruff, D.L.: A hierarchical control architecture for Constant Work-in- Process (CONWIP) production systems. J. Manuf. Opers. Manag., 2, 147–171 (1989)
22. 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)
23. Kundu, K., Rossini, M., Portioli-Staudacher, A.: Analysing the Impact of Uncertainty Reduction on WLC Methods in MTO Flow Shops. Prod. Manuf. Res., 6(1), 328–344 (2018)