Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnLine 54-1 (2021) 743–748

A Digital Twin-based Predictive Strategy for Workload Control

Lorenzo Ragazzini*, Elisa Negri*,
Marco Macchi*

*Dept. Management, Economics and Industrial Engineering, Politecnico di Milano, Milan, MI 20156
Italy (Tel: +39 0223993969; e-mail: lorenzo.ragazzini@polimi.it)

Abstract: **The paper aims at proposing a card controlling model to improve the standard CONWIP procedure, granting a similar system throughput while reducing Work In Progress (WIP) levels. To achieve this objective, the authors developed a Digital Twin-based production control system including a reinforcement learning algorithm (i.e. Q-Learning). The Digital Twin is responsible for short term predictions of the behavior of the system aimed at a what-if analysis with different numbers of cards. As there is lack of evidence of research related to Digital Twin applications for production control and for order release systems in particular, we aim at proposing this as an initial work to start the exploration of problems in this control area. The proposed model has been tested both in a Job Shop and in a Flow Shop systems with promising results.**

Keywords: **Digital Twin, Production Control, Workload Control, Order Release, Card Controlling, Reinforcement Learning**

1. INTRODUCTION

In the last decade, Cyber-Physical Systems (CPS) enabled a new level of connection between production equipment and industrial information systems. This is leading to emergence of relevant characteristics for the smart factories, among them virtualization (Napoleone, Macchi and Pozzetti, 2020). Digital Twin (DT) is indeed at the basis of this characteristic, and is potentially bringing about ideas for novel manufacturing control practices.

Within manufacturing, the Digital Twin (DT) is considered as a probabilistic multidomain simulation of a production system characterized by the synchronization between the digital and physical systems, with the constraint of real-time data elaboration (Negri, Fumagalli and Macchi, 2017). In addition, the connection between the DT and its physical counterpart must allow the flow of data in both directions (Kritzinger *et al.*, 2018). To close the control loop and cope with the highlighted bilateral communication requirement, Negri et al. states that a supervisory control feature of the DT should exist and lie in an intelligence layer detached from the DT itself (Negri *et al.*, 2020). The authors agree that the controlling agent located in this layer elaborates the information coming from the monitoring level to make the decisions which will feed the physical model. In the context of manufacturing systems DT simulation allow to replicate, evaluate, and validate shop-floor production processes, providing results useful to perform an effective optimization of production policies (Zhuang, Liu and Xiong, 2018).

Some effort has been made to study the potentiality of the DT for production planning and scheduling. Negri et al. developed a DT which was integrated with a scheduling tool based on a Genetic Algorithm and with a module for health prediction of the equipment (Negri *et al.*, 2019, 2021). In this framework, scheduling optimization is more robust since it considers production asset degradation through failure rate predictions in time. It may help in predicting and detecting disturbances, improving scheduling performances and enabling to trigger rescheduling promptly, as envisioned also by other researchers (Zhang, Tao and Nee, 2020). These are just a few examples indicative of possible uses of DT to support the short term prediction and control problems.

Despite the concrete targets envisioned in the examples, how to exploit Digital Twins (DT) for decision making in the control of an industrial plant is still an open challenge in research. Nevertheless, it is worth making further studies as the monitoring of production data is of paramount importance in order to develop an effective control of production processes (Urbina Coronado *et al.*, 2018). Therefore, a DT enabling online simulation and analysis may be used with this purpose.

This paper proposes a progress within the well-known theory on order release; to this regard, it investigates how the DT may be effectively used in order to improve the dynamic control of a production system. The paper is structured as follows: Section 2 sets the founding principles on which the production control has been dealt in literature, Section 3 states the research design, Section 4 proposes the innovative DT-based model for production control, Sections 5 and 6 respectively describe the experimental setup and the results, Section 7 reports some concluding remarks.

2. BACKGROUND ON PRODUCTION CONTROL

2.1 Order release

Within production control, the decision concerning the time point at which an order is allowed to be processed is addressed by order release procedures (Lödding, 2013). Order release can be classified according to:

- the criteria defining the conditions based on which decisions are made,

10.1016/j.ifacol.2021.08.183

- the degree of detail, which determines if the decisions apply to the entire order or to individual operations,
- the trigger logic, which determines when decisions about the order release will be initiated.

Order release can be classified as periodic, which is done at regular time steps, or event-oriented, which is triggered after a specific event (Lödding, 2013).

Traditionally, push and pull paradigms can be identified within production management (Hopp and Spearman, 2004). In push systems, decision regarding when to produce are based on short-term plans and the orders are "pushed" into the production system according to those plans. Pull systems instead receive the trigger to move jobs only when one of the subsequent resources requires the material to be processed, in this way "pulling" them (Garetti *et al.*, 2016).

2.2 CONWIP

Among pull approaches, CONWIP is considered more robust, more flexible, and easier to implement than the others, while sharing the advantages of these over push systems with respect to WIP control. Manufacturing companies interested in controlling inventory levels in uncertain and dynamic environments consider these characteristics particularly important (Framinan, González and Ruiz-Usano, 2003).

The idea behind CONWIP is that maintaining a constant Work in Progress (WIP) within the shop floor is a good way to optimize the throughput of the system. Unfortunately real systems rarely fulfil the hypotheses of CONWIP (i.e. similar routings and processing times among jobs and no setup times) (Hopp and Spearman, 2008). For this reason, in practice not much can be achieved using the simplest formulation of CONWIP, but it is possible to adapt CONWIP to achieve more reliable results (Hopp and Spearman, 2008). In general, CONWIP works by controlling the overall inventory in the system relying on cards or signals. Originally, a card was attached to a container of parts at the beginning of the line. When the container reached the end of the line, the card was removed and attached once again to a container in the beginning of the line (Spearman, Woodruff and Hopp, 1990). Pull systems are managed by choosing WIP levels and observing the resulting throughput; therefore, system performance is critically dependent on the WIP levels selected. In those systems that use cards (sometimes also referred to as *kanban*) to manage WIP, this is done by changing card counts. The number of cards, which define the WIP level must be large enough to ensure a throughput performance sufficient to meet customer demand (Hopp and Roof, 1998).

Specifying the number of cards in a CONWIP system consists of two procedures that may be simultaneously addressed. These are card setting, and card controlling. *Card setting* refers to setting the number of cards that makes the system perform acceptably according to predefined performance measures. It does not consider that the number of cards can change. *Card controlling* focuses instead on the development of rules to change the current number of cards of a CONWIP system (Framinan, González and Ruiz-Usano, 2003).

Liu and Huang developed a procedure based on automatic control and queueing theory (Liu and Huang, 2009). The authors set the target throughput as the reference input of the control system and compute CONWIP parameters considering each station to behave like a M/G/1 queue. The results show that the proposed procedure is competitive, and the authors analysed through a simulation study how the processing time impacts the performance of the CONWIP system. Renna et al. proposed a methodology which relies on the analysis of customer demand to adjust the number of cards (Renna, Magrino and Zaffina, 2013). The authors evaluate fluctuations of customers' demand according to two moving averages with different horizons. The signal of the controller is developed from the comparison between the two moving averages. The procedure developed by Tardif and Maaseidvaag uses 'extra cards' to change the WIP level. These 'extra cards' are added to the system when the inventory level falls under a release threshold; they are withdrawn if the finished goods inventory level is above a capture threshold (Tardif and Maaseidvaag, 2001). Hopp and Roof proposed an adaptive method named Statistical Throughput Control (STC). By using cards and real-time throughput measurements, they were able to improve CONWIP adjusting WIP levels to meet target production rates under uncertainty (Hopp and Roof, 1998).

3. RESEARCH DESIGN

Although the CONWIP procedure has attracted a lot of attention from industry and academy since its introduction, most of the research regarding the number of cards refers to card setting and only few works have delt with card controlling procedures (Framinan, González and Ruiz-Usano, 2003). Moreover, the DT has been rarely applied for production control, and it seems that no authors have utilized it for improving order release procedures. This works aims at closing this gap by developing a novel DT-based production control model.

The research objective is to develop a new order release model for DT-based card controlling to improve CONWIP protocol according to future estimates and real-time measure of the production rate. This means that the proposed method aims at implementing both the real-time capability and predictability as relevant characteristics, joining them with virtualization; this is well aligned with the expected requirements of the smart factories (Napoleone, Macchi and Pozzetti, 2020).

This proposal will be validated by demonstrating that the WIP is reduced, while system throughput is improved or remains constant.

4. PROPOSED MODEL

The proposed order release model is based on three founding elements: (i) the CONWIP protocol, (ii) a DT of the production system and (iii) a Q-Learning algorithm. It allows to change the WIP level dynamically to improve system throughput and to reduce WIP level. In fact, the algorithm attempts to decide between three different scenarios of WIP level: whether to increase or decrease the WIP level by one unit, or to keep it constant. Although the model is based on the STC developed by Hopp and Roof (Hopp and Roof, 1998), target throughput does not have to be set a priori, thus allowing to dynamically maximize it. The optimization model is triggered each time an order is completed and leaves the system. In case a product leaves the system while an optimization process is still ongoing, a new order is simply released to compel with the WIP level which was set previously. CONWIP card controlling problem has been modelled as a Markov Decision Problem (MDP) considering the possibility to increase or decrease the card number by one unit at a time, or to keep its value constant. Thus, it can be optimized by using reinforcement learning (Sutton and Barto, 2018), shortly RL in the reminder. The complete framework is shown in Fig. 1.

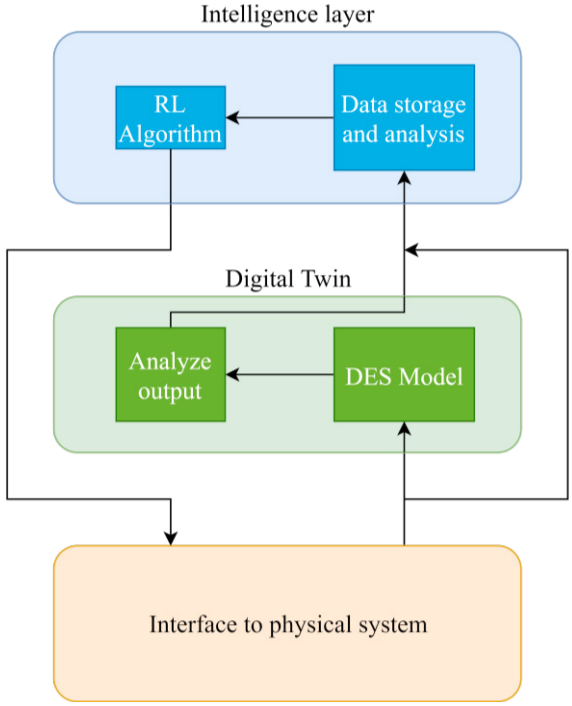

Fig. 1. Conceptual Framework

The flow of information in the proposed model shows that the system state S is gathered from the interface to the physical system and fed to the Discrete-Event System (DES) simulation model of the DT (described in Section 4.1). Simulation results are exported together with other relevant field data and processed in the intelligence layer. Here, a Q-Learning Algorithm makes decisions concerning the number of cards which will be set for the system (described in Section 4.2).

4.1 Digital Twin

The DT is based on a DES simulation model which can perform what-if analysis allowing to set different levels of WIP. From a DT viewpoint, the card controlling procedure is triggered when an order leaves the system. In that moment the DT is synchronized with the interface to the physical system: the state of the orders and their position on the shop floor is fed to the DT. Different simulations are performed starting from that time instant. Hereafter, the control agent is triggered to make a decision and the feedback signal returns to the interface to the physical system.

The DT simulation is run starting from the synchronization time t_0 for a fixed time interval, which has been set to 900 seconds (i.e. 15 minutes). For each of the three alternative card numbers n (+one, constant, or -one), the model performs 20 replications. The results of the three simulated scenarios are (i) the prediction of the average inter-output time of the orders λ_n [s] and (ii) the WIP level w_n [parts] for each card number n tested. $\lambda_{\min}$ is the minimum value among the found λ_n, which ensures the maximum throughput (Equation 1). The corresponding expected average WIP level is w_e (Equation 2).

$$\lambda_{\min} = \min \lambda_n \tag{1}$$

$$w_e = w(\operatorname{argmin}(\lambda_n)) \tag{2}$$

4.2 Intelligence layer

As stated in previous sections, the DT must be connected by a bilateral communication with its physical counterpart. For the purpose of this work, the feedback signal going from the virtual to the real space consists in a control action setting the system card number (i.e. the allowed WIP level). The agent in charge of doing this lies in the intelligence layer of the proposed DT. The supervisory control is then implemented at this layer.

To the purpose of this work, the intelligence layer includes a Reinforcement Learning (RL) algorithm. In fact, a Q-Learning algorithm was developed due to its ease of implementation and to the existence of a proof of its fit to production control (Dittrich and Fohlmeister, 2020).

Every decision (or action) is evaluated according to a reward function r_t plus the expected future rewards of subsequent decisions, discounted by the discount rate γ (Watkins and Dayan, 1992). This optimization process is based on the well-established action-value function by (Watkins and Dayan, 1992), that is reported in Equation 3:

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha\left[r_n + \gamma \max_b Q(s_{t+1}, b)\right] \tag{3}$$

where the evaluation of each possible state-action pair with an action a_t in a system state s_t at time t is calculated.

The proposed Q-Learning algorithm observes state s described by the actual card number n and factor Δ, which is the difference between measured and predicted inter-output times of the orders over the standard deviation of the first measured inter-output times. Actions can be represented by the integer values [+1, 0,-1].

The inter-output time of orders μ_t in [s] is measured and recorded in the intelligence layer. Its standard deviation is computed considering a moving average over the last m jobs completed, as in Equation 4:

$$\sigma_{t_0} = \sqrt{\sum_{t=t_0-m+1}^{t_0}\left(\mu_t - \bar{\mu}_{t_0}\right)^2} \tag{4}$$

Where t_0 is current time instant and $\bar{\mu}_t$ is the mean value of μ_t in the interval $[t-m, t]$, as defined in Equation 5:

$$\bar{\mu}_t = \frac{1}{m}\sum_{t=t_0-m+1}^{t_0} \mu_t \tag{5}$$

Factor Δ is computed as in Equation 6, inspired by Hopp and Roof (Hopp and Roof, 1998):

$$\Delta = \frac{\overline{\mu}_{t_0} - \lambda_{\min}}{\sigma_{t_0}} \tag{6}$$

And bounded to the interval [+3,-3], extreme values included.

The reward r_t signal is computed as in Equation 7:

$$r_t = -\frac{\overline{\mu}_t - \lambda_{\min}}{\overline{\mu}_t} - \frac{w - w_e}{w} \tag{7}$$

The subsequent action is selected according to an ϵ-greedy policy (Sutton and Barto, 2018), which works by sampling a random number $u \sim [0; 1]$ and choosing action a as in Equation 8:

$$a = \begin{cases} \text{argmax}\big(Q(s,a)\big), & if\ u \leq \epsilon \\ \text{random action}, & if\ u > \epsilon \end{cases}, \tag{8}$$

where ϵ is a parameter which must be set to grant a correct balance between exploration and exploitation.

5. EXPERIMENTAL SETUP

The framework has been tested on two different systems: a Job Shop and a Flow Shop, each composed of five workstations WS and five queues Q. Schematic representations of the systems under analysis are depicted in Fig. 2 and in Fig. 3.

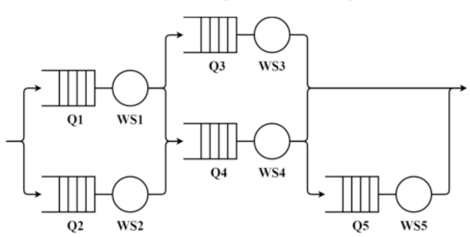

Fig. 2. Schematic representation of the Job Shop

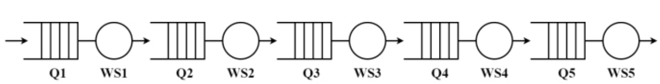

Fig. 3. Schematic representation of the Flow Shop

The routings across the Job Shop are set randomly, but each order has a fixed routing. Jobs share the same routing in the Flow Shop. Both for the Job Shop and for the Flow Shop, two sources of variability are considered. The first one is related to the production mix, as the average processing times are uniformly distributed in the interval [30, 120] seconds for each order and for each workstation. In addition, a less significant amount of variability is directly related to the variance of processing times. In the simulation model they are distributed around their mean value according to a normal distribution, as in (Huang, Wang and Ip, 1998), with standard deviation equal to 5. Other relevant assumptions include no transportation nor setup times, and the fact that new orders are always available in the pre-shop pool and they are released under a FIFO policy. The intelligence layer is implemented in MATLAB® while the DES model at the basis of the DT is built in Simulink using SimEvents toolbox, following the software selection methodology presented in (Fumagalli *et al.*, 2019).

6. RESULTS

A comparison between the proposed order release procedure and standard CONWIP was performed. Tables 1 and 2 compare the WIP and TH of the proposed model against a standard CONWIP approach at different card number levels. Experiments were conducted running 30 simulations lasting twelve hours each for every alternative (rows in the tables). Results reported were measured after the completion of the training of the algorithm, which took ten simulation runs lasting three days of simulated time.Graphical comparisons between the standard CONWIP and the proposed model are shown in Fig. 4 and in Fig. 5 to illustrate the behaviour over time of the latter.

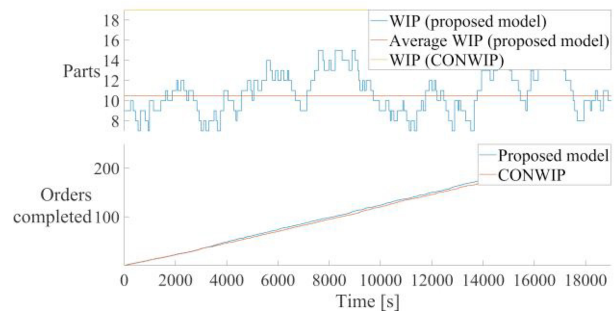

Fig. 4. Comparison in the Job Shop

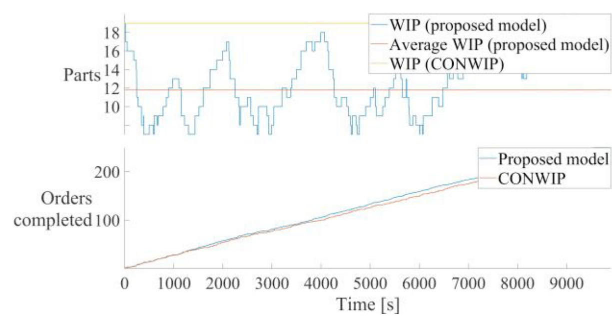

Fig. 5. Comparison in the Flow Shop

Blue and red lines in the lower part of Fig. 4 and of Fig. 5 represent the number of orders completed over time in the same scenario with the proposed order release model and with standard CONWIP respectively. In the upper part of each figure, the blue line shows the WIP level over time while the red one corresponds to its average under the proposed order release model; the yellow line represents WIP level under CONWIP. It is possible to see that both in the Job Shop and in the Flow Shop scenarios the algorithm allows the reduction of WIP level without affecting the production rate. In fact, throughput remains slightly higher for both the analysed systems. Besides, it is clear that the proposed model is continuously varying the WIP level to allow to keep a lower average WIP with respect to standard CONWIP. Fig. 6 and Fig. 7 graphically outline the results, where dots and dashes represent mean values and their confidence intervals. Table 1 and Table 2 report full numerical results.

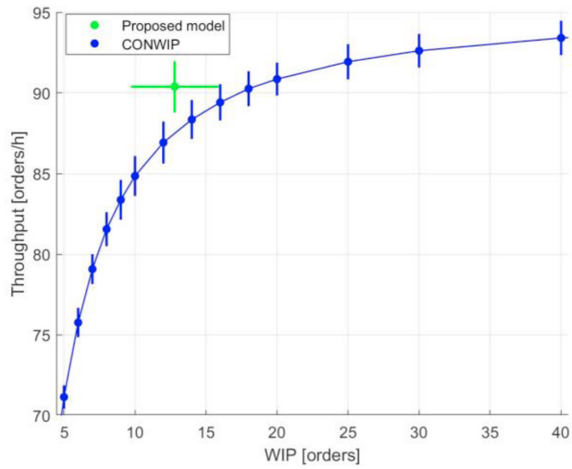

Fig. 6. Job Shop experimental results

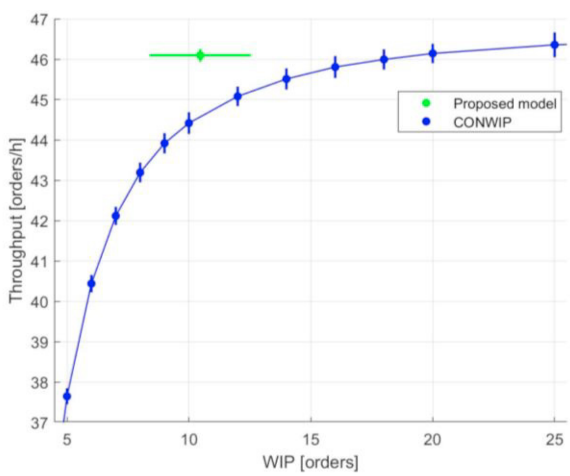

Fig. 7. Flow Shop experimental results

Table 1. Job Shop experimental results

		TH (parts/minute)		WIP (parts)	
		Mean	*Standard deviation*	*Mean*	*Standard deviation*
Proposed model		90.387	1.591	12.791	3.083
CONWIP (n)	*5*	71.137	0.715	5	0
	6	75.760	0.910	6	0
	7	79.073	0.902	7	0
	8	81.549	1.057	8	0
	9	83.370	1.293	9	0
	10	84.847	1.241	10	0
	12	86.923	1.305	12	0
	14	88.352	1.210	14	0
	16	89.413	1.126	16	0
	18	90.262	1.084	18	0
	20	90.857	1.020	20	0
	25	91.934	1.092	25	0
	30	92.616	1.045	30	0
	40	93.410	1.067	40	0
	50	94.052	1.119	50	0

Results in Table 1 show that the performances of the proposed order release procedure in Job Shop configuration are superior to CONWIP. As a matter of fact, it is possible to reach a throughput level which is not more than 5% lower than the expected maximum throughput but allowing to significantly reduce WIP. If compared to a CONWIP setting with a similar throughput, it grants a 45% reduction of the WIP.

Table 2. Flow Shop experimental results

		TH (parts/minute)		WIP (parts)	
		Mean	*Standard deviation*	*Mean*	*Standard deviation*
Proposed model		46.098	0.159	10.469	2.084
CONWIP (n)	*2*	18.529	0.129	2	0
	4	33.007	0.184	4	0
	6	40.439	0.215	6	0
	8	43.193	0.245	8	0
	10	44.417	0.270	10	0
	12	45.080	0.242	12	0
	14	45.511	0.263	14	0
	16	45.807	0.273	16	0
	18	45.995	0.252	18	0
	20	46.143	0.240	20	0
	25	46.357	0.309	25	0
	30	46.451	0.327	30	0
	35	46.490	0.349	35	0
	40	46.513	0.366	40	0
	50	46.511	0.361	50	0

Results from Table 2 show that the proposed model is able to improve the performance of the system also when applied to a simple Flow Shop. In fact, on average, throughput levels similar to the one allowed by the proposed algorithm could be achieved in a CONWIP system with an approximately double WIP level. Moreover, the throughput reduction with respect to its expected maximum is about 1% in this configuration.

7. CONCLUSIONS

This work aims at showing the capabilities of the DT within production control and, in particular, the order release system. Its impacts on the improvement of the performances of the production system under control, are also tested in the planned experiments, both in a flow shop and in a job shop. The DT is developed to drive the production control system by setting the production target and by monitoring production. In this sense this work closes the found gap in terms of proposing a DT-based production control strategy which is able to reduce the WIP while ensuring close-to-maximum throughput levels.

Being this an initial work to investigate a yet unexplored field, more work has still to be done. Research works shall be devoted to the application of similar models to different pull approaches while also considering other performances. Future development should involve the application of different and more efficient RL algorithms, with the purpose to include additional information into the decision-making algorithm such as, for example, workstation utilization levels and queue lengths. In addition, a benchmark activity should be performed to compare the performance of the proposed model with other card controlling procedures. Finally, developing and testing the model in industrial environments could allow to prove the relevance of the Digital Twin in the field of production control for manufacturing companies.

REFERENCES

Dittrich, M. A. and Fohlmeister, S. (2020) 'Cooperative multi-agent system for production control using reinforcement learning', *CIRP Annals*, 69(1), pp. 389–392.

Framinan, J. M., González, P. L. and Ruiz-Usano, R. (2003) 'The CONWIP production control system: Review and research issues', *Production Planning and Control*, 14(3), pp. 255–265.

Fumagalli, L. *et al.* (2019) 'Framework for simulation software selection', *Journal of Simulation*, 13(4), pp. 286–303.

Garetti, M. *et al.* (2016) 'Synchro-push: A new production control paradigm', *Proceedings of the Summer School Francesco Turco*, 13-15-Sept(figure 1), pp. 150–155.

Hopp, W. J. and Roof, M. L. (1998) 'Setting WIP levels with statistical throughput control (STC) in CONWIP production lines', *Int. J. Prod. Res.*, 36(4), pp. 867–882.

Hopp, W. J. and Spearman, M. L. (2004) 'To pull or not to pull: What is the question?', *Manufacturing and Service Operations Management*, 6(2), pp. 133–148.

Hopp, W. J. and Spearman, M. L. (2008) *Factory Physics: foundation of manufacturing management*.

Huang, M., Wang, D. and Ip, W. H. (1998) 'Simulation study of CONWIP for a cold rolling plant', *Int. J. Prod. Econ.*, 54(3), pp. 257–266.

Kritzinger, W. *et al.* (2018) 'Digital Twin in manufacturing: A categorical literature review and classification', *IFAC-PapersOnLine*. Elsevier B.V., 51(11), pp. 1016–1022.

Liu, Q. and Huang, D. (2009) 'Dynamic card number adjusting strategy in card-based production system', *Int. J. Prod. Res.*, 47(21), pp. 6037–6050.

Lödding, H. (2013) *Handbook of Manufacturing Control*, *Handbook of Manufacturing Control*.

Napoleone, A., Macchi, M. and Pozzetti, A. (2020) 'A review on the characteristics of cyber-physical systems for the future smart factories', *J. Manuf. Syst.*. Elsevier, 54(August 2019), pp. 305–335.

Negri, E. *et al.* (2019) 'A Digital Twin-based scheduling framework including Equipment Health Index and Genetic Algorithms', *IFAC-PapersOnLine*, 52(10), pp. 43–48.

Negri, E. *et al.* (2020) 'MES-integrated digital twin frameworks', *J. Manuf. Syst.*. Elsevier, 56(May), pp. 58–71.

Negri, E. *et al.* (2021) 'Field-synchronized Digital Twin framework for production scheduling with uncertainty', *Journal of Intelligent Manufacturing*. Springer US, 32(4), pp. 1207–1228.

Negri, E., Fumagalli, L. and Macchi, M. (2017) 'A Review of the Roles of Digital Twin in CPS-based Production Systems', *Procedia Manufacturing*. 11(June), pp. 939–948.

Renna, P., Magrino, L. and Zaffina, R. (2013) 'Dynamic card control strategy in pull manufacturing systems', *International Journal of Computer Integrated Manufacturing*, 26(9), pp. 881–894.

Spearman, M. L., Woodruff, D. L. and Hopp, W. J. (1990) 'CONWIP: A pull alternative to kanban', *Int. J. Prod. Res.*, 28(5), pp. 879–894.

Sutton, R. S. and Barto, A. G. (2018) *Reinforcement Learning: An Introduction*, *The Lancet*. MIT Press.

Tardif, V. and Maaseidvaag, L. (2001) 'An adaptive approach to controlling kanban systems', *European Journal of Operational Research*, 132(2), pp. 411–424.

Urbina Coronado, P. D. *et al.* (2018) 'Part data integration in the Shop Floor Digital Twin: Mobile and cloud technologies to enable a manufacturing execution system', *J. Manuf. Syst.*. The Society of Manufacturing Engineers, 48, pp. 25–33.

Watkins, C. J. C. H. C. H. and Dayan, P. (1992) 'Technical Note Q-Learning', *Machine Learning*, 8, pp. 279–292.

Zhang, M., Tao, F. and Nee, A. Y. C. C. (2020) 'Digital Twin Enhanced Dynamic Job-Shop Scheduling', *J. Manuf. Syst.*. Elsevier, (April), pp. 0–1.

Zhuang, C., Liu, J. and Xiong, H. (2018) 'Digital twin-based smart production management and control framework for the complex product assembly shop-floor', *Int. J. Adv. Manuf. Technol.*, 96(1–4), pp. 1149–1163.