



A novel throughput control algorithm for semi-heterarchical industry 4.0 architecture

Silvestro Vespoli¹ · Guido Guizzi¹ · Elisa Gebennini² · Andrea Grassi¹

Accepted: 24 June 2021
© The Author(s) 2021

Abstract

Modern market scenarios are imposing a radical change in the production concept, driving companies' attention to customer satisfaction through increased product customization and quick response strategies to maintain competitiveness. At the same time, the growing development of Industry 4.0 technologies made possible the creation of new manufacturing paradigms in which an increased level of autonomy is one of the key concepts to consider. Taking the advantage from the recent development around the semi-heterarchical architecture, this work proposes a first model for the throughput control of a production system managed by such an architecture. A cascade control algorithm is proposed considering work-in-progress (WIP) as the primary control lever for achieving a specific throughput target. It is composed of an optimal control law based on an analytical model of the considered production system, and of a secondary proportional-integral-derivative controller capable of performing an additional control action that addresses the error raised by the theoretical model's. The proposed throughput control algorithm has been tested in different simulated scenarios, and the results showed that the combination of the control actions made it possible to have continuous adjustment of the WIP of the controlled production system, maintaining it at the minimum value required to achieve the requested throughput with nearly zero errors.

Keywords Industry 4.0 · Decentralized production control · Mass customization · Dynamic production control · Optimal control

✉ Silvestro Vespoli
silvestro.vespoli@unina.it

Guido Guizzi
g.guizzi@unina.it

Elisa Gebennini
elisa.gebennini@unimercatorum.it

Andrea Grassi
andrea.grassi@unina.it

¹ Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale, Università degli Studi di Napoli Federico II, Piazzale V. Tecchio 80, 80125 Napoli, Italy

² Faculty of Economics, Universitas Mercatorum, Piazza Mattei 10, 00186 Rome, Italy

1 Introduction

Current and future market scenarios drive companies' attention to customer satisfaction, which in turn moves companies to pursue increased product customization and quick response strategies for maintaining competitiveness. This behaviour has led to a radical change in the approach to production toward the pursuit of customer-needs-based value creation and no longer aiming solely at the reduction of costs. Already in 1989, Davis introduced this trend, referring to it as a production strategy focused on the broad provision of personalized products and services and naming it "mass customization" (MC) (Davis 1989). However, the fulfilment of such a paradigm requires the definition of new production concepts in order to achieve the requested increased level of flexibility, reconfigurability, and performance (Dashchenko 2006; Mourtzis and Doukas 2012; Mourtzis et al. 2012)

Many scientists have identified the fourth industrial revolution, known as Industry 4.0 (I4.0), as a response to these new challenges (Ivanov et al. 2020; Brad et al. 2018; Prinz et al. 2019). In the I4.0 context, information technology (IT) merges with operations technology (OT) to make possible the achievement of the expected performance targets in a highly customized market scenario (Grassi et al. 2020a; Thames and Schaefer 2016). This link is made possible thanks to the main innovations introduced by Industry 4.0, identifiable in the concepts of cyber-physical systems (CPS) and the internet of things (IoT) (Jeon et al. 2020; Liu and Xu 2017). There has been an abundance of development and research related to Industry 4.0 technologies, but to the best of the author's knowledge, there is ongoing research activity into methodologies for efficiently using the data flows and interconnection capabilities provided by CPSs (Allgöwer et al. 2019; Gushev 2020; Oks et al. 2019). The CPS concepts itself is still debated within the literature. However, in this paper, CPS will refer to "engineering systems characterized by the integration of communication, control, and computation within a natural and/or man-made system governed by the laws of physics" (Gushev 2020).

The network of connected machines made possible by the introduction of the CPS concept may have a great impact on industrial production management, subverting consolidated manufacturing and production control (MPC) systems and making possible the implementation of innovative approaches based on the delegation of a decision-making quota to the shop-floor level. This trend involves a gradual shift away from classical centralized production planning and inventory control architectures, such as the MRP/MRP-II systems, to more decentralized ones focused on the scalability of the control system along different managerial levels, each of them characterized by a specific quota of autonomous decision-making. To this end, the literature contains numerous studies where researchers have investigated centralized approaches to tackling scheduling and inventory production control issues and pointing out their limitations (Habib and Chimsom 2019; Monostori et al. 2006; Windt and Jeken 2009).

Traditionally, in a centralized MPC system approach, the aim is to maximize production resource utilization while trying to meet the promised due date. The main problem of this approach is that firms are forced to lengthen the order crossing time by imposing fixed and long lead times between subsequent production stages, thus leading to an increase in stock levels within the production system, which helps, in turn, to compensate for production variability. In this way, each resource works as if it were isolated, theoretically achieving its own maximum production rate. However, the final result is that the production is systematically anticipated, provoking undesired queues and delays in the system (Bendul and Knollmann 2016; Grassi et al. 2020c; Knollmann and Windt 2013).

Moreover, the operational scheduling problem that, in the traditional centralized approaches is the methodology adopted to maximize production resource utilization, assumes a deterministic context, working on a frozen horizon in which known jobs are defined and assumed to be immutable. This results in (i) the dimension of the problem tending to increase as much as it is addressed at higher levels in the manufacturing control system, as a centralized approach implies, and (ii) the intrinsic NP-hard nature of scheduling problems resulting in planned order schedules tending to become unfeasible in the short term due to perturbations occurring at the shop-floor level (i.e. delays in production activities, unavailability of machines and operators, variability of production rates, and so on). This ends in a continuous call to reschedule activity and, in turn, increased unpredictability in the manufacturing system.

There are several examples in the scientific literature of attempts to solve this problem. For instance, some important studies dealt with assembly line sequencing and the optimization of production scheduling in manufacturing processes by requiring set-ups (Kendrick et al. 2017; Lou and Van Ryzin 1989; Minguillon and Lanza 2019; Plaga et al. 2019; Rolf et al. 2020; Sharifnia et al. 1991; Fragapane et al. 2020; Walker and Bright 2013). Other studies, however, have instead used heuristic methods for the scheduling of flexible flow-shop systems, aiming to find the overall optimum system through better utilization of resources (Paternina-Arboleda et al. 2008). Still other researchers have addressed the problem by adopting a control theory based approach (Ivanov et al. 2021; Dolgui et al. 2014; Ivanov and Sokolov 2019; Ivanov et al. 2018).

One of the main pillars of Industry 4.0 is the introduction of more autonomy in the system, paving the way to the adoption of a more decentralized MPC system (Guizzi et al. 2017). In this way, decisions are split among different managerial levels, each covering a confined physical part of the system while also addressing specific functional objectives. Only the lowest level executes precise scheduling activity based on the current detailed knowledge of the system state, while the upper levels control the general performance of the system (e.g. the average throughput) on a longer time scale and based on less detailed information, thus remaining tolerant to variations. Again, some examples are present in the literature (Bendul and Blunck 2019; Gonzalez et al. 2019; Jairo et al. 2019; Moghaddam and Deshmukh 2019). In particular, there is growing interest in hybrid MPC architectures that are neither centralized nor decentralized, in which the complex MPC system is broken down from a functional point of view. Grassi et al. (2020c) recently proposed a semi-heterarchical MPC architecture in which three different functional levels are introduced: (i) knowledge-based enterprise resource planning (KERP), which represents the business level and is also accountable for cloud interaction; (ii) high-level controller (HLC), which is liable for the general performance of a monitored system; and (iii) low-level controller (LLC), which is the operative level of a particular production unit.

By taking advantage of a semi-heterarchical architecture and integrating Industry 4.0 innovation, the presented work contributes to the design of the highest level of such an architecture. Extending the work of Vespoli et al presented at the MIM 2019 conference, the original contribution of this paper consists of the development of a control methodology allowing the HLC to follow performance targets imposed by the upper level (which may be an ERP system or the KERP system of the above-mentioned semi-heterarchical architecture). In particular, the performance target kept under control is the throughput that can be arbitrarily set by the upper level itself on the basis of order arrivals.

The proposed control system, designed to operate in an MC scenario in which the variability of job processing times is high due to the variability of the production mix resulting from

customization requests, showed promising ability to keep the production system controlled to the throughput target value while minimizing the WIP value.

The remainder of the paper is organized as follows: Sect. 2 presents the problem statement, while Sect. 3 describes the proposed high-level control system, discussing its integration in a semi-heterarchical architecture. Section 4 shows the effectiveness of the proposed approach by means of simulated experimental campaigns, and Sect. 5 concludes the paper.

2 Problem statement

Consider a manufacturing system called to produce customized orders. Without loss generality, we consider a flow-shop manufacturing system in which each job moves forward following the same sequence of working stations (i.e., the same production route). Jobs are all different, meaning that even if the technological route is respected, their processing times at stations may differ significantly.

Looking at system performance, we will refer to throughput (TH) as the average output of a production process (machine, workstation, line, plant) per unit time (e.g., parts per hour) and to cycle time (CT) as the average time from release of a job at the beginning of the routing until it reaches an inventory point at the end of the routing (i.e., the time the part spends as WIP) (Spearman et al. 1990). If the system is observed for a sufficiently long time, then Little's law holds:

$$WIP = TH \cdot CT. \quad (1)$$

In an MC context, the complexity and high level of customization in customer orders result in an increased variability of job processing times within the production system (Fogliatto et al. 2012). Hence, to facilitate mathematical modelling while allowing for good representation of the variability of the MC and Industry 4.0 scenarios, we assume that job processing times are taken from an exponential distribution, given its memoryless property and its significant coefficient of variation.

Variability affects performance, so the adopted production control mechanism has to be able to contend with the effects of variability and show robustness in control capabilities. Among the different production control mechanisms produced by both scientists and practitioners in recent years is the constant-work-in-process (CONWIP) control system, based on direct WIP control over a long loop that covers a series of stations. It is a hybrid push/pull technique that limits the total number of jobs in the production system under processing at the same time, keeping the cycle time under better control and increasing the predictability of the system. In the CONWIP, the throughput is an observable performance parameter that can be modelled in mathematical terms and modulated by directly acting on the WIP level authorized in the system.

Hopp and Spearman (2011) studied the dynamics of a CONWIP system and produced mathematical models describing its behaviour in the functions of control variables and boundary conditions. In particular, they investigated the behaviour of the best condition available (no variability and balanced line, with the same performing times at the stations) and the worst possible condition. These scenarios are of interest because they represent the best and worst possible performance that a CONWIP production system may obtain. Moreover, the researchers evaluated the behaviour of the same system working in a practical case (i.e., the practical worst case, or PWC) in which the performing times of jobs at the stations are expo-

Table 1 Basic factory dynamics (Hopp and Spearman 2011)

Performance scenario	Cycle time (CT)	Throughput (TH)
Best case	$CT_{min} = \begin{cases} T_0, & \text{if } w \leq W_0 \\ \frac{w}{r_b}, & \text{Otherwise} \end{cases}$	$TH_{max} = \begin{cases} \frac{w}{T_0}, & \text{if } w \leq W_0 \\ r_b, & \text{Otherwise} \end{cases}$
Worst case	$CT_{max} = w \cdot T_0$	$TH_{min} = \frac{1}{T_0}$
Practical worst-case	$CT_{PWC} = T_0 + \frac{w-1}{r_b}$	$TH_{PWC} = \frac{w}{W_0+w-1} \cdot r_b$

nentially distributed, maintaining the balanced line condition in terms of average processing times. Table 1 shows a summary of the resulting laws, in which:

- T_0 represents the raw processing time of the line (the sum of the average processing times of each workstation at the steady state),
- r_b represents the bottleneck rate of the line at the steady state, and
- W_0 represents the critical WIP of the line (i.e., the WIP level for which the line with a defined T_0 and r_b achieves maximum throughput with a minimum cycle time if no variability exists in the system).

As will be further clarified, the Hopp and Spearman's laws reported in Table 1 will be used as performance benchmarks of a traditional CONWIP flow-shop production line scheduling architecture composed of the HLC and LLC levels, as shown in Fig. 1.

The architecture being considered for the implementation of the dynamic production control developed in the present study, shown in Fig. 1, was derived from the semi-heterarchical architecture proposed by Grassi et al. (2020c). It is based on the three management levels in the production control breakdown from both physical and functional points of view: KERP, HLC, and LLC.

The KERP represents an evolution of the classical ERP system. It is at the top level of the structure, so its role covers the acquisition of customer orders and the negotiation of the delivery date, based on the knowledge of the general performance level at which the production system is operating in terms of TH and CT (i.e., it controls the profitability of the production). Accepted orders are then put in a queue of virtual orders waiting for admission to the lower levels. KERP is no longer liable for detailed production planning; it only defines the orders to be released to production and, most importantly, the target performances to be achieved from the lower HLCs.

The next level, HLC, is a sort of general performance controller of a production sub-system that inherits the performance targets from the upper level (KERP) and monitors the system state in terms of TH and CT , taking control actions as needed to dynamically maintain performance within the expected targets. The HLC then operates to quickly react to state variations of the controlled sub-system and/or to changes in the TH and CT inherited targets to keep actual performance as stable as possible and in line with expectations.

Finally, the LLC consists of a job-ready queue (JRQ), the Dispatcher, and the physical production system. Given the WIP-based control strategy adopted by the HLC, scheduling is addressed here with a dynamic approach, first, by means of a logical unit, the Dispatcher, that decides which order in the JRQ is the best one to be released into the production sub-system every time a completed one exits. Therefore, the LLC is affected by the HLC's decisions and at the same time has the responsibility to establish priorities according to which pending jobs will be processed in such a way as to better compensate for variability, thus improving synchronization in the physical system.

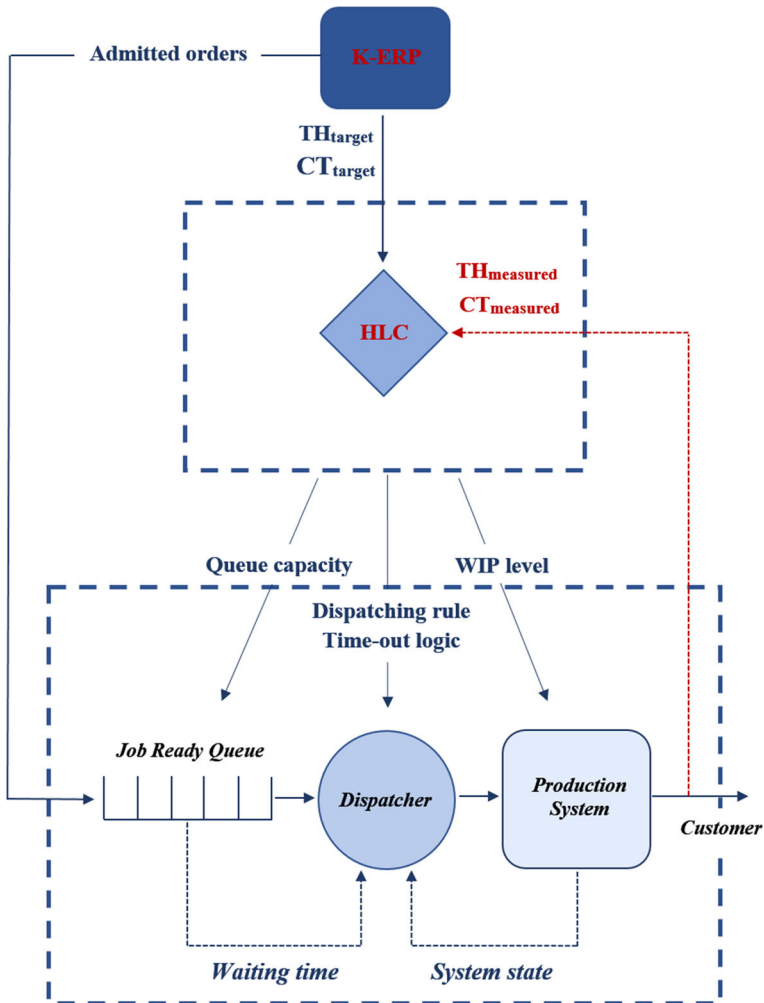


Fig. 1 The semi-hierarchical architecture being considered (adapted from Grassi et al. 2020c)

In this work, we propose a cascade control for the HLC level that is able to keep the monitored throughput in line with the targets inherited from the KERP in a production context characterized by high variability, such as in the I4.0 context (Fig. 2). In order to facilitate understanding of the proposed control scheme, we first briefly detail the effects of the four different control knobs at the HLC level, as already analysed in Grassi et al. (2020b):

1. *JRQ size* This is the number of orders in the JRQ that have been admitted and are waiting for actual release into production. The size of the JRQ influences the effectiveness of the Dispatcher. A small JRQ size would result in a dispatcher having limited choices, while conversely, a large JRQ would leave the Dispatcher a wider range of choices but would also increase the total time required for a job to cross the entire system (Grassi et al. 2020b).

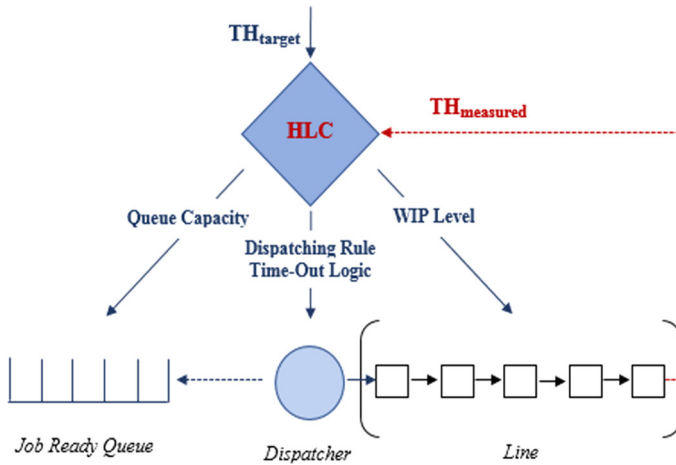


Fig. 2 Detail of the HLC control

2. *System WIP* This refers to the number of jobs allowed in the production system, and it is the most important control knob of the HLC. A high level of WIP better compensates for the variability in the system, allowing for increased utilization of stations, thus making possible the achievement of a higher TH with the cost of an increase of the CT . Conversely, low WIP levels may help to obtain lower CT value at the expense of a loss in TH .
3. *Dispatching rule* This is the rule according to which the Dispatcher, as the heart of the LLC, dynamically updates the priority given to the jobs waiting for processing in the JRQ. In opting for one rule rather than another, the set of information that the Dispatcher considers in making its decisions is determined, along with the re-sequencing mechanism adopted. In the following, we will consider two different dispatching rules for the Dispatcher, as proposed in Vespoli et al. (2019) and deeply analysed in Grassi et al. (2020b), so only briefly discussed here:
 - Open Loop Dispatch Control—chooses the job to be admitted into production without considering feedback from the production system. This decision is made considering only the known characteristics of the job.
 - Closed Loop Dispatch Control—chooses the job to be admitted in production taking into consideration both the characteristics of the job and the state of the monitored production system.
4. *Time-Out* This is the maximum time that a job is allowed to wait in the JRQ. Jobs passing this time are forced to be released by the Dispatcher. This can be a fixed value (statically chosen) or dynamically adjusted by an appropriate computation mechanism, updated on the basis of the contingent situation of the production system (Grassi et al. 2020b).

Therefore, the HLC dynamic control we are proposing is based on the vertical control aspect of the considered semi-heterarchical architecture, where a given HLC has to maintain the performances of the monitored production sub-system on a particular inherited target (in terms of TH and CT) by appropriately varying its own control knobs.

3 The proposed HLC control approach

The HLC represents the top-level production controller, since it is directly connected to the KERP, with which it exchanges information about the desired performance levels to achieve. These performance levels are the expected TH and CT , estimated at the KERP level, with respect to a high-level knowledge of the production system and the queue of orders promised to customers. This work addresses only throughput control, leaving cycle time control to future works.

In order to keep the monitored throughput in line with the target inherited from the KERP level, it should be noted that the most relevant control knob is the allowed WIP value for the production system. The idea of searching for a direct relationship between the WIP and the TH is not new. For instance, Little's law may be used to calculate the WIP value that a system should have in order to maintain a predefined TH level, given the knowledge of its CT . Other methodologies and tools that are able to statically estimate the best WIP value given a particular series of jobs have been proposed and analysed in studies contained in the body of scientific literature (Thürer et al. 2012). However, to the best of the author's knowledge, the proposed algorithm and methodologies approach the problem with a static estimation of the WIP value instead of a dynamic evaluation of the WIP level. This kind of design approach may be useful when the variability of the production system is controlled, but it does not guarantee adequate performance in an MC scenario.

In this study, the objective was to keep the throughput in line with the targets inherited from the KERP in an MC scenario. Thus, a cascade control scheme was developed for the HLC level. This control scheme consists of an optimal control (based on the best-known analytical model of the controlled system) plus a PID controller. The rationale behind this choice is to couple the knowledge of the production system with the most-established industrial control loop algorithm.

Figure 2 presents the detail of the HLC control scheme, pointing to the role of the HLC and the dynamic adjustment activity delegated to it to pursue the TH_{TARGET} . The control system here was developed considering a balanced CONWIP production line in which jobs' processing times taken from an exponential distribution are assumed. Different working conditions can be further considered, provided that general (even approximated) performance analytical models are available. In this paper, the PWC laws will be used in the optimal control algorithm part, given that it provides the exact analytical model that fits with the considered working condition as modelled by Spearman et al. (1990).

Referring to T_0 as the raw process time, which is the sum of the mean effective process times of the stations in the line; r_0 as the bottleneck rate of the line, defined as the production rate of the station with the most utilization; W_0 as the critical WIP level for the line, which is the WIP required for a line with no variability to achieve the maximum throughput (r_b) with minimum cycle time (T_0); and \bar{t} as the given average processing time at stations, we can write

$$\begin{aligned} W_0 &= r_b \cdot T_0 \\ T_0 &= m \cdot \bar{t} \\ r_b &= \frac{1}{\bar{t}} \end{aligned} \quad (2)$$

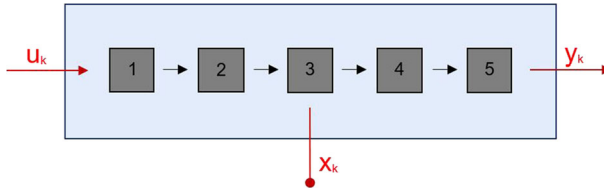


Fig. 3 Production system WIP dynamics

Hence, according to Spearman et al. (1990), the best long-term predictor of the considered production line throughput $TH_{PWC,k}$ is

$$TH_{PWC} = \frac{w}{W_0 + w - 1} r_b \tag{3}$$

where w represents the WIP value set for the production line.

Considering a generic time step k , let us indicate with $TH_{TARGET,k}$ the throughput value imposed by the KERP system. Then, the objective of the control algorithm is to dynamically estimate the WIP value so that the TH of the line is close to the desired one for each considered period. Therefore, the goal is to minimize, over the considered time horizon, the sum of the absolute deviations between the expected and target throughputs, which will also act as a sort of performance indicator:

$$\min \sum_{k=0}^N |TH_{PWC,k} - TH_{TARGET,k}| \tag{4}$$

Figure 3 shows the reference schematics of the considered production line: the output of the control u_k , which represents the input of the production line, sets the number of jobs to be introduced in the production system at the period k . The x_k , instead, is the state variable representing the WIP in the system at the beginning of the period k , while y_k represents the number of jobs leaving the production system at the period k . Obviously, the WIP control action is operated by directly controlling u_k .

Due to the assumed PWC working condition, an estimator of y_k , named \hat{y}_k , can be written as

$$\hat{y}_k = \frac{\Delta T_k}{\bar{t}} \tag{5}$$

where \bar{t} represents the given mean processing time at a station, and ΔT_k the time between k and $k + 1$. Hence, referring to Fig. 3, the WIP dynamics from $k = 0$ to $k = 1$ can be written as

$$x_{k+1} - x_k = u_k - \hat{y}_k \tag{6}$$

Finally, the optimal control problem may be summarized as

$$\begin{cases} \min_{\{u_k\}_{k=0}^N} \sum_{k=0}^N |TH_{PWC,k} - TH_{TARGET,k}| \\ x_{k+1} = x_k + u_k - y_k \\ x(0) = x_0 \end{cases} \tag{7}$$

Trying to graphically understand the behaviour of the proposed control algorithm, let us refer to Fig. 4. When the $TH_{TARGET,k}$ changes, the proposed law provides a forecast of the

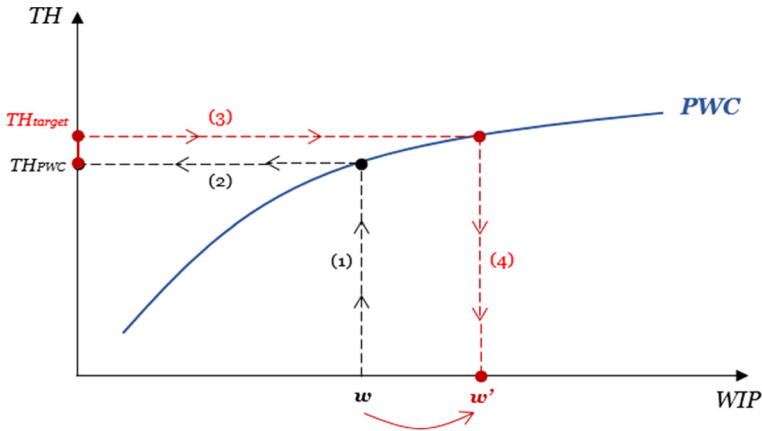


Fig. 4 TH_{PWC} vs TH_{TARGET}

value of u_k such that the WIP in the production system moves from w to w' . To do this, we have to impose that $TH_{PWC,k} = TH_{TARGET,k}$.

Hence, with reference to the Eq. 3, we may write

$$\frac{x_k + u_k - \hat{y}_k}{W_0 + x_k + u_k - \hat{y}_k - 1} r_b = TH_{TARGET,k} \quad (8)$$

Hence, with some mathematical steps, the formulation of the control law can be found:

$$u_k = \hat{y}_k - x_k + \frac{TH_{TARGET,k} \cdot (W_0 - 1)}{r_b - TH_{TARGET,k}} \quad (9)$$

Analysing the control law found in Eq. 9 in the function of the $TH_{TARGET,k}$ (see Fig. 5), it may be observed that the function is discontinuous when $TH_{TARGET,k} = r_b$. This is an expected behaviour because, when the TH_{TARGET} value is close to the r_b value, this translates into a nearly impossible target request for the production system, given that it is asked to achieve the TH value of the best case (refer to Table 1) even if there is variability in the system. Therefore, mathematically, this results in

$$\lim_{TH_{TARGET,k} \rightarrow r_b^-} u_k = +\infty \quad (10)$$

$$\lim_{TH_{TARGET,k} \rightarrow r_b^+} u_k = -\infty \quad (11)$$

Such situations are not practically feasible because it would mean that the HLC should permit an infinite amount of WIP on the line. Obviously, such a $TH_{TARGET,k} > r_b$ is unachievable because a production system cannot perform at a TH higher than its bottleneck. For the sake of completeness, we observe that if $TH_{PWC,k} - TH_{TARGET,k} > 0$, then $u_k < 0$. Since we cannot reduce the WIP by physically removing jobs, in this situation, the control value is set to $u_k = 0$.

In summary, the proposed HLC optimal control law, valid only if $TH_{TARGET} \geq \gamma \cdot r_b$ and schematized in Fig. 6, is

$$u_k = \max \left\{ 0, \hat{y}_k - x_k + \frac{TH_{TARGET,k} \cdot (W_0 - 1)}{r_b - TH_{TARGET,k}} \right\} \quad (12)$$

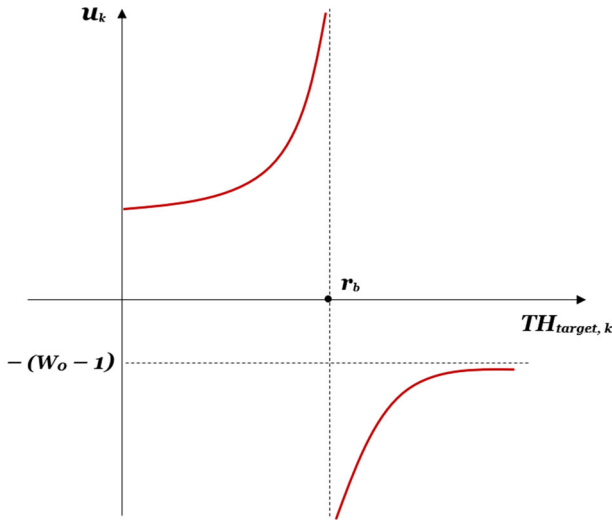
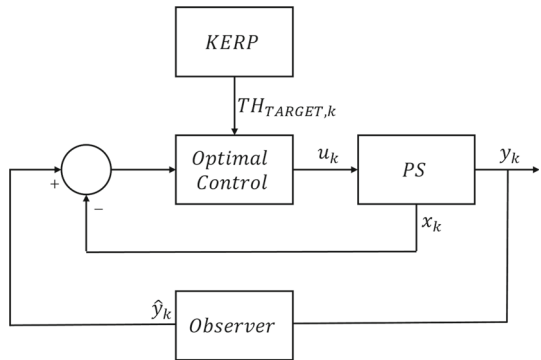


Fig. 5 The u_k (Eq. 9) graphical behaviour

Fig. 6 The optimal control HLC scheme



where $\gamma \in [0, 1]$ is a safety value that limits the maximum TH allowed in the system (i.e., to avoid a TH_{TARGET} equal to r_b).

Therefore, if the functioning of the system falls within the PWC hypothesis (i.e., in the case of a perfectly balanced line with exponential processing time distribution), the proposed law is able to control the system at its best. If, on the contrary, the assumptions of the PWC are not verified, the proposed law may cause the considered system to work in over-capacity conditions (e.g., showing a TH value always greater than the TH_{TARGET} due to less variability in the processes with respect to the exponential variability of the PWC), as well as in under-capacity conditions (i.e., the process variability is greater than the PWC variability). In both scenarios, the defined control system lacks an efficient TH feedback lever. In fact, the looped-back value is the amount of WIP actually taken out of the system (a sort of variation of the state), rather than the actual measured TH value. If the proposed law makes possible a more predictable control of the production system, it will not effectively hit the target TH_{TARGET} value because the PWC condition is no longer valid.

In order to overcome these limitations, a cascade regulator was introduced by adding a PID controller to the optimal control law. The aim is to equip the control system with a feedback

of the recorded TH . Using the feedback of the TH value from the production process, the PID controller compares it with the TH_{TARGET} value; the difference, or so-called *error signal*, is then used to determine the value of the output variable. Hence, the PID evaluates the output according to the following:

- The value of the error signal (proportional action)

$$u_P = K_P \cdot e(t) \quad (13)$$

- The past values of the error signal (integral action)

$$u_I = K_I \cdot \int e(t) dt \quad (14)$$

- How fast the error signal varies (derivative action)

$$u_D = K_D \cdot \frac{de(t)}{dt} \quad (15)$$

where

$$e(t) = TH - TH_{TARGET} \quad (16)$$

Figure 7 shows the final control scheme, where it can be seen that there are two different feedback actions. The first is represented by \hat{y}_k (i.e., the estimated number of completed jobs in a time interval ΔT_k), while the second is the observed TH value. The latter is not directly fed back from the production system because it is not possible to measure it directly. Instead, it must be observed over an interval of time δT that is, in general, significantly greater than ΔT_k because it has to average the effects of variability. The higher the δT considered, the more stable the observed TH will be, but the reaction of the control will be slow.

Therefore, as will be shown in the following section, the choice of the δT value influences the behaviour of the PID control action. In fact, even if the sole PID controller could cope with WIP adjustment needs, its corrective action would start only after the effect of variations in TH occurred. The effect of the control actions, therefore, would have to first propagate to the controlled variable, and given the variability in the considered system, the control action of the PID alone would not be satisfactory. This is why a hybrid solution is proposed where, first, a near optimal WIP value is anticipated through the use of the optimal control (i.e., based on theoretical knowledge of the production line) and, second, it is further corrected through the use of the PID to account for the differences between the theoretical PWC condition and the real one.

4 Experimental tests and results

Due to the complexity of the proposed approach and the number of changes that an implementation in an industrial environment would require, we analysed the proposed dynamic control algorithm in a simulated scenario. To make sure that the built simulation model can consistently simulate the performance of a real system and to ensure that the simulation time chosen is sufficiently long to have consistent results, we conducted a validation phase against the above-mentioned PWC case. The objective of the validation phase was to verify that after a two-year-long simulation of the production system, the recorded performances in terms of TH and CT can be considered statistically representative of the theoretical performance proposed by Spearman et al. (1990), as already reported in Table 1.

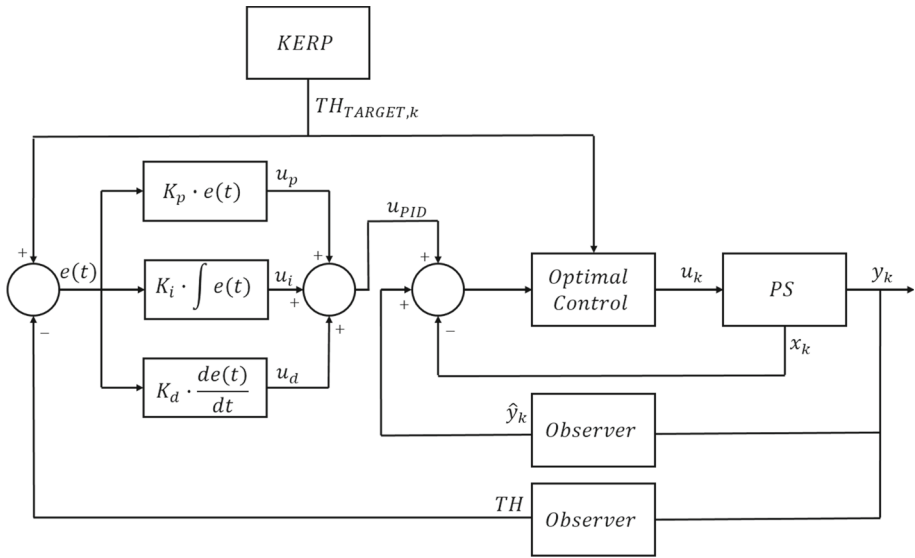


Fig. 7 The proposed HLC control scheme

After the validation phase, the behaviour of the proposed cascade control algorithm in different simulation scenarios was analysed by adjusting the variability of the generated jobs' processing times, thanks to the help of a gamma distribution that included the exponential one as a particular case. Nine different scenarios were analysed where there were variations in the logic involved in the LLC, alternating between three levels (FIFO, open loop, and closed loop), and the variability of the job processing times. The promising results showed that the proposed dynamic approach was able to maintain the production line at a predefined TH_{TARGET} with a limited standard deviation value.

4.1 Experimental methodology

In order to show the effectiveness, behaviour, and dynamics of the proposed control approach, a hybrid simulation tool based on the mutual use of discrete event simulation (DES) and multi-agent simulation in the AnyLogic environment was developed and used as a test rig.

In Fig. 8, the “main” agent of the developed simulation tool is shown. Modelling resources and jobs as agents made possible a wide parameterization of variables, including those related to the physical configuration of the system. In this case, the study focused only on some of them, without preventing the possibility of using the developed tool in future applications from different points of view.

Figure 9 shows a screen capture of the discrete event part of the simulation tool responsible for job generation and flow management of the production system. Specifically, it is composed of the following:

- A source element (agent generator), named “JobGenerator”, able to generate job agents within the line
- A queue element (order queue), named “buffer”, representing the JRQ from which the jobs still not released into production may be chosen by the Dispatcher to enter production



Fig. 8 The developed simulation tool



Fig. 9 Details of the developed simulation tool

- A hold element, named “HLCRelease”, where the control action of the proposed HLC control scheme is actually carried out (i.e., the u_k action). This block limits the flow of the generated jobs, allowing only the predefined u_k job to flow in the system per period
- Two restricted area elements, named “ProductionWIPStart” and “ProductionWIPEnd”, liable for maintaining the system’s WIP, controlled on the basis of the WIP level imposed from the HLC (this control action is complementary to the action of the previous “HLCRelease” element)
- A sink, to destroy “Job” agents just completed
- Various ancillary blocks to monitor line performance for statistical purposes

Coming back to the simulation tool analysis shown in Fig. 8, the Dispatcher is not immediately visible in the scheme. This one, in fact, is a dynamic control algorithm invoked every time a job enters the JRQ and every time a control signal is issued from the HLC. For the same reason, the processing machines are also absent: modelled as agents, they are in the restricted area elements.

Finally, the simulation tool has been enriched with some charts that make it possible to monitor, minute-by-minute, the trend of some indices and performance indicators. At the top left, the real-time TH value, measured by the observer (in blue), and the TH_{TARGET} imposed by the KERP system (in red) are shown. At the top right, the job cycle time, evaluated as the difference between the time a job enters the system and its completion, is displayed. The reported value represents the average value over all the jobs that have been produced in

the same ΔT considered for the evaluation of the TH . In the bottom left corner, the current allowed WIP value for the system is found (i.e., the value of x_k), while in the bottom right corner, the error sent back to PID $e(t) = TH - TH_{TARGET}$ is shown.

Once the simulation model is built, in order to make sure that the model can consistently simulate the performance of a real system and, above all, that the simulation time chosen is sufficient to have consistent results, we conducted a validation phase against the above-mentioned PWC. For this test, we disabled the HLC control action and the Dispatcher logic, focusing the analysis on the TH value (expressed in jobs per hour) according to the PWC assumptions. After the evaluation of the TH for different (fixed) WIP values, a t -Student test was conducted to validate the results obtained by the simulator. The developed simulation tool statistically proved to produce the same TH and CT mean values of the one calculated with the PWC law proposed by Spearman et al. (1990) in a two-year simulation time run, with a confidence level of 95%. More detail about the validation phase can be found in Grassi et al. (2020b), due to the similarities of the methodology and simulation tools involved.

4.2 Results

To better point out the impact of the proposed HLC control action, a scenario in which jobs are always available for processing has been assumed. Hence, the unique constraints for the TH performance is the chosen WIP level in the production system since the JRQ is always full of jobs available for production. The intention is to show the behaviour of the introduced HLC control algorithm when facing a transient state (like the initial one) and a steady state, while tackling the variability of the introduced job.

For every experimental working condition, the jobs are generated with the same characteristic: every job has to be processed on all the machines in the system, with a variable processing time generated from a gamma distribution with a fixed mean of 10 minutes. As known, the gamma distribution is a two-parameter family of continuous probability distributions, which includes the exponential one as a special case. In this paper, we will refer to its parameterization with a shape parameter α and an inverse scale parameter β , called the rate parameter. With that formulation, the mean is defined as $\mu = \frac{\alpha}{\beta}$, and, as noted, it was considered fixed at 10 minutes.

In order to explore the capabilities of the proposed HLC control, the shape parameter of the distribution was varied to simulate scenarios differing from the PWC case. In fact, choosing values of the shape parameter $\alpha < 1$, scenarios with more variability than the exponential one were obtained; with a value $\alpha = 1$, the particular case of the exponential distribution was represented; and with values of $\alpha > 1$, scenarios with reduced variability were addressed. Clearly, when the α parameter changed, the β parameter was re-computed to ensure the processing time mean value remained unchanged.

Nine different scenarios varying the logic involved in the LLC among three levels (FIFO, open loop, and closed loop) and the variability of the job processing times by varying the α factor form of the gamma distribution were chosen, as shown in Table 2. For what concerns the PID controller parameters tuning, the Ziegler–Nichols method (Ziegler and Nichols 1942), among the most used and appreciated for its simplicity, was adopted in each scenario. This choice does not exclude the use of more sophisticated PID tuning algorithms, and the authors left the possible use of other tuning methods to future research and evaluation.

Table 3 shows the results of the simulation campaign. In the first and second columns, the factors that defined the different experimental scenarios are found, while the following three columns show the estimated PID parameter values. The next two columns report the value

Table 2 Experimental factors and levels

Experimental factor	Level	Measure unit
Mean processing time	10	[min]
JRQ dimension	10	[jobs]
TH_{TARGET}	$0.6 \cdot r_b = 3.6$	[jobs/hour]
Dispatching logic	FIFO-Open Loop-Closed Loop	[logic]
Factor form α	0.75–1–5	[]

Table 3 Simulation results

Dispatching logic	α	K_P	K_I	K_D	$STD[e(t)]$	$E[WIP]$
FIFO	0.75	1.2	0.006	60	0.57	8.04
Open Loop	0.75	1.2	0.006	30	0.50	7.22
Closed Loop	0.75	1.2	0.006	10	0.55	6.88
FIFO	1	0.6	0.003	10	0.41	6.22
Open Loop	1	0.6	0.003	10	0.35	5.85
Closed Loop	1	0.6	0.003	10	0.39	6.08
FIFO	5	0.3	0.002	5	0.10	3.87
Open Loop	5	0.3	0.002	5	0.09	3.84
Closed Loop	5	0.3	0.002	5	0.09	3.93

of the standard deviation $STD[e(t)]$ of the PID error $e(t)$ (here identified as a performance index) and the average WIP value recorded, respectively. The average value of the error is not shown as it was maintained at zero by the proposed controller in all the considered scenarios.

Analysing the results of Table 3, we may note that the best tuning gain values for the PID depended substantially on the α form factor of the processing times distribution, while they remained substantially independent from the different dispatching logic adopted. The only exception is shown in the high variability scenarios ($\alpha = 0.75$), where a higher K_D value is preferred. In particular, the K_D value is higher for the FIFO logic than the other dispatching logic scenarios. This can be explained by the behaviour of the derivative action that tried to quickly compensate for variations of the error signal, due to its speed of change, without waiting for the error to become significant (proportional action) or persist for some time (integral action). In fact, the Dispatcher is conceived so as to do just that to compensate for the error through a more careful choice of the next job to be admitted into the system, while the PID tries to compensate for this variability by introducing more WIP into the system. In this regard, the average WIP value needed by the system to achieve the TH_{TARGET} value in the case of $\alpha = 0.75$ for the FIFO rule was 8, while through the use of the improved rules, it was possible to achieve the same throughput with about 12% less WIP. As variability increases, it was found that adopting improved rules at the dispatcher level allowed the system to work with even less WIP than the FIFO scenario, especially when a TH_{TARGET} close to r_b was set.

Finally, with regard to the TH control, the control architecture showed good effectiveness in the simulated scenarios, with the average error between TH_{TARGET} and TH kept at zero, while also keeping a reduced standard deviation of the actual TH value around the target.

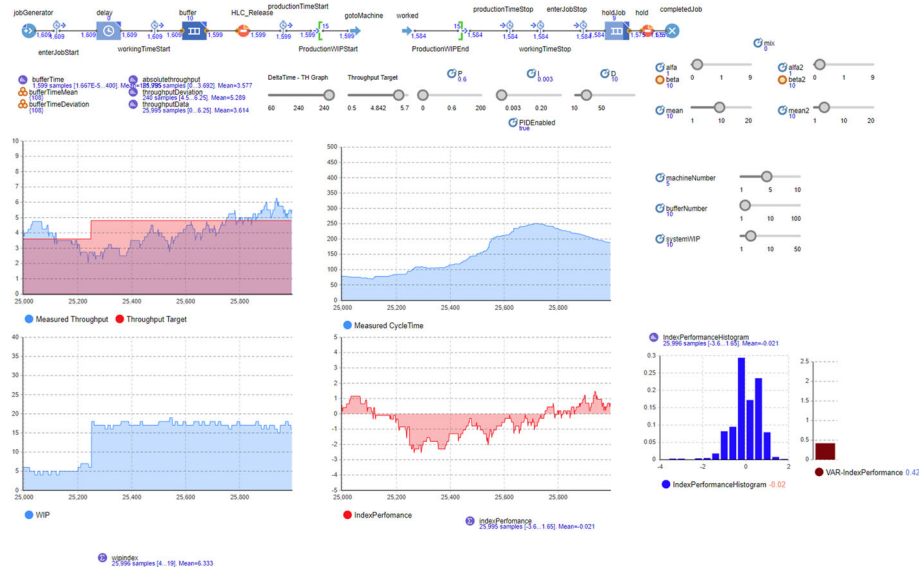


Fig. 10 The control action response to a TH change

The registered standard deviation was implicitly due to the intrinsic variability in the system as a consequence of the variable processing times.

Next, considering the scenario with $\alpha = 1$ and with the dispatching logic set to open loop, we wanted to evaluate how the considered system responded to a TH_{TARGET} variation. In particular, we proposed a change from the previously considered value of $TH_{TARGET} = 0.6 \cdot r_b = 3.6$ [job/h] to a $TH_{TARGET} = 0.8 \cdot r_b = 4.8$ [job/h]. The action expected to bring the system to the new TH_{TARGET} value as quickly as possible is a sharp change in WIP. In Fig. 10, it is possible to see the variations of WIP imposed by the controller after the change of TH_{TARGET} was imposed. The red line in the first chart reports the variation of TH_{TARGET} from 3.6 to 4.8, while the graph at the bottom shows the consequent change of the WIP imposed by the controller. The first control action was rapidly taken by the optimal control, projecting the change in TH_{TARGET} and quickly proposing a higher WIP value. After some minutes, the PID completed the refinement of the value, maintaining a continuous compensation for the variability of the introduced jobs. The time to reach the new steady state was evaluated to be around 400 [min].

We also analysed the behaviour of the proposed control architecture when a change in the variability of the stochastic process determining station processing times was imposed, while maintaining the same TH_{TARGET} . Here, we considered a shift in the production mix where the station performing times changed from a gamma distribution with $\alpha = 1$ to one with $\alpha = 5$, generating a situation with less variability. As can be seen in Fig. 11, as a consequence of the reduction in variability, the production system initially started to over-perform, given that the current level of WIP was related to the previous variability condition. In this particular case, the optimal control does not provide any change to the controlled system due to the fact that its modelling is here independent of process variability. Instead, the PID controller, receiving a feedback error signal, started its control adjustment action by progressively reducing the WIP in the system until the TH was brought to the target and the error zeroed. In this case, the system responded more slowly than with the TH_{TARGET}



Fig. 11 The control action response to a TH change

change previously simulated. The reason is that, here, the optimal control as conceived was not able to predict the TH variation on the basis of the characteristics of the jobs entering the system, so only the PID part of the controller intervened once the error was recognized. In addition, because the TH signal is delayed by a time that, on average, is the cycle time of the system, the controller responds with a systemic delay. In other words, the controller recognized the change in the entry variability after it propagated along the whole system, thus producing an effective variation of the TH . In Fig. 11, it can be seen that the system needed a time of 800 [min], twice the value observed in the case where a variation in the TH_{TARGET} was imposed, to find the new stable condition.

In general, the control system proved to be effective in maintaining the system TH at the requested TH_{TARGET} even when there were changes in the entry variability (i.e., non-controllable variations coming from the market). Thank to this, the production system is brought by the control actions to always work in its best possible condition, with the requested throughput TH_{TARGET} at the minimum possible WIP, translating to the minimum possible cycle time or, in other words, with the maximum possible responsiveness.

5 Conclusion

The ability to effectively allocate production resources is critical for maintaining competitiveness in modern market scenarios characterized by customized and dynamic demand. The technologies introduced in Industry 4.0 enable data exchange among production system entities, allowing the implementation of new approaches to production management and control. As stated in the introduction, the semi-heterarchical architecture was chosen as one of the most promising for the Industry 4.0 context among the various possible MPC architectures.

This paper proposed a first throughput control algorithm to be implemented in the HLC of the Grassi et al. (2020c)'s semi-heterarchical architecture. This control consists of a cascade algorithm comprising first an optimal control regulator and a PID controller that are able to perform a second control action, with the aim of providing feedback about the achieved throughput. The results showed that the combined control actions made it possible to have a continuous adjustment of the WIP for the controlled production system, maintaining it at the minimum WIP required to achieve the requested TH , with a nearly zero error.

This work can be considered a first step toward the development of the MPC semi-heterarchical architecture, which differs significantly from the traditional centralized scheduling and inventory production control systems currently used by practitioners. This architecture moves us closer to the long-awaited decentralization of Industry 4.0. Because of the promising results of this study, future research may concentrate on more complex production systems, such as the job-shop system. It would be especially advisable to begin with the development of analytical models capable of predicting the performance of such a system, as well as the ability to forecast the effects involved by changes in processing time variability.

Funding Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allgöwer, F., Borges de Sousa, J., Kapinski, J., Mosterman, P., Oehlerking, J., Panciatici, P., et al. (2019). Position paper on the challenges posed by modern applications to cyber-physical systems theory. *Nonlinear Analysis: Hybrid Systems*, 34, 147–165. <https://doi.org/10.1016/j.nahs.2019.05.007>.
- Bendul, J., & Knollmann, M. (2016). The lead time syndrome of manufacturing control: Comparison of two independent research approaches. *Procedia CIRP*, 41, 81–86. <https://doi.org/10.1016/j.procir.2015.08.104>.
- Bendul, J. C., & Blunck, H. (2019). The design space of production planning and control for industry 4.0. *Computers in Industry*, 105, 260–272. <https://doi.org/10.1016/j.compind.2018.10.010>.
- Brad, S., Murar, M., & Brad, E. (2018). Design of smart connected manufacturing resources to enable changeability, reconfigurability and total-cost-of-ownership models in the factory-of-the-future. *International Journal of Production Research*, 56(6), 2269–2291. <https://doi.org/10.1080/00207543.2017.1400705>.
- Dashchenko, A. (2006). *Reconfigurable manufacturing systems and transformable factories, first edit edn*. Springer, Berlin. <https://doi.org/10.1007/3-540-29397-3>
- Davis, S. M. (1989). From “future perfect” : Mass customizing. *Planning Review*, 17(2), 16–21. <https://doi.org/10.1108/eb054249>.
- Dolgui, A., Kovalev, S., Kovalyov, M., Nossack, J., & Pesch, E. (2014). Minimizing setup costs in a transfer line design problem with sequential operation processing. *International Journal of Production Economics*, 151, 186–194. <https://doi.org/10.1016/j.ijpe.2013.10.013>.
- Fogliatto, F., Da Silveira, G., & Borenstein, D. (2012). The mass customization decade: An updated review of the literature. *International Journal of Production Economics*, 138(1), 14–25. <https://doi.org/10.1016/j.ijpe.2012.03.002>.

- Fragapane, G., Ivanov, D., Peron, M., Sgarbossa, F., Strandhagen, J. (2020). Increasing flexibility and productivity in industry 4.0 production networks with autonomous mobile robots and smart intralogistics. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-020-03526-7>
- Gonzalez, S. R., Zambrano, G. M., & Mondragon, I. F. (2019). Semi-heterarchical architecture to AGV adjustable autonomy within FMSs. *IFAC-PapersOnLine*, 52(10), 7–12. <https://doi.org/10.1016/j.ifacol.2019.10.003>.
- Grassi, A., Guizzi, G., Santillo, L., & Vespoli, S. (2020a). The manufacturing planning and control system: A journey towards the new perspectives in industry 4.0 architectures. *International Series in Operations Research and Management Science*, 289, 193–216. https://doi.org/10.1007/978-3-030-43177-8_10.
- Grassi, A., Guizzi, G., Santillo, L. C., Vespoli, S. (2020b). Assessing the performances of a novel decentralised scheduling approach in Industry 4.0 and cloud manufacturing contexts. *International Journal of Production Research*, 0(0):1–20. <https://doi.org/10.1080/00207543.2020.1799105>.
- Grassi, A., Guizzi, G., Santillo, L. C., & Vespoli, S. (2020c). A semi-heterarchical production control architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 24, 43–46. <https://doi.org/10.1016/j.mfglet.2020.03.007>.
- Guizzi, G., Vespoli, S., Santini, S. (2017). On the architecture scheduling problem of Industry 4.0. In *CEUR workshop proceedings*, vol. 2010.
- Gushev, M. (2020). Dew computing architecture for cyber-physical systems and IoT. *Internet of Things*, 11, 100186. <https://doi.org/10.1016/j.iot.2020.100186>.
- Habib, M., & Chimsom, C. (2019). Industry 4.0: Sustainability and design principles. In *Proceedings of the 2019 20th international conference on research and education in mechatronics, REM 2019*, <https://doi.org/10.1109/REM.2019.8744120>.
- Hopp, W. J., & Spearman, M. L. (2011). *Factory physics*. Waveland Press, Inc.
- Ivanov, D., & Sokolov, B. (2019). Simultaneous structural-operational control of supply chain dynamics and resilience. *Annals of Operations Research*, 283(1–2), 1191–1210. <https://doi.org/10.1007/s10479-019-03231-0>.
- Ivanov, D., Sethi, S., Dolgui, A., & Sokolov, B. (2018). A survey on control theory applications to operational systems, supply chain management, and industry 4.0. *Annual Reviews in Control*, 46, 134–147. <https://doi.org/10.1016/j.arcontrol.2018.10.014>.
- Ivanov, D., Tang, C., Dolgui, A., Battini, D., & Das, A. (2020). Researchers' perspectives on industry 4.0: Multi-disciplinary analysis and opportunities for operations management. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2020.1798035>.
- Ivanov, D., Sokolov, B., Chen, W., Dolgui, A., Werner, F., & Potryasaev, S. (2021). A control approach to scheduling flexibly configurable jobs with dynamic structural-logical constraints. *IIEE Transactions*, 53(1), 21–38. <https://doi.org/10.1080/24725854.2020.1739787>.
- Jairo, R., Jimenez, J. F., & Zambrano-Rey, G. (2019). Directive mode for the semi-heterarchical control architecture of a flexible manufacturing system. *IFAC-PapersOnLine*, 52(10), 19–24. <https://doi.org/10.1016/j.ifacol.2019.10.013>.
- Jeon, B., Yoon, J. S., Um, J., & Suh, S. H. (2020). The architecture development of industry 4.0 compliant smart machine tool system (smts). *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-020-01539-4>.
- Kendrick, B. A., Dhokia, V., & Newman, S. T. (2017). Strategies to realize decentralized manufacture through hybrid manufacturing platforms. *Robotics and Computer-Integrated Manufacturing*, 43, 68–78. <https://doi.org/10.1016/j.rcim.2015.11.007>.
- Knollmann, M., & Windt, K. (2013). Control-theoretic analysis of the Lead Time Syndrome and its impact on the logistic target achievement. *Procedia CIRP*, 7, 97–102. <https://doi.org/10.1016/j.procir.2013.05.017>.
- Liu, C., & Xu, X. (2017). Cyber-physical machine tool: The era of machine tool 4.0. *Procedia CIRP*, 63, 70–75. <https://doi.org/10.1016/j.procir.2017.03.078>.
- Lou, S., & Van Ryzin, G. (1989). Optimal control rules for scheduling job shops. *Annals of Operations Research*, 17(1), 233–248. <https://doi.org/10.1007/BF02096607>.
- Minguillon, F. E., & Lanza, G. (2019). Coupling of centralized and decentralized scheduling for robust production in agile production systems. *Procedia CIRP*, 79(i), 385–390. <https://doi.org/10.1016/j.procir.2019.02.099>.
- Moghaddam, M., & Deshmukh, A. (2019). Resilience of cyber-physical manufacturing control systems. *Manufacturing Letters*, 20, 40–44. <https://doi.org/10.1016/j.mfglet.2019.05.002>.
- Monostori, L., Váncza, J., & Kumara, S. (2006). Agent-based systems for manufacturing. *CIRP Annals - Manufacturing Technology*, 55(2), 697–720. <https://doi.org/10.1016/j.cirp.2006.10.004>.
- Mourtzis, D., & Doukas, M. (2012). Decentralized manufacturing systems review: Challenges and outlook. *Logistics Research*, 5(3–4), 113–121. <https://doi.org/10.1007/s12159-012-0085-x>.

- Mourtzis, D., Doukas, M., & Psarommatis, F. (2012). A multi-criteria evaluation of centralized and decentralized production networks in a highly customer-driven environment. *CIRP Annals - Manufacturing Technology*, *https://doi.org/10.1016/j.cirp.2012.03.035*.
- Oks, S. J., Jalowski, M., Fritzsche, A., & Möslin, K. M. (2019). Cyber-physical modeling and simulation: A reference architecture for designing demonstrators for industrial cyber-physical systems. *Procedia CIRP*, *84*, 257–264. <https://doi.org/10.1016/j.procir.2019.04.239>.
- Paternina-Arboleda, C., Montoya-Torres, J., Acero-Dominguez, M., & Herrera-Hernandez, M. (2008). Scheduling jobs on a k-stage flexible flow-shop. *Annals of Operations Research*, *164*(1), 29–40. <https://doi.org/10.1007/s10479-007-0257-2>.
- Plaga, S., Wiedermann, N., Anton, S. D., Tatschner, S., Schotten, H., & Neue, T. (2019). Securing future decentralised industrial IoT infrastructures: Challenges and free open source solutions. *Future Generation Computer Systems*, *93*, 596–608. <https://doi.org/10.1016/j.future.2018.11.008>.
- Prinz, F., Schoeffler, M., Lechler, A., & Verl, A. (2019). A novel i4.0-enabled engineering method and its evaluation. *International Journal of Advanced Manufacturing Technology*, *102*(5–8), 2245–2263. <https://doi.org/10.1007/s00170-019-03382-1>.
- Rolf, B., Reggelin, T., Nahhas, A., Lang, S., & Müller, M. (2020). Assigning dispatching rules using a genetic algorithm to solve a hybrid flow shop scheduling problem. *Procedia Manufacturing*, *42*(2019), 442–449. <https://doi.org/10.1016/j.promfg.2020.02.051>.
- Sharifnia, A., Caramanis, M., & Gershwin, S. (1991). Dynamic setup scheduling and flow control in manufacturing systems. *Discrete Event Dynamic Systems: Theory and Applications*, *1*(2), 149–175. <https://doi.org/10.1007/BF01805561>.
- Spearman, M., Woodruff, D., & Hopp, W. (1990). Conwip: A pull alternative to kanban. *International Journal of Production Research*, *28*(5), 879–894. <https://doi.org/10.1080/00207549008942761>.
- Thames, L., & Schaefer, D. (2016). Software-defined cloud manufacturing for industry 4.0. *Procedia CIRP* *52*:12–17, <https://doi.org/10.1016/j.procir.2016.07.041>, the Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016)
- Thürer, M., Stevenson, M., Silva, C., Land, M. J., & Fredendall, L. D. (2012). Workload control and order release: A lean solution for make-to-order companies. *Production and Operations Management*, *21*(5), 939–953. <https://doi.org/10.1111/j.1937-5956.2011.01307.x>.
- Vespoli, S., Grassi, A., Guizzi, G., & Santillo, L. C. (2019). Evaluating the advantages of a novel decentralised scheduling approach in the industry 4.0 and cloud manufacturing era. *IFAC-PapersOnLine*, *52*(13), 2170–2176. <https://doi.org/10.1016/j.ifacol.2019.11.527>.
- Walker, A., & Bright, G. (2013). Stabilisation and control of configurable product manufacturing through biased decision feedback decoupling. *Journal of Manufacturing Systems*, *32*(1), 271. <https://doi.org/10.1016/j.jmsy.2012.11.002>.
- Windt, K., & Jeken, O. (2009). Allocation flexibility: A new flexibility type as an enabler for autonomous control in production logistics. In *Proceedings of the 42th CIRP conference on manufacturing systems*.
- Ziegler, J. G., & Nichols, N. B. (1942). Optimum settings for automatic controllers. *Transactions of the ASME*, *64*(11), 759–768. <https://doi.org/10.1115/1.2899060>.