On the modelling of a decentralized production control system in the Industry 4.0 environment

Andrea Grassi, Guido Guizzi,
Liberatina Carmela Santillo, Silvestro Vespoli*

Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale (DICMaPI)
Università degli Studi di Napoli Federico II, Napoli, NA, 80125 Italy
*Corresponding Author (e-mail: silvestro.vespoli@unina.it)

Abstract: The paper deals with a decentralized production control in an Industry 4.0 environment. In such a kind of systems, the capability to deliver a high level of product customization together with reduced response time is crucial to maintain competitiveness and to increase profit. A semi-heterarchical architecture, formed by three levels, in which the first is responsible for meeting business objectives, the second to maintain target system general performances, and the third to tackle operative scheduling problems, is first discussed as a framework for the future implementation in an Industry 4.0 environment. Successively, the problem to model the system form a dynamic point of view is addressed directly at the second architectural level. This paper, in particular, contributes to the semi-heterarchical architecture development, by proposing a first mathematical model of the shop-floor of a such a system, involving the use of the population dynamic modelling. Finally, the results of the first implementation in a simulated environment are reported. Copyright © 2019 IFAC

Keywords: Industry 4.0, industrial production systems, industrial control, decentralized control, CONWIP.

1. INTRODUCTION

The growing globalization process, the continuous evolution of the competitive industrial scenario, the new challenges transversally crossing innovation, quality, costs and time-to-market, represent opportunities that a modern company cannot lose. The ability to satisfy increasingly personalized market demand in a short time and with limited costs can be considered a fundamental principle for the competitive revival of industrialized countries against the emerging countries characterized by lower technological development but with lower social and labour costs (Panetto, Iung, Ivanov, Weichhart, & Wang, 2019, Yao & Lin, 2016). A change in the production concept is therefore needed, with the aim to create value by pursuing the needs of customers timely and no longer pointing at the mere cost reduction, that is a typical objective of low industrialized countries. In this context, it is essential to acquire the ability to effectively allocate the available resources, as well as the ability to revisit and revolutionize the methods and the controlling approaches of the production systems, to respond to new market criticalities appropriately, thus maintaining the competitive position and ensuring effectiveness and efficiency at the same time (Fogliatto, Da Silveira, & Borenstein, 2012).

The impact of a production paradigm focused on the enhanced product customization and on the shortening of time-to-market is so important to justify the creation of a new industrial paradigm: the Fourth Industrial Revolution (or Industry 4.0). In this regard, while the previous industrial revolutions were characterized mainly by significant advances in technology, this one pursues logistical managerial objectives, linked to the new need of product customization and the consequent flexibility required by a production plant (Oesterreich & Teuteberg, 2016). Industry 4.0, therefore, aims to solve the long-standing contradiction between the individuality of on order production and the savings made through economies of scale. Like every previous revolution, Industry 4.0 is also characterized by the introduction of new enabling technologies: above all, the Cyber-Physical Systems (CPS) and the Internet of Things (IoT) (Hermann, Pentek, & Otto, 2016; Riedl, Zipper, Meier, & Diedrich, 2014).

Regarding Industry 4.0 technologies, a lot has been developed. However, to the best of the author knowledge, there is still an open gap in the literature about the methodologies for efficiently using the data flow provided by the Industry 4.0 CPSs. As mentioned above, Industry 4.0 is intended to make it possible the transition from a classic production paradigm (Mass Production) to a customization-oriented one (Mass Customization). This transition must be supported, however, also by a shift in the logic of the Manufacturing Planning and Control (MPC) system. The main problem of a Mass Customization market scenario is the variability entering the system, generated by the customization requests of the customers. And the problem is that this variability, at the moment, can be faced only with a classical MPC system, such as Manufacturing Resource Planning (MRP-II, in the following simple MRP), which has a strongly hierarchical and centralized structure (ANSI/ISA95) (Guizzi, Vespoli, & Santini, 2017; Moeuf, Pellerin, Lamouri, Tamayo-Giraldo, & Barbaray, 2018). The problems involved by the existing MRP System, characterized by centralized scheduling and inventory production control system has been investigated in the literature, pointing out several limits (Bendul & Blunck, 2019).
Therefore, in this industrial context, it is advisable to gradually shift from a centralized MPC system, like the MRP, to a decentralized one focused on the scalability of the control system along with different management levels, each of them with a different quota of autonomous decision-making.

Numerous control strategies implying autonomous and independent control concepts have been proposed in the scientific literature. Among these, Dolgui et al. introduced a new research branch based on the implementation of the classical Control Theory in the scheduling and inventory production control system (Dolgui, 2017; Dolgui, Ivanov, Sethi, & Sokolov, 2019; Ivanov, Dolgui, Sokolov, Werner, & Ivanova, 2016; Sokolov, Dolgui, & Ivanov, 2018). In particular, Sokolov et al. (2018) analyzed the advantages and limits of different control computational methods and algorithm for the solution of short-term scheduling in an optimal manner developing a first optimal control algorithm.

Therefore, we may consider that the hierarchical structure has been widely studied, showing its potential and, above all, its limits when the system complexity increase. Similarly, strongly decentralized architectures, although not yet fully operational, show all their limits due to the degree of complexity to be transferred to each entity. This leads to a situation in which entities can only chase local optimizations, thus trying to solve complex problems through a simple parcelling of the decision-making. And this is not a feasible way to cope with the complexity of a manufacturing MPC system (Bendul & Blunck, 2019).

The recent growing interests about the hybrid MPC system architecture showed that realistic solutions for an Industry 4.0 scenario are the intermediate architecture, i.e. oligarchic and semi-heterarchical ones (Grassi, Guizzi, Santillo, & Vespoli, 2020). These types of architectures face a complex system not by dividing it into several smaller problems, but by functionality and decision-making skills.

This paper contributes to the semi-heterarchical architecture development, by proposing a first mathematical model of the shop-floor of a such a system. Starting from the work of Vespoli et al. (2019) here is proposed an analytical approach to the High-Level Controller (HLC), required for the future study on its controllability.

2. A SEMI-HETERARCHICAL MANUFACTURING PLANNING AND CONTROL SYSTEM ARCHITECTURE

The research about the possible MPC architecture is not a new topic in manufacturing. Already in 1996, Duffie & Prabhu (1996) proposed a completely heterarchical approach for the management of highly distributed production systems, anticipating the ideas and basic principles of control architectures. In their work, they also proposed a first complete taxonomy of possible management architectures, distinguishing among four possible types: Hierarchical, Oligarchical, Semi-Heterarchical and Heterarchical. As a highlighted from them, an excessive degree of decision-making autonomy of the single entities of the system does not allow the achievement of satisfactory global system objectives (Philipp, Böse, & Windt, 2006). For these reasons, we considered a Semi-Heterarchical architecture to develop a new MPC system. This structure is neither specifically hierarchical nor heterarchical but mixed, allowing the MPC system to manage high customization production systems, ensuring tighter flow times, through the introduction of Pull-like control logic in the 4.0 environment.

This MPC system may be considered as a derivation of Vertical Integration Hybrid Systems (Cochran & Kim, 1998) in which we should have at least two functional levels of management control: a higher one with a medium/long-term production planning (such as modern ERP computerized systems), operating with a Push logic; and a second level which effectively schedules the production, operating with a Pull logic.

Extending this concept, Grassi et al. (2020), proposed an MPC system architecture that consists of three different management level (or control levels), as shown in (Figure 1):

- The first level that may be represented by a classical ERP that receives the order from the market and decide about their acceptance. It holds the coordination activities of the plant, maintaining the products cycle time under control (i.e. it controls the profitability of the production), while it is no longer liable for the solution of the scheduling problem. Monitoring production performances (in terms of Throughout (TH) and Cycle Time (CT)) and the orders in progress at the lower levels, it defines the subsequent orders to be released in production and the target performances to be followed.

![Fig. 1 - The proposed Semi-Heterarchical MPC system control architecture (inspired by Grassi et al. (2020))](image-url)
• The High-Level Controller (HLC) level inherits the target performances (in terms of TH and CT) and the orders to be admitted in production from the upper level. This level has a significant detail of the system, to be able to represent a logical entity to control the system itself: practically, it can be viewed as a CPS in communication with the upper level and the other HLCs of the plant. The HLC has different objective compared with the upper level: it has to control and maintain the inherited performances. To this extent, it implements two separate, not alternative, control approach:
  - A horizontal one, in which the HLCs cooperates to enhance the workload balancing among the lines (e.g., exchanging orders received);
  - A vertical one, in which the HLC controls the parameters of the lower level (i.e. set the WIP and the number of job in the Ready Queue), dynamically changing the WIP level of the production system (without establishing the production order of the job).

In this level the advantages of a semi-heterarchical architecture arise: the controller inherits the performance target and the jobs to be produced from the upper level, but it has the opportunity to cooperate with similar production lines of the firm, without involving the decision of the upper level, establishing a horizontal front of optimization among the controllers of the same level.

• The Low-Level Controller (LLC), previously proposed and analyzed by (Vespoli, Grassi, Guizzi, & Santillo, 2019), represents the physical part of the shop floor level of the production system. This level has a more detailed knowledge of the system, and it is the level in which, for the first time, the scheduling problem is tackled.

3. THE HIGH-LEVEL CONTROLLER (HLC) DYNAMIC MODEL

If with the introduction of Low-Level Controller (LLC) an increase in TH has been obtained, both with a decrease in CT and average WIP within the line, as analyzed by Vespoli et al. (2019), here, we want to keep the focus on the implementation of the vertical aspect of the introduced High-Level Controller (HLC). So, while, on the one hand, the LLC tries to balance the variability introduced into the system dynamically choosing the job to be admitted into production, on the other hand, the HLC have to control the system to a certain value of Throughput (TH) and Cycle Time (CT).

Thus, in order to allow this control action, it is necessary to identify the dynamic model of the production system. From the analysis of Fig. 1 is clear that the main HLC control knob is the system’s WIP level and, more specifically, the number of jobs admitted to production in a given moment.

Therefore, taking advantage of the CONWIP experience proposed by Hopp & Spearman (2011), we are proposing to switch from the CONstant WIP concept to the CONtrolled WIP one, in which the WIP changes dynamically, depending on the TH and CT performance target imposed from the upper level.

Consider now a manufacturing system called to produce customized orders. Without loss of generality, a system in which orders (jobs) have to be processed with the same technological sequence is considered, which is, a classical flow-shop system. Every job is different from the others, meaning that, even if the technological cycle is respected, the consistency and the specifics of each operation may change, resulting in different processing times.

We assume that the flow shop system is forced to work with the imposed WIP level set by the HLC based on its own proper rules. In order to find a first dynamic model of the system, we assume that the HLC is called to operate only with the WIP level, which represents one of its control knobs (Figure 2).

![Fig. 2 – The High-Level Controller Framework (Grassi et al., 2020)](image)

This hypothesis will let the production line working as in the CONWIP case, whose behaviour is well known in the literature (Hopp & Spearman, 2011) from a mathematical point of view, and it may be used as a starting point for the modelling process of the introduced CONtrolled WIP concept.

For the CONWIP case, the Best condition available (no variability and balanced line) is known, as well as the Worst possible condition. Moreover, the behaviour of such a controlled production line working in a practical case, in which the processing times of jobs are exponentially distributed is also known.

In [Table 1] a summary of the derived laws is reported.

Following is the meaning of the parameters:

- $T_0$ represents the Raw Processing Time of the line (the sum of long-term average process time of each workstation);
• \( r_b \) represents the Bottleneck Rate of the line (it is the rate of the workstation that have the highest long-term utilization);

• \( W_0 \) represents the Critical WIP of the line (it is the WIP level for which a line, with a defined Raw Processing Time and Bottleneck Rate, achieve the maximum throughput and the minimum cycle time without any variability).

It is now essential to identify the variables representing the state of the system under consideration. In this study, we propose to use as the system state both the WIP level and the number of jobs that exit the system in a certain \( \Delta t \) of time. This latter is representative of the throughput, being it a measure of the output within the considered \( \Delta t \) period.

Table 1 - Basic Factory Dynamics (Hopp & Spearman, 2011)

<table>
<thead>
<tr>
<th>Performance Scenario</th>
<th>Cycle Time (CT)</th>
<th>Throughput (TH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Case</td>
<td>( C_{T_{\text{min}}} = \frac{t_0}{w} ) if ( w \leq W_0 ) ( \frac{w}{r_b} ) otherwise</td>
<td>( T_{\text{H}_{\text{max}}} = \frac{w}{r_b} ) if ( w \leq W_0 ) ( \frac{1}{r_b} ) otherwise</td>
</tr>
<tr>
<td>Worst Case</td>
<td>( C_{T_{\text{max}}} = wT_b )</td>
<td>( T_{\text{H}_{\text{min}}} = \frac{1}{r_b} )</td>
</tr>
<tr>
<td>Practical Worst-Case</td>
<td>( C_{T_{\text{PW}}} = T_0 + \frac{w - 1}{r_b} )</td>
<td>( T_{\text{H}_{\text{PW}}} = w \frac{w}{W_0 + w - 1/r_b} )</td>
</tr>
</tbody>
</table>

The objective is to model the behaviour of such a system and, hence, to describe the known dynamic model to implement in the Observer block of Figure 2.

As highlighted before, in a production flow system like the one hypothesized, the only control lever - once the jobs have entered the production system - is represented by the WIP Level or, more properly, the quantity of "extra" jobs to be admitted into the production system at \( k \) period. To understand it, imagine the Production System as a "box" (Figure 3) within which the WIP is able to exhibit its dynamic evolution. At the generic period \( k \) are defined:

- \( x_k \): as the system WIP at the beginning of the period \( k \). Within the production system, the jobs may be distributed in a certain way among the workstations and the respective queues;
- \( u_k \): as the control input, that is, the number of additional jobs allowed to enter the Production System for the period \( k \);
- \( y_k \): as the output, which represents the number of jobs leaving the Production System at the period \( k \), as they underwent with all the process phases.

The WIP dynamics of the production system considered is like a "mass balance": its variation is represented by the difference between the number of jobs admitted in the system minus the number of them that are completed and are exiting from the system.

Then, we propose to assume \( x_2 = y_k \) as the second state variable of the system, directly related to the first derivative of the WIP. As a matter of fact, the variation of \( x_2 \) depends on \( x_1 \) (that is, the level of WIP) and on the parameters characteristic of the system (bottleneck rate, raw process time, the variability of processing times, etc.) that, as we will show in the following, it can be derived, in a first approximation, from the equation of the Practical-Worst Case in Table 1.

Hence, the variation of \( x_2 \) can be expressed by a function \( f(t, x(t)) \) having the following proprieties:

- \( \lim_{x_1 \to 0} f(t, x_1(t)) = 0 \), meaning that when the WIP goes to 0, the output is forced to go to 0 as well;
- \( \lim_{x_1 \to +\infty} f(t, x_1(t)) = \frac{r_b}{\beta} \), meaning that, also with an infinite WIP in the system, the maximum production rate is the bottleneck of the line, representing the best-case performance.

As understandable from the consideration above, the \( f(t, x(t)) \) equation has strong behaviour similarities with the function involved in the population dynamics modelling. In particular, it is of interest to analyze the behaviour of single-species continuous-time population models, like the Logistic Growth Function, the Gompertz Growth Equation, Richards Growth Equation and the Bertalanffy Growth Equation (Marsi Libelli, 2016).

Between the mentioned population dynamics models, the Gompertz Growth Equation showed the best behaviour when compared with the acquired data from a simulation environment. The reason is to be found in its ability to take into account the ageing of the population through reduced reproductive capacity over time. In a flow-shop this effect is justified by the saturation on the bottleneck station, resulting in a declining rate of growth over time.

The Gompertz model in its differential form is

\[
\frac{dx}{dt} = r_b \cdot x \cdot (\ln K - \ln x)
\]

which can be solved analytically to yield

\[
x(t) = K \cdot e^{(e^{-b+rt})}
\]

resulting in a decreasing growth rate that varies in time as

\[
\dot{x}(t) = x(t) \cdot \left( \frac{r_b}{b+rt} - b \right)
\]
As it is possible to see, in the Gompertez model we found two new parameters to be estimated: the growth rate \( r_g \) and the carrying capacity \( K \). The latter plays a major role as it represents the maximum density that the population can reach in that environment. In our case, it represents the maximum throughput that the system can reach in the considered flow-shop system. To this extent, it should depend on the assumption made in the system, in terms of processing time mean and variability, number of WIP in the system (i.e., the \( x_1 \) value) and a number of the workstation within it.

In the case of a system which the processing times of jobs are exponentially distributed with the same mean processing times (that are the Practical Worst-Case hypothesis), it is easy to show that the \( K \) value is equal to the steady-state TH given from the equation in Table 1:

\[
K = \frac{x_1}{W_0 + x_1 - 1} r_b
\]

It only remains to establish the value of \( r \), whose modelling is more complex. In fact, it is inversely proportional to wide parameters involved, such as the job processing time mean, the length of the system (in term of workstation number), the number of WIPs within the system (i.e., the \( x_1 \) value), and the \( \Delta t \) of time considered for the estimation window for the TH. The latter, in particular, change the measure of the dynamic involved as it increases, the system shows more stable TH value, but with less reactivity to the occurrence.

Given this, it is possible to build a first version of the model:

\[
x(t) = f(t,x(t),u(t)) = \begin{pmatrix} -x_2(t) + u(t) \\ r_g(x_1) x_2 (\ln K(x_1) - \ln x_2) \end{pmatrix}
\]

Then, the final model is represented by:

\[
\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} -x_2(t) \\ r_g(x_1) x_2 (\ln K - \ln x_1) \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u(t)
\]

4. VALIDATION EXPERIMENT

In order to validate the proposed model, a simulation tool was developed and used as a test rig within Anylogic environment. In particular, we considered the case of a Flow-Shop system with a fixed number of workstation (five, in the simulated scenario). Hence, we carried out the experiment varying the value of the WIP in the system (leaving it fixed during the simulation run), the processing time means of the job (exponentially distributed) and the value of \( \Delta t \) windows time for the estimation of the TH of the system.

In Figure 4, the result window of the simulation tool is shown. It is possible to observe in the left, the graph of the Absolute CT (calculated as a mean value between all the registered CT from the job) and the Moving Average CT (calculated as a mean value between all the registered CT from the job within the \( \Delta t \) window). On the right, the same graph is evaluated for the TH.

Thanks to the simulation tool, a wide set of data have been collected and then used for estimating the parameter of the proposed model. In particular, in Figure 5 and Figure 6, it is possible to see the run of the Parameter Estimation tool from the Simulink suite, that validates the assumption made on the model. In particular, it is shown that for all the conducted experiment the \( K \)-value is often the same as the TH in Practical Worst-Case, while the \( r_g \) value decreases accordingly to the increase of the WIP in the system, the mean processing time and the considered \( \Delta t \) as supposed.
The proposed model is a first example of dynamic modelling for the flow-shop production system and may represent a good starting point for the future research. It showed to provide accurate estimation of the system state, also during the critical transition phase of the Throughput of the Production System. However, its study must be depended, in order to find the dependencies of the $r_2$ growth rate in function of the system parameter. Additionally, it was validated only in a scenario with a fixed WIP value, showing promising result but, it would be of interest, for future studies, to understand how it effort to dynamic WIP variation.

REFERENCES


