Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

A DIGITAL TWIN FOR PRODUCTION CONTROL BASED ON REMAINING CYCLE TIME PREDICTION

Giovanni Lugaresi

Dep. of Mechanical Engineering
KU Leuven
Celestijnenlaan 300
3001 Leuven
BELGIUM

Pedro Luís Bacelar Dos Santos
Alex Chalissery Lona
Monica Rossi
Andrea Matta

Dep. of Mechanical Engineering
Politecnico di Milano
Via La Masa, 1
20156, Milano
ITALY

Eduardo Zancul

Dep. of Production Engineering
University of São Paulo
Av. Prof. Luciano Gualberto, 1380
Butantã, São Paulo
BRAZIL

ABSTRACT

The recent industrial context pushed manufacturers to invest heavily in digitization for a more efficient use of their equipment and scarce resources. The digitization of industrial environments allows the establishment of digital decision-support tools such as digital twins, to exploit the shop-floor data for making more accurate decisions considering the real system state. Existing literature focuses on the development of specific digital twin components as well as methods that are typically developed and tested without an integration within a digital twin architecture. This paper proposes a complete digital twin framework with the purpose of aiding production planning and control operations. The focus is on the design of a production control service that manages the material flow in the real system using simulation-based predictions of the remaining cycle time. Preliminary experiments are done by applying the digital twin architecture on a lab-scale model, demonstrating the applicability of the proposed approach.

1 INTRODUCTION

Recently, production facilities across the globe have been investing consistently in the digitization of their equipment. New technologies such as the internet-of-things, cyber-physical systems, and artificial intelligence are being applied to optimize the performance of manufacturing systems (Zhang et al.). Among the tools and methods for decision support, the digital twin (DT) stood out as one of the most promising. Within the context of production planning and control, a DT can be defined as *"a virtual representation of a production system that is able to run on different simulation disciplines that is characterized by the synchronization between the virtual and the real system [...]"* (Negri et al. 2017). For discrete manufacturing systems and within the scope of production planning and control, the digital model can be represented by discrete event simulation. The capability of a digital twin of being continuously aligned with its physical counterpart has been frequently highlighted as essential for its correct functioning (Sakr et al. 2021). The utility of the real-time alignment is evident when the DT is used to make predictions based on the current system state and provide useful feedback in the form of improvement actions (Hyre et al. 2022). These processes are useful services that the DT provides to its physical counterpart, while the internal operations such as synchronization and model validation are essential for the appropriate functioning of a valid DT (Lugaresi et al. 2022). In the literature, the importance of production control services that provide a dynamic scheduling capabilities in a production system has been frequently underlined (Xu and Xie 2021). Several methods have been proposed, for instance using machine learning for scheduling policies (Leng et al. 2022). However, the contributions that study the integration of simulation-based services in a fully integrated DT framework are scarce. This paper presents the development of a complete DT capable of making production control decisions exploiting the results of real-time predictions on the system performance. The work consists in designing a digital twin architecture with proper services to guarantee a production control service, in which real-time streams of data from the shop floor feed simulation experiments, and the results are exploited online to provide feedback to the shop floor. This work also tests the architecture on a lab-scale physical system that reproduces the main dynamics of real production systems. The remainder of this paper is organized as follows. Section 2 contains a short state of the art on the key elements included in a DT for production planning and control. Section 3 introduces the digital twin framework taken as reference in this work. Section 4 describes the methodology used for the development of the production control service. Section 5 details the experiments that have been done to test the developed architecture. Section 6 concludes the paper with final remarks.

2 STATE OF THE ART

In recent years, there has been an increasing interest in DT frameworks, proven by the significant number of publications on this topic. However, no common agreement has been achieved until now about a generic architecture. For example, Hyre et al. (2022) recommended an agnostic framework that starts with the virtual representation of the system, then replicating it into digital objects and testing the sub-components individually. Similarly, Leng et al. (2022) emphasize that the flexibility of digital twins depends on the reconfigurability and dynamic scheduling performance. According to Papacharalampopoulos et al. (2021), the framework of a DT should have modeling, diagnostic, and prognostic functions so that the DT can be adaptive with real-time optimization. Sakr et al. (2021) states that a digital twin includes a digital shadow that acquires real-time data from the physical system to feed a digital replica, which is used for operating the rest of the DT functions.

Physical-to-digital alignment is one of the most cited research challenges and regards the importance to keep the model updated in real-time according to changes that occur in the physical system (Lugaresi et al. 2021). Sakr et al. (2021) mentions that usage of static models in evolving systems would require manual re-configurations. Traoré (2021) discusses two types of synchronization: (1) event based, which synchronizes the model at the occurrence of each event, and (2) time based, which applies synchronizations according to a fixed frequency. Tan and Matta (2022) introduce the problem of synchronization as a trade off between computation effort and improvement in simulation model accuracy. Both state-dependent and state-independent policies are defined to decide the temporal allocation of synchronization actions. In order to maintain a digital twin aligned with his physical counterpart, synchronization is not enough, it is also necessary to validate the model. Lugaresi et al. (2022) developed an online validation procedure for digital twins based on Dynamic Time Warping to quantify the difference between the digital and physical systems based on sequences of events.

Once the proper functioning of a digital twin is guaranteed by its alignment with the physical system, it is possible to exploit it to provide useful services for the production. Zhang et al. () suggests that the main DT functions are to analyze production performance indicators, optimize production processes dynamically and give feedback in real-time. The ISO 23247 (Shao 2021) defines the scheduling and routing as one of the most important application of the DT. This is true when taking into consideration the positive aspects of a dynamic decision making in real-time, such as the optimization of cycle time and asset utilization (Sivakumar 1999). According to Li and Yu (2017), an optimized scheduling also implies achieving a better performance, throughput and movement of work pieces.

One way to accomplish the DT benefits is to design an automated decision-making flow based on predictions of the Remaining Cycle Time (RCT) of a part in production. According to van Dongen et al. (2008), the RCT can be defined as *"the amount of time left to finish a cycle operation for an entity from a given stage of the system"*. Estimating the RCT is not trivial and depends on the changing dynamic of the system. For example, imagine a customer waiting in queue at a counter. If a new counter is opened in parallel, the customer will likely experience a lower waiting time (i.e., lower RCT). On the other hand, if the server at the counter takes a break and stops serving customers momentarily, the waiting time increases (i.e., higher RCT). This unpredictable and complex dynamics of a system requires a dedicated system to continuously update the predictions (van Dongen et al. 2008). In the available literature, several methods were implemented to improve the RCT predictions. In the recent years, applications using machine learning models have gained relevance. For instance, Tirkel (2011) implemented a machine learning model for the prediction of average cycle time for a production lot. The paper applies methods based both on decision trees and neural networks, considering that the latter consistently show higher accuracy than the former. One way of estimating the RCT of a entity is by looking at its average cycle time and deducting the time that the entity has already spent in production (van Dongen et al. 2008). However, this approach does not include an update of the actual RCT in real-time. Hence, the authors proposed different types of non-parametric regression based on multiple estimators extracted from a real log. Similarly, Yang et al. (2022) implemented a DT framework capable of calculating the RCT using different ML models. Choueiri et al. (2020) employed a combination of a transition system and linear regression to develop a hybrid model aimed at predicting RCT in the manufacturing process. Friederich et al. (2023) applied a discipline of process mining called *predictive process monitoring* which relies on regression methods based on event data to predict the RCT.

Despite effective in their own domains of application, the aforementioned approaches have been mostly applied as standalone methods and within controlled conditions. To the best of the authors' knowledge, no approaches have included the RCT within a data-driven DT framework. This work aims to exploit the simulation-based RCT predictions as basis to complete a DT framework as production control service.

3 DIGITAL TWIN FRAMEWORK

The framework developed in this work is based on the one suggested by Lugaresi et al. (2022), which allows for the integration of the essential components of a DT such as synchronization and validation. Figure 1 outlines the components and the data flows of the DT that supports real-time data acquisition and control of the physical system. Data is continuously collected through sensor devices in the manufacturing system. The manufacturing execution system allows for the data collection and stores both raw and aggregated data in a database. In the *Data Layer*, specific services guarantee the physical-to-digital alignment of the DT by verifying the characteristics of a digital model. The database is a central element of this layer, and stores data both from the real system and from its virtual model. The synchronization component uses data to gather the current system state, namely the allocation of work-pieces along the system (e.g., buffer levels) and updates the system configuration in the digital model. The synchronization component implements an adapted version of the methods proposed by Passarin, Edoardo and Verucchi, Francesco (2022). Validation uses data to verify if the digital model correctly represents (1) the topological features of the manufacturing system (e.g., physical layout, the material flow connections), and (2) the stochastic behaviour of the system (i.e., processing time distributions). The validation component implements the procedures presented by Lugaresi et al. (2022). If the digital model is not valid, the validation component updates it with the most recent parameters. Each time the digital model is updated, it is also saved in a model store repository, which allows to keep track of all the models that are used within the DT. Finally, the

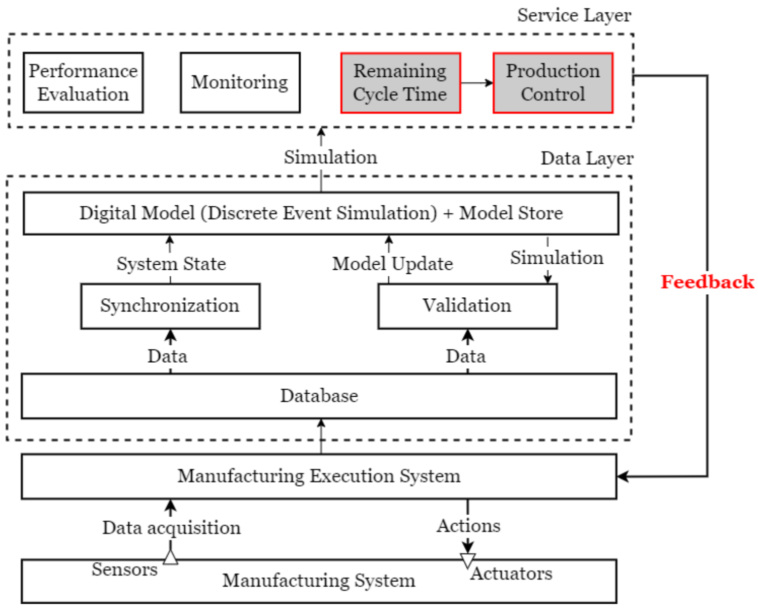


Figure 1: Digital twin framework used as reference in this work.

Service Layer includes components that use the most updated digital model to generate useful services for the production system. Multiple services can reside independently in the DT architecture (e.g., monitoring, performance evaluation). The services provide feedback to the manufacturing execution system, which can convert them into actions on the physical system. In general, each service works at a specified frequency that depends on the physical system and the production planning and control problem under study. The synchronization, validation and control frequencies are denoted as f_s, f_v and f_c, respectively. Section 4 describes the development of the production control service.

4 PRODUCTION CONTROL SERVICE

This section illustrates the methodology used in the production control service within the DT architecture.

4.1 Remaining Cycle Time Definition

Let us consider a generic production system that has to produce $|\mathbb{J}|$ parts. The RCT of a part $j \in \mathbb{J}$ at time t is defined as the expected amount of time required to terminate its production. We denote with T_j the time at which the part j exits the system. Under ideal conditions, T_j is known and the RCT becomes a linear function of time. In real conditions, T_j depends on specific conditions of the system and its environment (e.g., variable processing times, chosen production paths). The physical system conditions at time t can be represented compactly by a vector $\mathbf{x}(t)$ and the exit time can be expressed as a function $T_j(\mathbf{x}(t))$. The RCT of the j-th part at time t can be expressed as a function of of the expected exit time, as in Equation (1).

$$RCT_j(\mathbf{x}(t)) = E\left[T_j(\mathbf{x}(t))\right] - t. \tag{1}$$

4.2 Production Control Service

The proposed production control service exploits the prediction of the RCT to manage a dispatching policy of parts in a manufacturing systems. Without loss of generality, let us refer to a manufacturing system with $|\mathbb{M}|$ machines, connected one another by conveyors. A subset of the machines are identical and alternative in a production recipe, hence constituting parallel routes. The subset $\mathbb{M}_B \subseteq \mathbb{M}$ collects all the machines

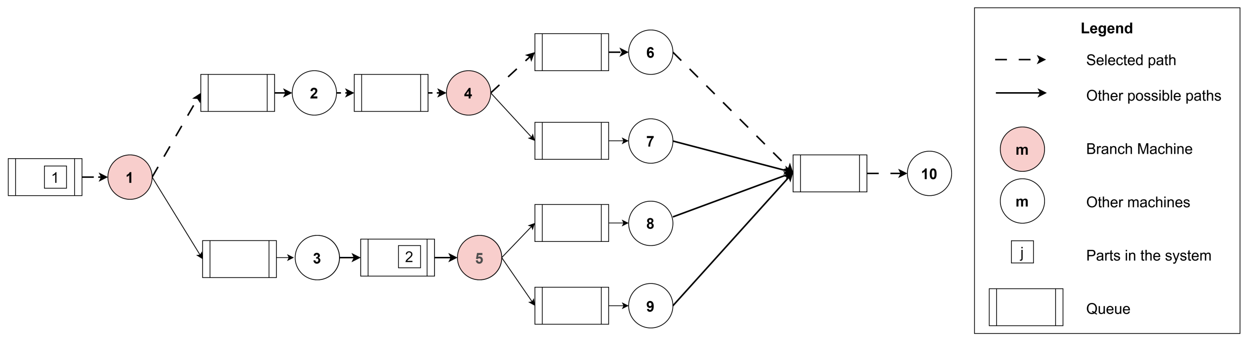


Figure 2: Example of feasible paths for a part at a branching point at $t = 0$.

that are followed by a parallel route. At each moment in time t, a subset of the parts $\mathbb{J}_B(t)$ represents the parts that are located at a branching point machine $m \in \mathbb{M}_B$. Moreover, each part in the system can follow a path ϕ_{jt}. The path indicates the sequence of machines that must be visited in the remaining portion of the production of part j, namely in the time window $[t, T_j]$. For instance, Figure 2 illustrates a system with 3 branching point machines (i.e., $\mathbb{M}_B = \{1,4,5\}$), and the first part $j = 1$ at time $t = 0$ has to follow the path $\phi_{10} = (1,2,4,6,10)$. The production paths can be updated along the production. The aim of the proposed DT-based approach is to find the fastest path possible for each part at anytime. This corresponds to verifying that the remaining production path of each part is still optimal whenever there is the possibility to change it, which corresponds to the moment a part is waiting at a branching point machine. The DT runs simulation experiments to predict the RCT for each possible path. In each experiment, the simulation model is initialized in order to take into account the last known real system state. Based on the set of RCT predictions, the path with the smallest expected remaining time is chosen for the specific part in consideration. This decision is then implemented in the physical system through a feedback action.

The production control procedure is shown in the Figure 3. The following steps are deployed:

- **Gather real system state.** The DT obtains the real system state, and observes which parts are queuing at branching point machines $m \in \mathbb{M}_B$ and updates the set $\mathbb{J}_B(t)$. Figure 2 shows an application example of possible paths. At $t = 0$, there are two parts at branching points ($\mathbb{J}_B = \{1,2\}$).
- **Scenario generation.** Each part $j \in \mathbb{J}_B(t)$ is assigned a combination of feasible paths $\phi \in \mathbb{P}_{jt}$. The paths are generated as a combination without repetition of the possible sequences of machines that a part can visit in the remaining portion of the system. In this work, the method to generate the paths has been adapted from the Depth First Search method (Santhosh and Sastry 2023). For example, in the situation depicted in Figure 2, the paths are $\mathbb{P}_{10} = \{(2,4,6,10);(2,4,7,10);(3,5,8,10);(3,5,9,10)\}$ and $\mathbb{P}_{20} = \{(8,10);(9,10)\}$.
- **Simulation.** For each possible path $\phi \in \mathbb{P}_{jt}$, a simulation experiment is executed. The simulation predicts the RCT for each path, which in the remainder is indicated as $RCT(\phi_{jt})$. Considering Figure 2, $|\mathbb{P}_{10}| = 4$, hence four simulation experiments will be performed for part $j = 1$, and $|\mathbb{P}_{20}| = 2$ experiments are executed for $j = 2$.
- **Solution check.** An analysis is executed within the results of the previous step to check if it is worthy to direct the j-th part in any of the analysed paths. This is done by comparing the predicted RCT with a default value ($\tilde{RCT}_j$), which is calculated considering a round robin policy, i.e., an alternating policy without an intelligent allocation. For each path $\phi \in \mathbb{P}_{jt}$, the gain is calculated using Equation 2 and compared with a defined threshold Δ. Specifically, if $\delta(\phi_{jt}) \geq \Delta$, the path is chosen to be implemented.

$$\delta(\phi_{jt}) = 1 - \frac{RCT(\phi_{jt})}{\tilde{RCT}(\phi_{jt})}. \tag{2}$$

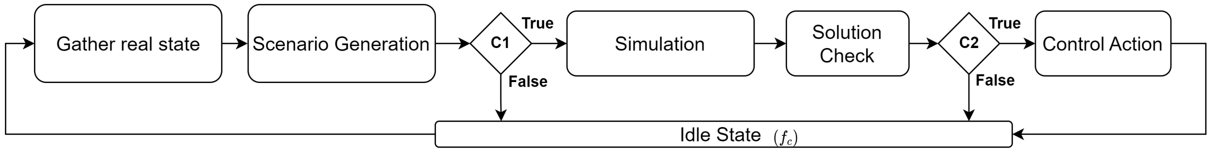
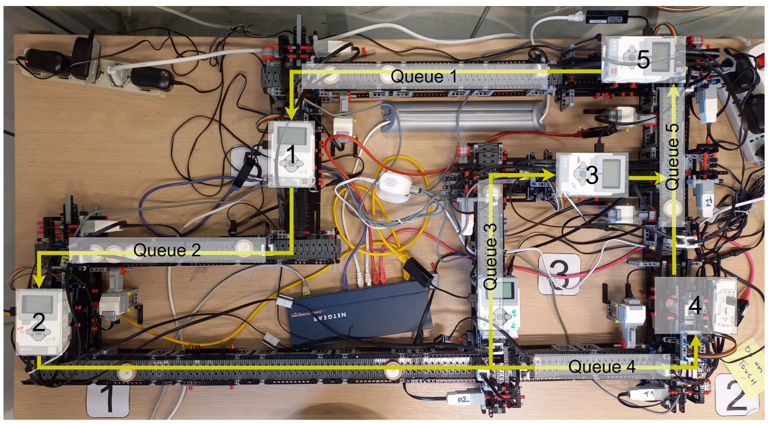
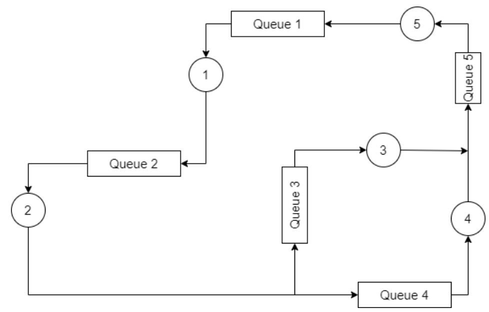


Figure 3: Information flow of the production control service based on RCT predictions (C1: at least one part queuing at any branching point machine; C2: the solution advantage is proven).



(a) Physical system



(b) Logical layout

Figure 4: Five-machine lab-scale model used for testing the proposed digital twin architecture.

- **Control action.** The chosen path is implemented in the system through a feedback message. The message sets the decision to be taken for each specific part to the right branch machine.
- **Idle state.** This is a state which is reached either after the completion of the service or in case any of the following conditions is not satisfied: C1 is true in case of $|\mathbb{J}_B(t)| > 0$, i.e., in the system there is at least one part queuing at any branching point machine $m \in \mathbb{M}_B$; C2 is true if $\delta(\phi_{jt}) \geq \Delta$. In this state, the control service waits for the next call based on the frequency f_c.

5 DIGITAL TWIN PROTOTYPE TESTING

This section explains the experiments done to investigate the applicability of the developed DT architecture.

5.1 Manufacturing System and Digital Twin Setting

The physical system is a lab-scale closed-loop manufacturing system with five machines and two parallel paths, as shown in Figure 4. The system is available at the Department of Mechanical Engineering of Politecnico di Milano (Lugaresi et al. 2021). All the machines have their own queues and have stochastic processing times. Specifically, the processing time of each machine follows a truncated normal distribution with the following parameters, respectively: $N(11,2), N(17,2), N(80,2), N(80,2), N(10,2)$. The branching machine is $m = 2$. Hence, each part that at time t is located in front of machine 2 can follow two alternative paths, $\mathbb{P}_{jt} = \{(2,3,5);(2,4,5)\}$. There are twelve pallets circulating in the system and all pallets start from the queue in front of machine $m = 1$ at $t = 0$. The default routing policy for the parallel machines is a round robin policy. The system allows for data collection and control via the Message Queuing Telemetry Transport protocol. The DT architecture described in Section 3 has been developed in python. The codes can be found in the github repository (Bacelar dos Santos and Chalissery Lona 2023a). The dashboard

Figure 5: Dashboard developed on the Siemens MindSphere platform.

shown in Figure 5 has been developed in the Siemens MindSphere cloud service platform to allow for data visualization and monitoring.

5.2 Problem Description

We represent a situation in which the two alternative machines $\mathbb{M}_B = \{3, 4\}$ present a slow behaviour which results in a higher mean production time (e.g., due to aging). As counteraction, an improvement is applied to the system with regards to only one of the two machines (e.g., installation of a new resource, targeted maintenance). The intervention results in having a machine with lower processing time. The effect of this change is reflected in the whole system dynamics, and the default round robin policy might not be optimal in the new situation. The application of the DT architecture can result be advantageous to (1) monitor and identify the system behaviour change, and to (2) readily adapt to the new situation by defining a new routing policy for the parallel paths.

5.3 Experiments

In order to assess the advantage of the proposed digital twin architecture, we have performed two experimental campaigns described in the following cases:

- **Case 1**: production on the system where neither a validation procedure nor a production control service are available, only synchronization running with $f_s = 0.5\,\text{Hz}$ (i.e., monitoring).
- **Case 2**: production on the system equipped with also validation and the production control service described in Section 4 for optimizing the routing policy at the branching point.

In both cases, at $t = 900\,s$ the processing time distribution of machine $m = 3$ changes to *N(25,2)* (i.e., a reduction in the mean). From preliminary experiments it has been determined that the noise due to the stochastic behaviour of the physical system does not impact significantly the performance of the DT and its service. Hence, only one iteration of each experiment has been done. In this prototype implementation, each time the RCT of a part is estimated, the number of replications is not deterministic and depends on the waiting time on the queue of the branching machine. Namely, all the waiting time is used to gather as many replications as possible. Table 1 illustrates a portion of results obtained in the first case. The RCT values are normalized in relation with the last replication of each part. The complete results can be found in the public dataset (Bacelar dos Santos and Chalissery Lona 2023b).

Part j	Time t (UNIX)	$RCT(\phi_1)$	$RCT(\phi_1)$ [avg]	$RCT(\phi_2)$	$RCT(\phi_2)$ [avg]
41	1681219744	247	247	652	652
42	1681219767	237	–	635	–
42	1681219781	247	242	630	632.5
43	1681219787	179	–	699	–
43	1681219793	180	179.5	689	694
44	1681219805	235	–	686	–
44	1681219814	178	206.5	691	688.5
45	1681219820	170	–	718	–
45	1681219827	177	–	719	–
45	1681219833	178	177	713	718

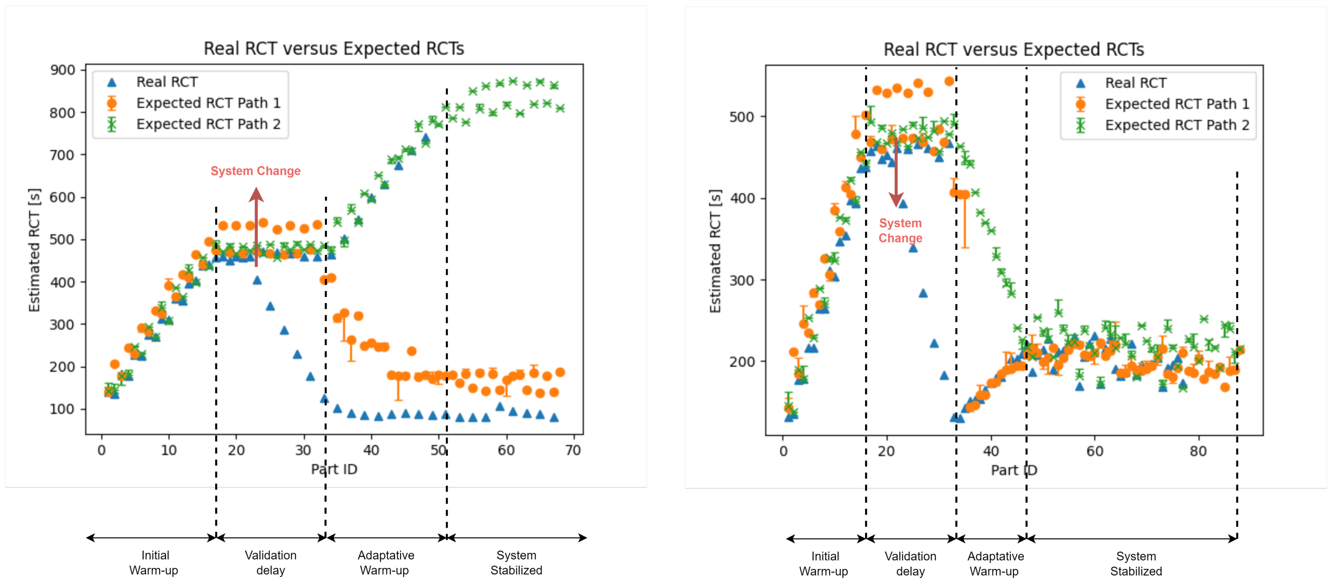
Table 1: Case 1 - Extract of results data for parts $j \in \{41, 42, 43, 44, 45\}$: $\phi_1 = (2,3,5)$, $\phi_2 = (2,4,5)$. Time values are in seconds.

5.3.1 Case 1: System without Digital Twin Control (Benchmark)

Figure 6a shows the comparison between the predicted and the actual RCT for each path $\phi \in \mathbb{P}_{jt}$. The reduction in the processing time of machine 3 increases the throughput of the respective path. Due to the default policy of branching machine $m = 2$, the parts starts accumulating in the upstream queue of machine $m = 4$. This results in lack of parts in the other queues of the system. At $t = 1075s$, machine 3 starts to face starvation due to lack of parts in the queue 3. This behavior can clearly be seen in the drop of machine 3 utilization in the Figure 7a. The RCT of parts visiting machine 4 becomes significantly higher that the one of parts visiting machine 3, as visible in Figure 6a. The reduction in utilization of machine 3 also propagates across the whole system as it is a closed-loop system. The system has been observed for $2098\,s$ in which 58 parts have been produced, for an average throughput of 99.52 parts per hour.

5.3.2 Case 2: System with Digital Twin Control

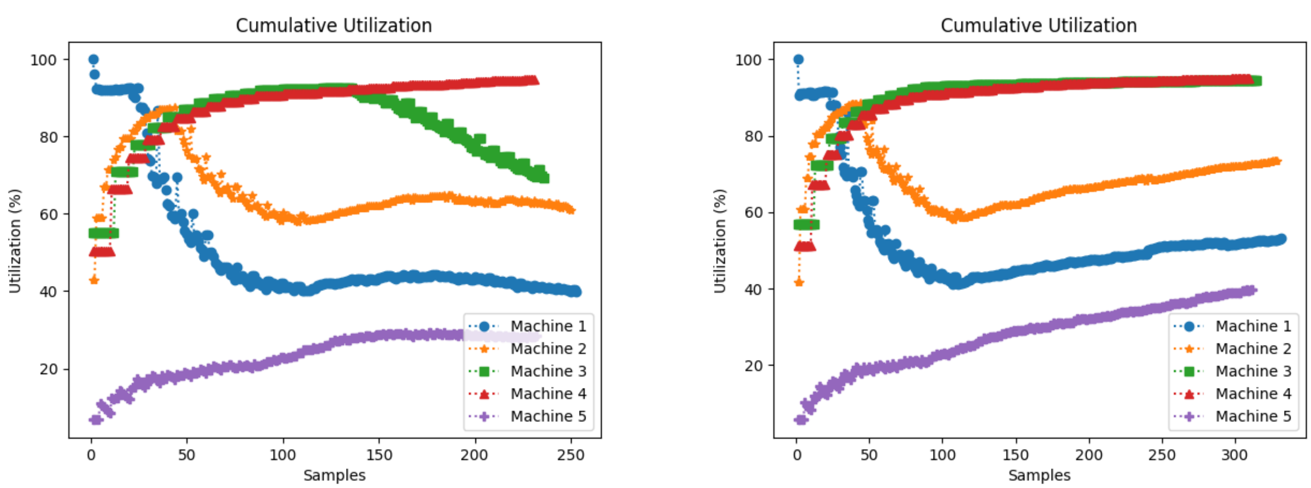
In this case, the validation and the production control services of the digital twin are also enabled, with $f_s = f_c = 0.5Hz$, $f_v = 0.011Hz$ and $\Delta = 0.01$. Figure 6b shows the comparison between the predicted and the actual RCT values of each produced part. Figure 6b also shows the two warm-up periods that happened in the system during the experiment: initial warm-up after the initialization of the physical system, second warm-up after the manual change of system behavior as a part of the experiment. At $t > 900\,s$, the validation component indicates that the model is not valid, and updates it to reflect the new conditions (Lugaresi et al. 2022). After the update, the production control service may use the latest digital model to predict the RCT. Hence, the accuracy of the RCT predictions is influenced by the validation frequency. This is visible in the delay in Figure 6b. The policy decided by the control service effectively distributes the parts across both paths in order to minimize their RCT. Indeed, in this scenario, more parts are sent to the faster machine $m = 3$. The control service balances the work load of the alternative machines in the system, which increases the system performance. Indeed, the difference between the RCT predicted for each possible path tends to converge as illustrated in Figure 6b. The overall effect of this can be seen in the rise of machine utilization across the system (Figure 7b): the utilization of both machine 3 and 4 converge to 90 %, differently from the diverging behaviour observed in Case 1. Finally, in this experiment the system was observed for $2075\,s$ and 77 parts have been produced, which results in an average throughput of 133.59 parts per hour.



(a) Case 1: DT without Production Control Service

(b) Case 2: DT with Production Control Service

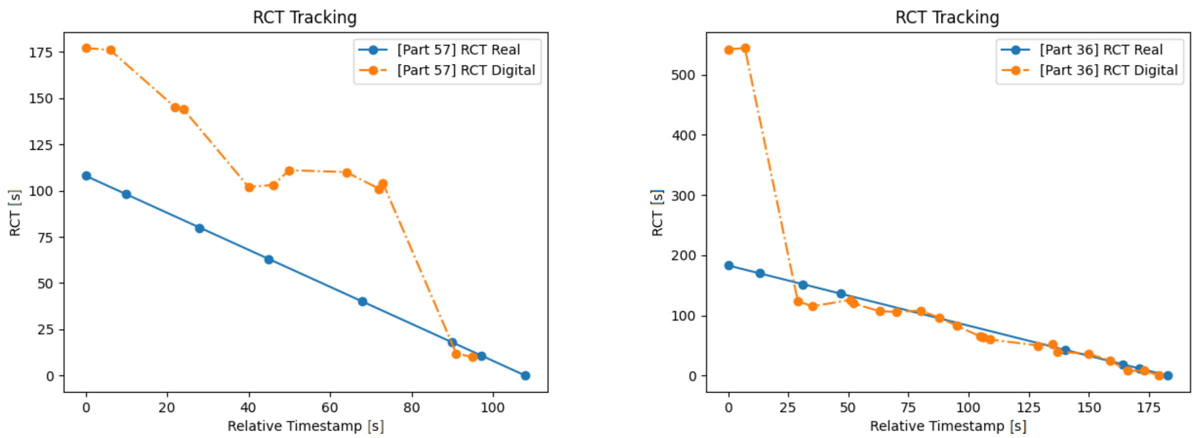
Figure 6: Comparison of Real RCT and Expected for both Case 1 and 2.



(a) Case 1: Machine utilization plot

(b) Case 2: Machine utilization plot

Figure 7: Machines Utilization for both Case 1 and 2.



(a) Part $j = 57$ in Case 1

(b) Part $j = 36$ in Case 2

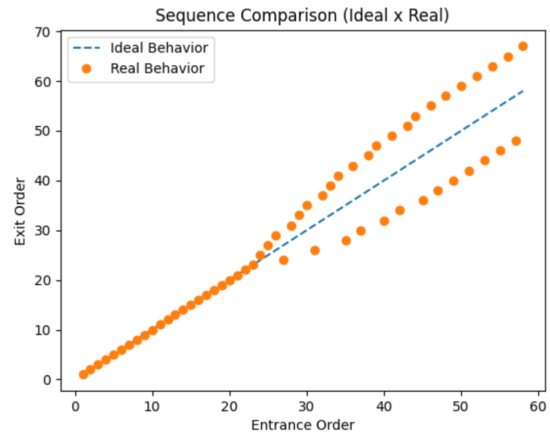
Figure 8: Comparison between actual and predicted RCT of a part.

5.4 Assessment of the Digital Twin Advantage

A comparison between the actual and the predicted RCT of a part in the system for both experimental cases is shown in Figure 8. The comparison demonstrates the advantage of controlling a system based on RCT prediction within a DT framework, which is proved by the fact that the predicted RCT values are closer to the actual values. Figure 8 shows the predicted and actual RCT values for two different parts selected from Case 1 and 2, respectively. The part used in the Figure 8 is selected appropriately to visualize the effect of mis-alignment of DT from the physical system and the subsequent error in RCT prediction because of it. The solid line of the Figure 8 is plotted after the end of the experiment, i.e., the exit time is known. The dashed line shows a more random behavior because it represents values obtained online, with an unknown exit time T_j. Figure 8a shows that in the first case the predicted RCT is significantly higher than the actual, due to the misalignment between the physical system and the DT. The RCT prediction becomes more accurate only toward the end of the production. On the other hand, Figure 8b shows how in the second case the RCT predictions start with an error and tend to correct values once the DT is aligned. It is also interesting to observe what happens to the positioning of parts in the system. For example, the Figure 9 shows the plot comparing the order of entrance and order of exit. In a perfectly balanced system, the order of entrance would tend to be equal to the exit one (in the plot, the diagonal line). The Figure 9a shows the comparison between the ideal behavior with the real one observed in the experiment of Case 1. after the becomes unbalanced, it never returns to a stable point. On the other hand, the Figure 9b show the same comparison in Case 2. The production control service enhances the workload balance of the system. Finally, we can state that the application of the production control services in the physical system allowed optimization of routing policy of parts as per the behavior and dynamics of the system. The policies set by the service significantly increased both the machine utilization and the system throughout. Indeed, in the second case a throughput 34% higher than the first case has been observed along with an increase of machine utilization in the system.

6 CONCLUSIONS

This work has developed and tested a complete digital twin architecture, with a focus on a production control service based on the estimation of the remaining cycle time via discrete event simulation. The test has demonstrated the applicability of the architecture in a controlled environment: the production control service is capable of controlling the routing policy, improving the machine utilization, optimizing



(a) Case 1: finished order do not follow the ideal behavior

(b) Case 2: finished order tends to follow the ideal behavior

Figure 9: Comparison of order of finished parts between Case 1 and Case 2 (Ideal behaviour: fully balanced system).

the distribution of parts along the system and reducing their average cycle time. This work is subject to several limitations that inspire future research endeavours. Currently, the performance predictions are dependent on the validation frequency, which imposes a delay between a change in the physical system and the moment when the updated digital model can be used. In future research, the optimal values of the frequency of each DT service should be investigated with proper methods and experiments. The proposed digital twin has been applied in a relatively small system, with limited complexity and full control on the information system. Realistic environments typically present a much higher complexity. Future research should consider the inclusion of the developed prediction service in a simulation-optimization component, in which proper methods should be used to manage the complexity and the computation effort.

REFERENCES

Bacelar dos Santos, P. L. and Chalissery Lona, Alex 2023a. "DtwinPy GitHub Repository". https://github.com/pedrolbacelar/Digital_Twin, accessed on 2023-04-21.

Bacelar dos Santos, P. L. and Chalissery Lona, Alex 2023b. "DtwinPy-Paper-WSC2023-RCT - Experiments Dataset in Zenodo". https://doi.org/10.5281/zenodo.7853820, accessed on 2023-04-21.

Choueiri, A. C., D. M. V. Sato, E. E. Scalabrin, and E. A. P. Santos. 2020. "An Extended Model for Remaining Time Prediction in Manufacturing Systems Using Process Mining". *Journal of Manufacturing Systems* 56:188–201.

Friederich, J., J. K. Lindeløv, and S. Lazarova-Molnar. 2023. "Predictive Process Monitoring for Prediction of Remaining Cycle Time in Automated Manufacturing: A Case Study".

Hyre, A., G. Harris, J. Osho, M. Pantelidakis, K. Mykoniatis, and J. Liu. 2022. "Digital Twins: Representation, Replication, Reality, and Relational (4Rs)". *Manufacturing Letters* 31:20–23.

Leng, J., Z. Chen, W. Sha, Z. Lin, J. Lin, and Q. Liu. 2022. "Digital Twins-based Flexible Operating of Open Architecture Production Line for Individualized Manufacturing". *Advanced Engineering Informatics* 53(January):101676.

Li, L., and Q. Yu. 2017. "Scheduling Strategy of Semiconductor Production Lines with Remaining Cycle Time Prediction". In *Proceedings of the 2017 Winter Simulation Conference*, edited by V. W. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 3679–3690. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lugaresi, G., V. V. Alba, and A. Matta. 2021. "Lab-scale Models of Manufacturing Systems for Testing Real-time Simulation and Production Control Technologies". *Journal of Manufacturing Systems* 58:93–108.

Lugaresi, G., S. Gangemi, G. Gazzoni, and A. Matta. 2022. "Online Validation of Simulation-Based Digital Twins Exploiting Time Series Analysis". In *Proceedings of 2022 Winter Simulation Conference (WSC)*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C. G. Corlu, L. H. Lee, E. P. Chew, T. Roeder, and P. Lendermann, 2912–2923. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

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

Papacharalampopoulos, A., C. K. Michail, and P. Stavropoulos. 2021. “Manufacturing Resilience and Agility Through Processes Digital Twin: Design and Testing Applied in the LPBF Case”. *Procedia CIRP* 103:164–169.

Passarin, Edoardo and Verucchi, Francesco 2022. “Online Synchronisation of Digital Twins a Control-based Methodology for Manufacturing Systems Applications”. https://www.politesi.polimi.it/handle/10589/189120, accessed on 2023-04-21.

Sakr, A. H., A. Aboelhassan, S. Yacout, and S. Bassetto. 2021. “Building Discrete-Event Simulation for Digital Twin Applications in Production Systems”. In *IEEE International Conference on Emerging Technologies and Factory Automation*. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Santhosh, G., and P. Sastry. 2023. “Efficient Depth-First Search Approach for Mining Injective General Episodes”. In *Proceedings of the 6th Joint International Conference on Data Science & Management of Data*, 1–9.

Shao, G. 2021. “Use Case Scenarios for Digital Twin Implementation Based on ISO 23247”. In *NIST Advanced Manufacturing Series 400-2*. Gaithersburg, MD, USA.

Sivakumar, A. I. 1999. “Optimization of a Cycle Time and Utilization in Semiconductor Test Manufacturing Using Simulation Based Online Near-real-time Scheduling System”. In *Proceedings of the 1999 Winter Simulation Conference*, edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 727–735. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Tan, B., and A. Matta. 2022. “Optimizing Digital Twin Synchronization in a Finite Horizon”. In *Proceedings of 2022 Winter Simulation Conference*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C. G. Corlu, L. H. Lee, E. P. Chew, T. Roeder, and P. Lendermann, 2924–2935. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Tirkel, I. 2011. “Cycle Time Prediction in Wafer Fabrication Line By Applying Data Mining Methods”. In *Proceedings of the 2011 Advanced Semiconductor Manufacturing Conference*, 1–5. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Traoré, M. K. 2021. “Unifying Digital Twin Framework Simulationbased Proof of Concept”. *IFAC-PapersOnLine* 54(1):886–893.

van Dongen, B. F., R. A. Crooy, and W. M. van der Aalst. 2008. “Cycle Time Prediction: When Will This Case Finally Be Finished?”. In *On the Move to Meaningful Internet Systems: OTM 2008: OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS, and ODBASE 2008, Monterrey, Mexico, November 9-14, 2008, Proceedings, Part I*, 319–336. Springer.

Xu, L.-Z., and Q.-S. Xie. 2021. “Dynamic Production Scheduling of Digital Twin Job-shop Based on Edge Computing”. *Journal of Information Science and Engineering* 37(1):93–105.

Yang, M., J. Moon, J. Jeong, S. Sin, and J. Kim. 2022. “A Novel Embedding Model Based on a Transition System for Building Industry-Collaborative Digital Twin”. *Applied Sciences (Switzerland)* 12(2).

Zhang, Y. F., Y. Q. Shao, J. F. Wang, and S. Q. Li. “Digital Twin-based Production Simulation of Discrete Manufacturing Shop-floor for Onsite Performance Analysis”. *IEEE International Conference on Industrial Engineering and Engineering Management*:1107–1111.

AUTHOR BIOGRAPHIES

GIOVANNI LUGARESI is Assistant Professor at the Department of Mechanical Engineering of KU Leuven (Belgium). His research interests include production planning and control, internal logistics, process mining, and robust optimization. His email address is giovanni.lugaresi@kuleuven.be.

PEDRO LUÍS BACELAR DOS SANTOS is a double-degree student of M.Sc in Management Engineering at Politecnico di Milano and Mecathronics Engineering at University of São Paulo. His research interests include Industry 4.0 and Entrepreneurship. His email address is pedrobacelar.official@gmail.com.

ALEX CHALISSERY LONA is a student of M.Sc in Mechanical engineering at Politecnico di Milano, Italy. His research interests include development and management of production systems and Industry 4.0. His email address is alex.chalissery@mail.polimi.it.

MONICA ROSSI Monica Rossi is Assistant Professor at Politecnico di Milano in the Department of Management, Economics and industrial Engineering. Her main research interests are in the areas of product development, innovation and design engineering, Lean Product Development and life cycle thinking. Her email address is monica.rossi@polimi.it.

EDUARDO ZANCUL is an Associate Professor at the University of São Paulo, School of Engineering. His research interests include advanced manufacturing, design processes, and design education. His email address is ezancul@usp.br.

ANDREA MATTA is Full Professor at Politecnico di Milano, where he currently teaches integrated manufacturing systems and manufacturing. His research area includes analysis, design and management of manufacturing and health care systems. He is Editor in Chief of the Flexible Services and Manufacturing Journal. His email address is andrea.matta@polimi.it.