

# Pattern-based Digital Twin for Optimizing Manufacturing Systems: A Real Industrial-case Application

Concetta Semeraro<sup>\*.1,2</sup>, Mario Lezoche<sup>2</sup>, Hervé Panetto<sup>2</sup>,  
Michele Dassisti<sup>1</sup>

<sup>1</sup> Department of Mechanical, Mathematics and Management (DMMM), Politecnico di Bari, Bari- Italy (e-mail: {concetta.semeraro, michele.dassisti} @poliba.it).

<sup>2</sup> Université de Lorraine, CNRS, CRAN, Nancy, France (e-mail: {concetta.semeraro, mario.lezoche, herve.panetto}@univ-lorraine.fr)

---

**Abstract:** The digital twin has received strong interests from researchers and industries since it allows predictive manufacturing by integrating the cyber and the physical space. An important prerequisite for the cyber-physical integration is a proper and highly-accurate digital model. Considering the complexity of digital modelling, the paper aims at developing and using predefined modelling patterns to enable building digital models independently of the specific application domain. The idea is explored and validated on a real case study by designing a set of patterns to create a digital twin model prototype adopted to control and optimize the manufacturing system taken into account.

*Keywords:* Digital Twin; CPS; Data-driven; Modelling Patterns; Predictive Manufacturing

---

## INTRODUCTION

The Smart Factory paradigm represents the "fourth industrial revolution" in manufacturing (Zuehlke 2010) that can be synthesized in systems networks integrating physical components and software for the improvement of manufacturing processes (Kagermann et al., 2013). The smart systems typically consist in different components including sensors for signal acquisition, communication units for data transmission between components, control and management units for decision-making, and actuators to perform the appropriate actions. In recent years, the emergence of Cyber-Physical Systems (CPS) has amplified the capacity of sensing the world through a network of connected devices using the existing network infrastructure, having the unexpected potential outcome of assuring sustainability in manufacturing operations. Cyber-physical system (CPS) aims at embedding computing, communication and controlling capabilities into physical assets to converge the physical space with the virtual space (Monostori et al. 2016) for a real-time operation control and management support. They suffer, on the other side, of missing adequate modelling techniques to harness the complexity of the physical world by taking into account the technological settings and the functional inter-correlations between them as well. Digital models constitute a major challenge in cyber-physical integration. The digital twin (DT) is intended to create high-fidelity virtual models for each physical entity (manufacturing settings; job-shop settings; etc.) to emulate their states and behaviours and controlling, optimizing (Grieves 2014) – in a word bring manufacturing

operation toward more sustainable statuses - as well as supporting the decision making. Generally, the modelling action has a specific application type, and this requires significant efforts to be afforded. This is the rationale of the present paper, facing the core challenge of the digital modelling of creating data-driven solutions employable mostly independently of the different application domains. The core idea discussed in the paper is to develop predefined modelling patterns for building digital models based on the function accomplished. The paper provides, in section 1, a state-of-the-art of main features and techniques employed for creating digital models. A set of patterns for improving the smartness of manufacturing processes and products is presented in section 2 and tested in section 3 with the help of a case study in an Italian SME Master Italy s.r.l. In section 4, a pattern has been instantiated to demonstrate how to design a digital twin prototype for controlling and predicting a specific problem of the process in analysis.

## 1. STATE OF ART

Several methods and tools have been proposed in the literature with the aim to develop digital model of complex system and their related performance. The approaches can be distinguished in: model-based, data-driven, and hybrid approaches (Tobon-Mejia et al. 2012).

The model-based approaches rely on the use of models to simulate the systems behaviour in different operating conditions. A model is a representation of a system for a specific purpose (Konikow and Bredehoeft 1992) but it is not

easy to develop and keep updated during the system life-cycle (Luo et al. 2008). In contrast to the model-based approaches, where a-priori knowledge is needed, in data-driven approaches the availability of large amount of data is required.

The data-driven approaches allow to integrate parameters across different domains. The data-driven approaches, in fact, aim at transforming data into relevant information and reliable behavioural models (Okoh et al. 2014). Data-driven approaches give the possibility to discover tacit knowledge from a set of data. The main challenges of data-driven approaches involve treating massive data sets and managing their high dimensionality (Kusiak 2018). The results and the performances depend on the quality of training data (Tidriri et al. 2016). It is typically time consuming to find valuable knowledge from a huge dataset (Zhao et al. 2019) and this requires a-priori knowledge model.

The hybrid approaches involve model and data (Tidriri et al. 2016). These not define a common framework to apply to different systems (Ghosh, Ng, and Srinivasan 2011). Most of the model-based, data-driven and hybrid approaches have been developed for specific application. The choice of the method to apply is typically tailored on the specific use case based on the specific needs. Developing time-tested solutions to recurring set of problems is closely related to object-oriented design pattern domain.

The object-oriented programming (OOP) is a programming language model that organizes software design around objects, rather than functions and logic (Gamma, Helm, and Johnson 1995). A design pattern in OOP is a general repeatable solution to a commonly occurring problem in software design (Gamma, Helm, and Johnson 1995). The design pattern (OOP) defines, develops, and implements repeatable solution to a commonly occurring problem in software engineering. However, the design patterns also have some trade-offs and limitations. The detection of patterns and the updating is developed by experts: it is not an automatic process. The main consequence of using design patterns is trying to apply many of them across the same objects. From the state of art analysis on OOP, it follows that it is necessary to make the digital modelling more structured and reliable. A pattern could be a good solution because it could be used several times to create models for different application domains. However, unlike patterns developed in OOP, patterns need to be created and updated automatically. The paper aims to develop a set modelling pattern detected automatically from data. The patterns need to be modelled for describing a system independently from its context of application, as it will be shown in sections 2 and 3.

## 2. PATTERN-BASED DIGITAL TWIN

The core aspect of the paper is to identify and formalise elements of "invariant" modelling contributing to the construction of information and functional systems to improve

the sustainability and the performances of manufacturing processes and products. Formal Concept Analysis (FCA) is applied to detect automatically patterns in a large set of data (Valtchev, Missaoui, and Godin 2004). FCA is a mathematical theory oriented at applications in knowledge representation (Wille 2002). It provides tools to group the data and to discover formal patterns by representing it: as a hierarchy of formal concepts organised in a semi ordered set named lattice or as association rules. The relation concept analysis (RCA) extends FCA to the processing of multi-relational datasets (Rouane-Hacene et al. 2013). Huge datasets, belonging to the real test case discussed in the next section, have been selected to detect all possible patterns. The use of Multi Relational Data Mining techniques has enabled the extraction of the tacit knowledge in the datasets of the processes analysed and the detection of data-driven patterns.

In view of the above, a list of data-driven modelling patterns has been presented in (Semeraro 2020) according to the methodology described in (Semeraro et al. 2019). The detected data-driven patterns are listed as follow: Filling pattern (P1); Re-start machine pattern (P2); Clamping force pattern (P3); Pression control pattern (P4); State pattern (P5). The filling pattern (P1) can be used whenever it is necessary to model the behaviour of a plunger or an injection system. The re-start machine pattern (P2) can be applied for describing the machine restart after machine downtimes. The clamping force pattern (P3) can be instantiated for describing the behaviour of a clamping system (e.g. the clamping of a mould) or for describing and preventing mechanical breakdowns. The pression control pattern (P4) can be useful for modelling hydraulic systems (e.g. the clamping of a mould) and related problems. The state pattern (P5) can be instantiated for simulating different states of the machine based on the evaluation of a set of parameters. A data-driven pattern is described by 1) ID Pattern; 2) Pattern Name; 3) Description; 4) An example of association rule; 5) Data View; 6) Model View (using System Modelling Language); 7) Applicability (Semeraro 2020) as shown in Fig.1. These formal elements, which are made concrete through the modelling of systems and procedures in SysML language, make it possible to represent the knowledge and its deep relationship with the manufacturing processes. In this way, they make the knowledge more easily reusable and are the basis of the efforts of standardisation. The formalized data-driven modelling patterns represent the design criteria for building simple and efficient end-user digital twins' interfaces to support the employees in decision making process. A pattern includes rules, constraints and deductions related to the production processes, such as the constraint of the processing capability of a certain equipment. These can be formalized in algorithms to make the DT able to evaluate, optimize and/or predict possible problems or failures. The formalized patterns can be used directly into different systems for design digital twins or

for other applications finalized to control and manage cyber-physical systems. It is possible because each pattern represents a specific function which can be modelled for the emulation of cyber-physical systems and thus contribute to the digital transformation of industrial production companies. In this way, data-driven modelling patterns can be combined, based on the specific application, to create easily digital models for realising its digital twin.

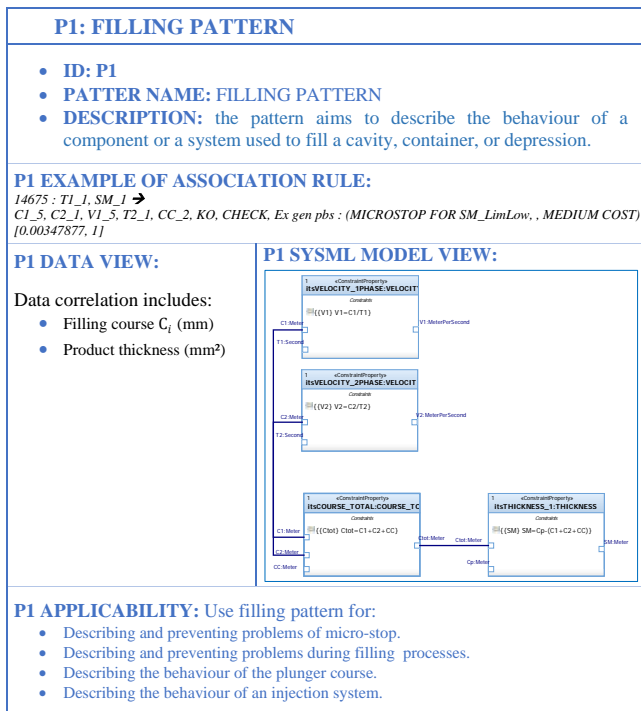


Fig. 1: Filling Pattern (P1)

### 3. CASE STUDY: MASTER ITALY s.r.l

Master Italy is an Italian SME that designs and produces accessories and components for doors and windows since 1986. The die casting aluminium is the process in analysis because it is affected by quality and maintenance problems (Dassisti, Semeraro, and Chimenti 2019). Die casting aluminium is a manufacturing process in which molten metal is poured or forced into steel moulds to shape the steel corners. The process cycle consists of four main stages:

1. *Melting*: the aluminium enters at the solid state and exits at the molten state. Once melted and taken up to proper temperature, the aluminium is transferred to the die cast machine.
2. *Injection*: the molten aluminium is injected into the mould, through a plunger. When the machine is ready for its next cycle, an automated ladle takes a prescribed volume of molten aluminium ( $V_{al}$ ) from the holding furnace and pours it into the mould. Once pouring is complete, the injection phases begin. The first injection phase is the slow phase (T1) where the plunger moves forward at a low speed (V1). After a prescribed distance

(C1), the plunger enters (C2) an intermediate speed phase (V2) where the speed is increased to fill the mould. Once this is complete (T2), the machine enters a fast phase where speed is greatly increased to fill the part cavity with aluminium (CC).

3. *Cooling*: the molten aluminium solidifies in the mould cavity. When the entire cavity is filled and the molten metal is solidified at a given pressure (PM) and clamping force (FC), the final shape of the casting is formed.
4. *Extraction*: an ejection mechanism pushes the product out of the mould cavity. When the injection cycle is completed and the machine is fully open, the die cast pushes out and the thickness of the die cast (SM) is controlled to detect quality defects.

The production line is controlled by the programmable logic controller (PLC) and by the manufacturing execution system (MES). PLC monitors the technological parameters, processes the data, and triggers outputs based on the setting values of each technological parameter. MES monitors the productivity, tracks machine down-time and measures the overall equipment effectiveness (OEE) of the production line in real-time. The main downtimes concern the micro-stop machine and the mechanical maintenance. The micro-stop represents a down-time less than five minutes. This kind of problem is caused by the plunger course during the injection stage. The filling pattern (P1) can be used to describe and prevent this type of failure. The mechanical breakdown causes a down-time of the machine more than two hours. It has a high impact on maintenance costs. Mechanical breakdown is typically generated by the clamping force during cooling stage. The pattern P3 can be applied to prevent mechanical maintenance. Finally, P5 pattern should be employed to simulate the behaviour of the machine and predict possible failures. The patterns have been instantiated in a web platform, as displayed in Fig.2. The platform is created in SQL language in Oracle database, with the aim of formalizing and instantiating the modelling patterns and creating analytical reports for the employees.

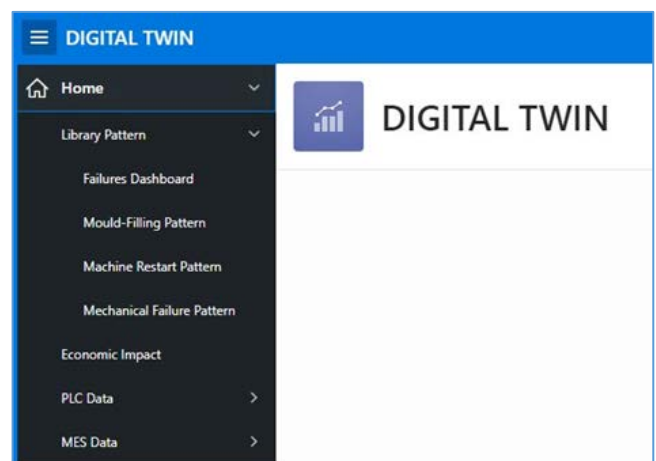


Fig. 2: Web Platform of Data-driven Patterns

P1 pattern in Fig.2 aims at providing a solution to describe the behaviour of a component or a system used to fill a cavity, container, or depression. The pattern presents the correlation between two parameters: the filling course (mm) and the product thickness (mm<sup>2</sup>). In the process in analysis, the instantiation of P1 pattern results in the correlation between the following parameters:

- Plunger course C1 (mm) in the first phase of the injection stage.
- Plunger course C2 (mm) in the second phase of the injection stage.
- Plunger course CC (mm) in the multiplied phase of the injection stage.
- Thickness of die cast SM (mm<sup>2</sup>).

The dashboard in Fig.3a allows to understand the daily micro-stop and the frequency of this problem over time, the duration, and the economic impact of each occurrence. It is possible to click on each occurrence to understand the degradation trends of SM over time, as shown in Fig.3b. The red line indicates the threshold value of SM (SM Min). When the value of SM is equal to or below the threshold value (SM Min) there is the recasting of the injection cycle (36 steel corners). The recast products are recorded in MES as scraps and this impacts on the quality performances of the process. The dashboard shows that a micro stop is always preceded by a recasting. It is necessary to predict the value of SM before the threshold value to avoid quality scraps (dashed circle in red). In particular, the dashboard in Fig.3c demonstrates that there is a relationship between SM, C1, C2 and CC. C1 and C2 are inversely proportional as a function of SM. It means that it is necessary to detect all possible combinations of C1, C2, CC and SM for defining and understanding the behaviour and failures of the process. All correlations are split into three area as presented in Fig.3d:

- The green area indicates all possible combination values between C1, C2 and CC when SM is greater than SM Min. These combinations define when the process functioning properly (STATE 1).
- The yellow area indicates all possible combination values between C1, C2 and CC when SM is close to the threshold value (STATE 2).
- The red area indicates all possible combination between values C1, C2 and CC when SM is equal to or below SM Min. These combinations indicate a failure of the system and consequent micro stop of the machine (STATE 3).

The yellow area allows to discover all combinations for predicting all possible occurrences of micro-stop avoiding in this way quality scraps.

#### 4. DIGITAL TWIN PROTOTYPE

A digital twin has been designed and developed, as shown in Fig.4, to predict problems of micro-stop based on the outcomes of P1 instantiation. The solution here proposed consists of a realistic production environment but “augmented” with intrinsic technological knowledge. With our approach, the physical settings interact with the digital space, according to specific properties and rules, to understand the behaviour of the process and the correlations between technological parameters. The digital twin in Fig.4 can get and evaluate a set of data, to emulate the behaviour of the process and recognize possible micro-stops based on the combinations presented in Fig.2d. The digital twin compares the values of C1, C2, CC and SM over time and distinguishes all possible states detecting in Fig. 4b that a micro-stop is going to happen in five minutes. In the red rectangle at the top left, the digital twin prescribes possible actions to the employees to prevent a micro-stop.

The digital twin has been designed to support the employees in decision-making process to:

- Prevent quality scraps.
- Analyse and correlate the symptoms and causes of failures and defects in production.
- Alert the operators through proper alarm systems about abnormal or out of tolerance situations.
- Support the choice of corrective actions to avoid possible failures.

#### 5. CONCLUSIONS AND FUTURE WORKS

The core idea of the paper is to extract patterns automatically from data and model these to achieve invariance in modelling. The paper defines and presents how to detect data-driven modelling patterns with the aim of building digital models valid in different applications and use cases. The contribution has been tested through a real implementation on an industrial case-study. A web platform and a digital twin prototype have been presented for showing how to instantiate and use a pattern.

The future research direction is to enrich the pattern’ semantics to create a comprehensive library of formalized data-driven patterns. It means to select different production lines to apply the same patterns for designing digital models for different applications and scopes. Other sets of data need to be selected such as logistic data, product data, customers’ data to extract new data-driven constructs. In this way it is possible to create a consistent library of patterns. In this way, data-driven patterns can be combined to create digital models for describing recurrent problems, behaviours, or situations.

### Daily Micro-Stop: Monthly Analysis

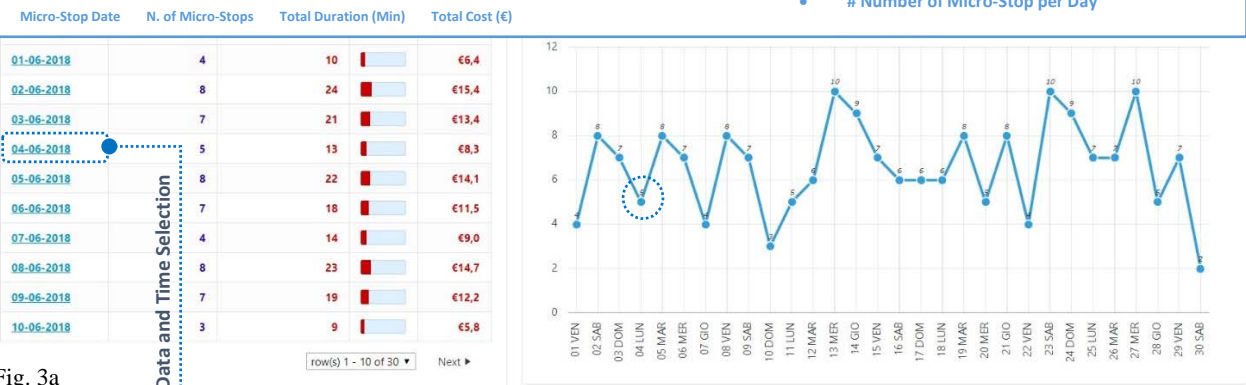


Fig. 3a

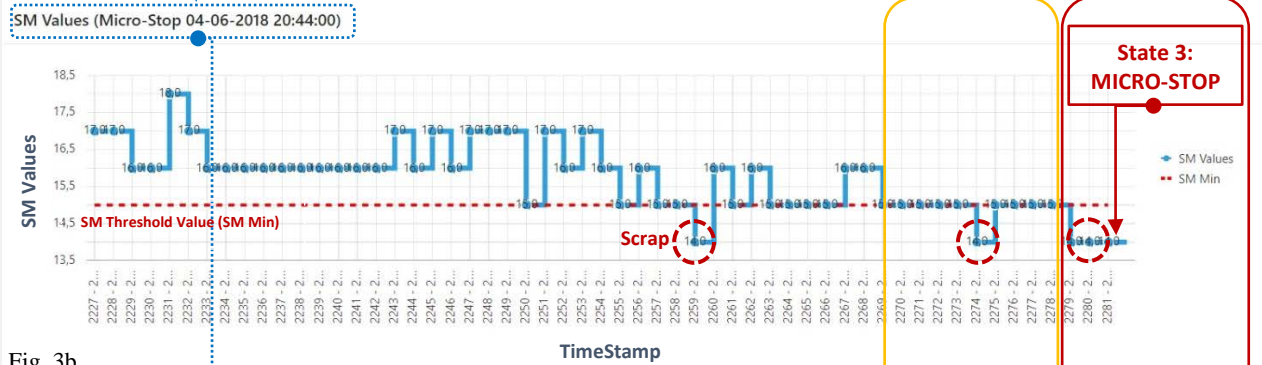


Fig. 3b



Fig. 3c

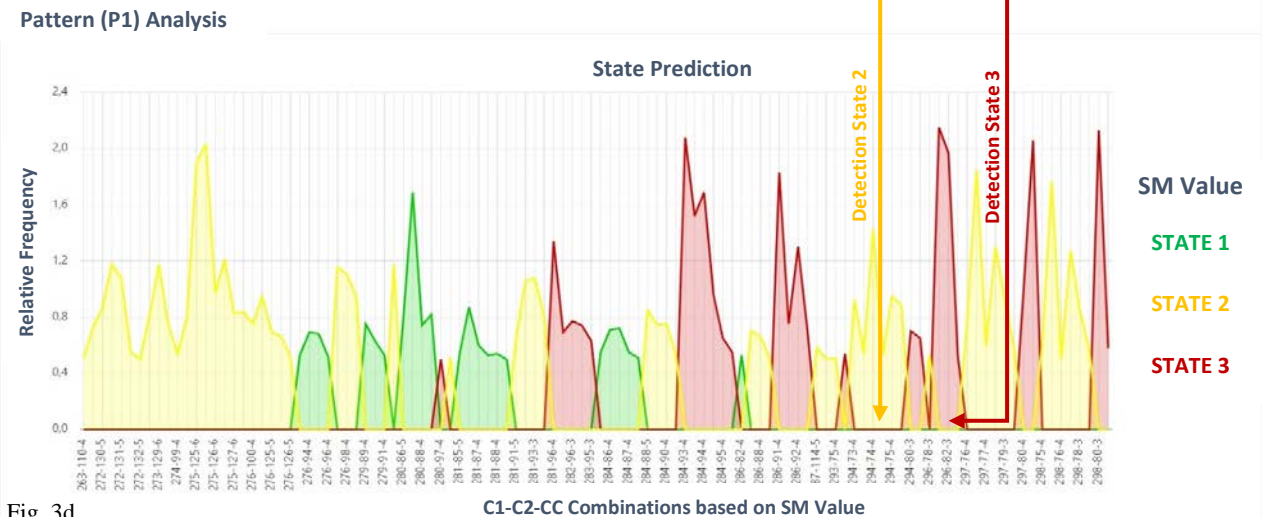


Fig. 3d



Fig. 3: PI Pattern Instantiation and Analysis

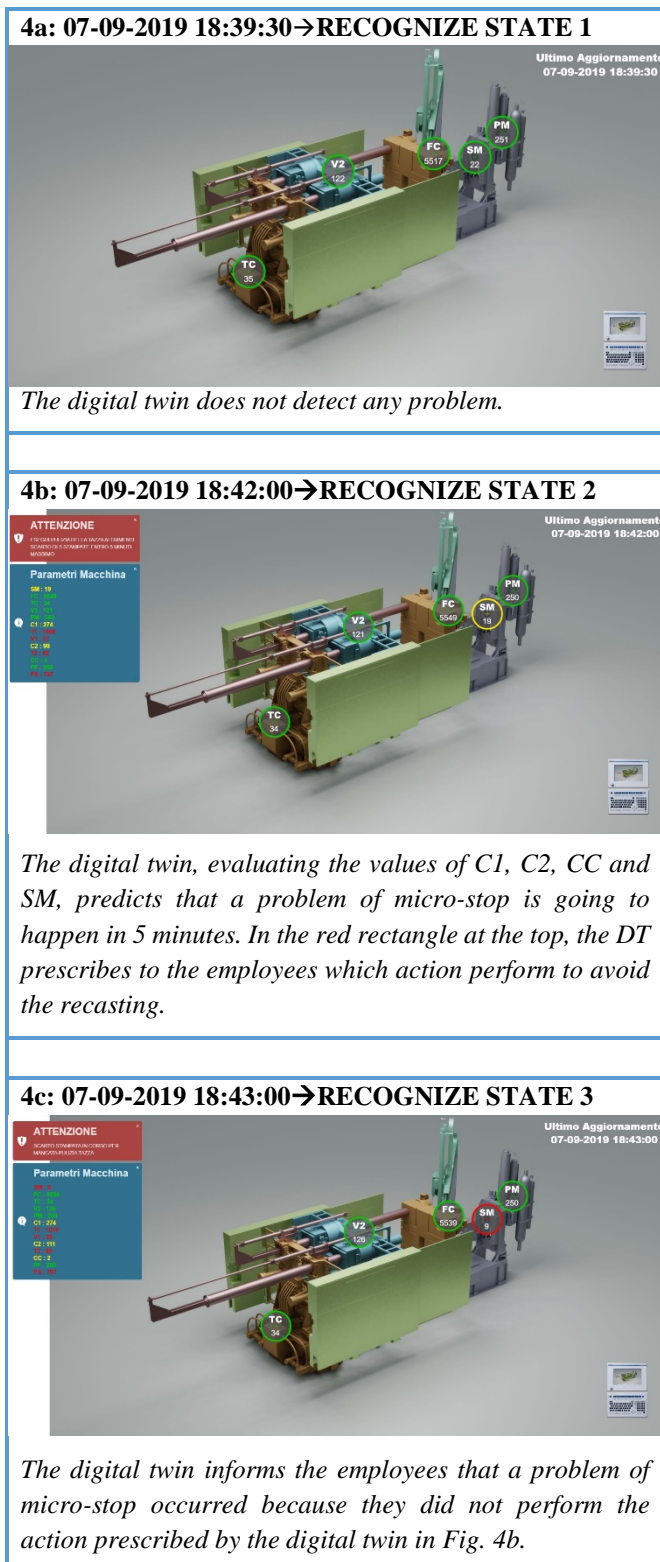


Fig. 4: Digital Twin Interface

## Acknowledgement

This research has been funded by the company “Master Italy s.r.l.”. We are deeply indebted with Michele Loperfido and Stefano Cafagna for their moral support and encouragements.

## References

- Dassisti, Michele, Concetta Semeraro, and Michela Chimenti. 2019. ‘Hybrid Exergetic Analysis-LCA Approach and the Industry 4.0 Paradigm: Assessing Manufacturing Sustainability in an Italian SME’. *Procedia Manufacturing* 33: 655–62.
- Gamma, Erich, Richard Helm, and Ralph E. Johnson. 1995. *Design patterns: elements of reusable object-oriented software*. Reading (Mass.), Etats-Unis d’Amérique, Canada, France.
- Ghosh, Kaushik, Yew Seng Ng, and Rajagopalan Srinivasan. 2011. ‘Evaluation of Decision Fusion Strategies for Effective Collaboration among Heterogeneous Fault Diagnostic Methods’. *Computers & Chemical Engineering* 35 (2): 342–55.
- Grieves, Michael. 2014. ‘Digital Twin: Manufacturing Excellence through Virtual Factory Replication’. *White Paper*.
- Konikow, Leonard F., and John D. Bredehoeft. 1992. ‘Ground-Water Models Cannot Be Validated’. *Advances in Water Resources* 15 (1): 75–83.
- Kusiak, Andrew. 2018. ‘Smart Manufacturing’. *International Journal of Production Research* 56 (1–2): 508–17.
- Luo, Jianhui, Krishna R. Pattipati, Liu Qiao, and Shunsuke Chigusa. 2008. ‘Model-Based Prognostic Techniques Applied to a Suspension System’. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38 (5): 1156–68.
- Monostori, László, Botond Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, and K. Ueda. 2016. ‘Cyber-Physical Systems in Manufacturing’. *Cirp Annals* 65 (2): 621–41.
- Okoh, Caxton, Rajkumar Roy, Jorn Mehnen, and L. Redding. 2014. ‘Overview of Remaining Useful Life Prediction Techniques in Through-Life Engineering Services’. *Procedia CIRP* 16: 158–63.
- Rouane-Hacene, Mohamed, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. 2013. ‘Relational Concept Lattice Analysis: Mining Concept Lattices from Multi-Relational Data’. *Annals of Mathematics and Artificial Intelligence* 67 (1): 81–108.
- Semeraro. 2020. ‘Contribution to the Formalization of Data-Driven Invariant Modelling Constructs of Cyber-Physical Systems, Co-Mentorship Phd of the Université de Lorraine (France) and of the Politecnico Di Bari (Italy)’.
- Semeraro, Concetta, Mario Lezoche, Hervé Panetto, Michele Dassisti, and Stefano Cafagna. 2019. ‘Data-Driven Pattern-Based Constructs Definition for the Digital Transformation Modelling of Collaborative Networked Manufacturing Enterprises’. In *Working Conference on Virtual Enterprises*, 507–15. Springer.
- Tidirri, Khaoula, Nizar Chatti, Sylvain Verron, and Téodor Tiplica. 2016. ‘Bridging Data-Driven and Model-Based Approaches for Process Fault Diagnosis and Health Monitoring: A Review of Researches and Future Challenges’. *Annual Reviews in Control* 42: 63–81.
- Tobon-Mejia, Diego Alejandro, Kamal Medjaher, Noureddine Zerhouni, and Gerard Tripot. 2012. ‘A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models’. *IEEE Transactions on Reliability* 61 (2): 491–503.
- Valtchev, Petko, Rokia Missaoui, and Robert Godin. 2004. ‘Formal Concept Analysis for Knowledge Discovery and Data Mining: The New Challenges’. In *International Conference on Formal Concept Analysis*, 352–71. Springer.
- Wille, Rudolf. 2002. ‘Why Can Concept Lattices Support Knowledge Discovery in Databases?’ *Journal of Experimental & Theoretical Artificial Intelligence* 14 (2–3): 81–92. <https://doi.org/10.1080/09528130210164161>.
- Zhao, Yang, Chaobo Zhang, Yiwen Zhang, Zihao Wang, and Junyang Li. 2019. ‘A Review of Data Mining Technologies in Building Energy Systems: Load Prediction, Pattern Identification, Fault Detection and Diagnosis’. *Energy and Built Environment*, November. <https://doi.org/10.1016/j.enbenv.2019.11.003>.
- Zuehlke, Detlef. 2010. ‘SmartFactory—Towards a Factory-of-Things’. *Annual Reviews in Control* 34 (1): 129–38.