A Digital Twin Framework for Predictive Maintenance in Industry 4.0

S. Mihai[1], W. Davis[1], D.V. Hung[1], R. Trestian[1], M. Karamanoglu[1], B. Barn[1]
R. V. Prasad[2], H. Venkataraman[2] and H. Nguyen[1]

[1]London Digital Twin Research Centre, Middlesex University, London, UK
{SM3488, WD085}@live.mdx.ac.uk, {D.VietHung, R.Trestian, M.Karamanoglu, B.Barn, H.Nguyen}@mdx.ac.uk

[2] Indian Institute of Information Technology, Sri City, India
{yrv.prasad, hvraman}@iiits.in

***Abstract*—The rapid advancements in manufacturing technologies are transforming the current industrial landscape through Industry 4.0, which refers not only to the integration of information technology with industrial production, but also to the use of innovative technologies and novel data management approaches. The target is to enable the manufacturers and the entire supply chain to save time, boost productivity, reduce waste and costs, and respond flexibly and efficiently to consumers' requirements. Industry 4.0 moves the digitization of manufacturing components and processes a step further by creating smart factories. Within this context, one of the key enabling technologies for Industry 4.0 is the adoption and integration of the Digital Twin (DT). However, most of the DT solutions provided by the current leading vendors are in fact digital models or digital shadows, and not digital twins. This is due to the fact that there is no common understanding of the definition of the DT amongst the leading vendors, and its usage is slightly different but showcased under the same umbrella of DT. In this paper, a DT framework is proposed that replicates the processes of a real production line for product assembly using the Festo Cyber Physical Factory for Industry 4.0 located at Middlesex University. Moreover, the paper introduces a viable framework for interlinking the physical system with its digital instance in order to offer extended predictive maintenance services and form a fully integrated digital twin solution.**

***Index Terms*—Predictive Maintenance, Digital Twin, Smart Manufacturing, Industry 4.0.**

I. INTRODUCTION

The Fourth Industrial Revolution brings about fast-paced growth in the ways industry operates through the introduction of automation technologies that drive production-based companies. The focus is now shifted towards increasing efficiency and sustainability in manufacturing machinery by leveraging Machine Learning (ML), Artificial Intelligence (AI), the Industrial Internet of Things (IIOT), and Cyber-Physical Systems (CPS). These technologies promise improved productivity by enabling communication between heterogeneous systems to transmit large amounts of data at high speeds, thus allowing real-time monitoring and analysis of machinery parameters. The Big Data is then processed using state-of-the-art techniques to extract meaningful information to be used for optimal decision-making.

Industry 4.0 is built upon intelligent systems and processes, sensor networks, low latency communication, distributed control systems and specialized automation technologies. As such, the complexity of up and coming solutions demands expert knowledge in order to monitor the state of the machinery, assess its performance, detect equipment failure and provide timely maintenance. The delivery of quality products and services strongly depends on the manufacturing production line machines' health, and neglect or poorly synchronised repair action can result in additional costs and resource waste, manifested as unplanned downtime and additional workforce. As such, in the context of Industry 4.0, an automated solution is required to provide monitoring, simulation, performance evaluation and threat analysis.

The concept of Digital Twin (DT) is not a new one, but nowadays it is evolving and becoming gradually more prominent in the industry as companies start to see real value in implementing it. There are various definitions for this technology in literature and between businesses, with some companies denominating the digital model of a physical entity as a digital twin, when in fact it would be more appropriate to call it a *digital shadow*. A working definition would be that the DT represents the two-way interaction between the twin entities which is a challenging task to achieve due to the highly complex configuration. The DT needs to be fully integrated within the Production Line Engineering (PLE) and interact with the environment and its physical processes. This integration becomes even more complex because of the heterogeneity of components and the tight interaction between software, networks/platform and physical components. Thus, the digital twin is not just a passive replica of the real system, but it is an active and reactive component that can continuously evaluate the current state of its real twin and provide expert recommendations in terms of optimising processes, predicting and scheduling maintenance, and improving the design and overall performance. Consequently, in this work, DT refers to the two-way interaction between the physical entity and its virtual replica, achieving full integration within the PLE, along with its surrounding environment and its physical processes. This integration indeed requires domain knowledge of the physical asset, as well as of several enabling technologies, but it is the successful implementation of this two-way connection that enables harnessing the Predictive Maintenance (PdM) capabilities of Digital Twins to increase production efficiency and avoid equipment failure.

In this context, this paper proposes a framework for connecting a real Festo Cyber Physical Factory for Industry 4.0 with its digital representation in order to facilitate the implementation of a Predictive Maintenance model, by compensating for the lack of run-to-failure data through error simulations run on the physical asset and its Digital Twin. The framework highlights the integration of the Digital Twin into a Predictive Maintenance architecture, with the objective of leveraging the virtual replica's operation history and simulation capabilities to achieve accurate and adaptive Predictive Maintenance. Additionally, the work describes the PdM use-case of the Festo CP Lab's Tunnel Furnace Station, facilitated by the DT.

The rest of this paper is structured as follows: Section 2 explores related works on this subject, Section 3 provides an overview of the Cyber Physical Factory's configuration, with an emphasis on the tunnel furnace station, Section 4 describes the proposed Digital Twin framework for Predictive Maintenance, and Section 5 showcases this work's conclusions.

II. Related Work

In its initial stages of development, the Digital Twin is seen as a reliable auxiliary tool for decision-making assistance in view of obtaining optimal functionality and productivity. The information-driven nature of the DT along with its capability of one-to-one replication of the physical asset's behaviour make it the go-to solution for process emulation. More of a digital shadow in this instance, the technology continuously monitors its physical counterpart, gauges its state, mirrors it, and provides meaningful insight and suggestions that an operator can take into consideration to decide on the best course of action that is to be followed. Furthermore, it learns behavioural patterns by studying the physical asset, then it uses that information to illustrate the functioning of the real object in hypothetical scenarios through simulations. This mimicking characteristic of the Digital Twin allows the testing of infant technologies, like 5G and Vehicle-to-Everything networks, before roll out, saving resources and time [1].

Once the two-way interaction is established, the digital shadow becomes a Digital Twin and its gainful feedback can be cultivated to make products, operations and infrastructure more flexible, reliable and predictable. Big Data analytics have been proven to be a reliable tool in process optimisation for production lines. Ferreiro et al. [2] show that power data is a good indicator of a machine's health status, reflecting even slight component wear, and that faults in the equipment can be identified by comparing real-time monitored process data to a healthy baseline. Wang et al. [3] proposed a data-driven Intelligent Immune System to react and adapt to manufacturing condition changes in order to optimise the machinery's energy consumption and increase productivity. The work demonstrates that the equipment's power consumption pattern can again be a comprehensive parameter for condition monitoring. Similarly, Liang et al. [4] put forward a Big Data-enabled anomaly detection method based on power data. The paper showcases the use of pre-processing mechanisms that store, categorise, and scale the data before extracting a set of critical features that can indicate anomalies. The framework also includes a threshold optimisation algorithm to allow accurate failure detection in dynamic working conditions. Liang et al. in [5] proposed a layered architecture for a low latency deployment of a Convolutional Neural Network (CNN) – based prognosis system. The proposed scheme consists of three layers that share responsibilities effectively, keeping high-speed processing capabilities on the terminal and fog layers, close to the manufacturing equipment, and leaving the training of the CNN to the cloud layer.

In order to achieve satisfying accuracy, Predictive Maintenance schemes that rely on Big Data more often than not require massive amounts of historical failure data. This is not always available, especially in old machinery where maintenance records have not been kept, or in equipment whose uptime is so crucial that no run-to-failure scenarios were allowed. In these instances, a digital solution to emulate the tool's behaviour in various states is needed. In this direction, Hsieh et al. [6] proposed a PdM scheme based on Virtual Metrology that can generate a baseline model for the monitored target device to reflect its healthy state using real process-related data. Then, the equipment's condition is monitored and when its state deviates from the baseline, a fault is registered and fed into the Remaining Useful Life (RUL) regression-based predicting module. As shown by Ortega in [7], the concept of Digital Twin can be used to obtain accurate fault detection and failure prediction, replacing the need for extensive historical run-to-failure data with the DT's ability to replicate the functional patterns of the real equipment. The research is built around a Digital Twin model of the drilling station of a Festo Cyber-Physical Laboratory (CP Lab). The work focuses on the drilling station, where an accelerometer has been installed to record the vibration signal that describes the drilling. The vibration data together with the energy consumption information provided by the CP Lab provide enough information to estimate the RUL using an exponential degradation statistical model. Aivaliotis et al. [8] proposed an approach based on a Digital Twin that mirrors the physical properties of the device, generating virtual sensor data that is synchronised and converged with the real sensors' data, then using Prognostics and Health Management (PHM) techniques to generate the RUL.

These applications require specialised expert knowledge of the operation of the target device in order to build its Digital Twin. In practice, manufacturing equipment is always adapting to new requirements depending on the product evolution. Barthelmey et al. [9] researched a solution to this issue, showcasing a flexible Digital Twin architecture with an update service to allow accommodation of new components in the physical asset and the twin. To enable PdM capabilities, the framework incorporates a Human Machine Interface (HMI) to allow operators to register and label failures as they happen, over time accumulating enough data to train a supervised Machine Learning model. Several other use-cases for the Digital Twin for Predictive Maintenance purposes are showcased by

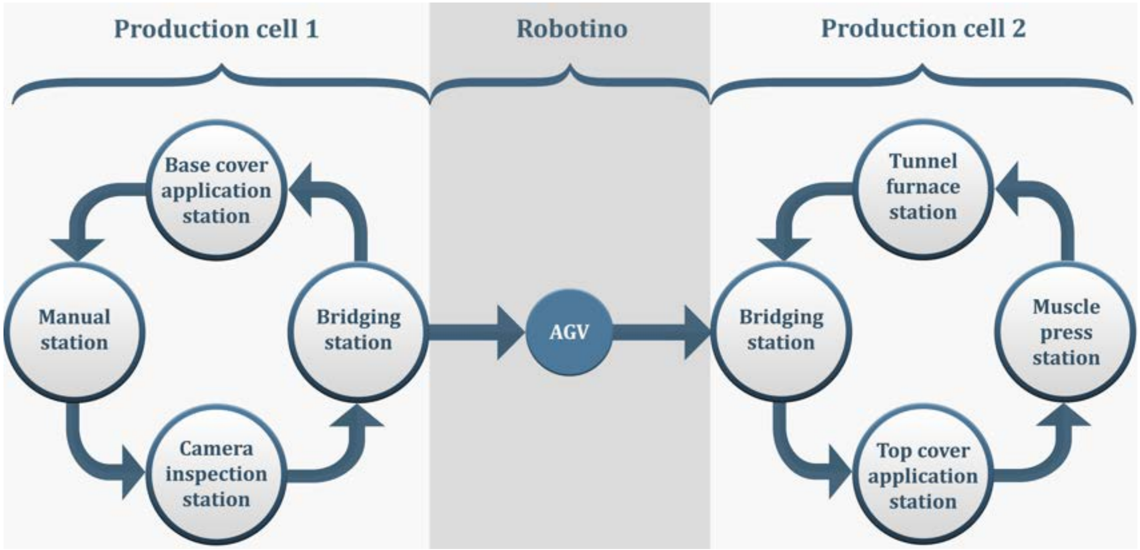

Fig. 1: Festo Cyber Physical Factory functional diagram.

Aivaliotis et al. in [10].

Another approach to fault detection that discards the need of expert domain knowledge is presented in [11], whose authors propose a diagnosis framework based on a Digital Twin that is created during the production phase of the real asset. In this instance, potential failure modes can be simulated in the virtual environment before production in order to train a Stacked Sparse autoencoder deep neural network to identify several stages of failure. Once the product is deployed and data can be collected in real time, Deep Transfer Learning is used to further adapt the parameters to the new distribution of the real data, obtaining exceptional accuracy.

In summary, the Predictive Maintenance frameworks studied in related works are greatly dependent on historical run-to-failure data. Some of them also use Digital Twins for the purpose of generating synthetic data that can be used for training Machine Learning algorithms, whether traditional ones or deep-learning. However, the underlying assumption in most of these works is that the configuration of the machine remains static over its whole lifetime, to ensure a constant distribution of the measurements used for training, validating and testing. This is often not the case in complex manufacturing pipelines, where the different order configurations would manifest themselves through varying sensor readings patterns. However, this paper proposes using the Digital Twin to its full potential by also relying on it to mirror the current operational state of the machine and its demands. By associating sensor readings to its corresponding configuration data, then Machine Learning algorithms can be safely trained and deployed on appropriate distributions of data, as instructed by the DT-provided configuration meta-data.

III. Festo Cyber Physical Factory Configuration

The Festo Cyber Physical Factory at Middlesex University is a didactic model of a product assembly line. The CP Lab consists of two production cells (or "islands") connected via an Automated Guided Vehicle (AGV – "Robotino"), and each production cell has three stations that perform individual assembly tasks. Additionally, the production cells each have a bridging station whose role is to pass the product to the AGV to transport to the next island. Figure 1 illustrates this setup and the functionalities of the stations.

Once an order is placed, a carrier tray is assigned to it for the whole duration of the process, and it travels between an island's stations via a conveyor belt. The carrier, the order number and assembly progress are checked at each station using RFID-based identification to ensure that the order is serviced once by every module.

The production process begins at the *Base cover application station*: IR sensors check if the carrier is empty or not; if it is, the base cover holder places a base cover on it. If the carrier is not empty, the station will not perform any action on the product. The second module is the *Manual Station*, where an operator is required to place the order-specific Printed Circuit Board (PCB) onto the base cover. Next is the *Camera inspection station*, where a quality check is performed to verify that the correct PCB has been selected. If the visual inspection is successful, the product is passed to the *Bridging station* where the AGV is waiting to transport the carrier to the second production cell. If the visual inspection is not successful, the order will be passed again through the first island's stations, where each of them will verify the progress of the product via RFID and act upon the order accordingly to successfully pass the camera inspection on the next attempt.

The product is then transported to the second island's *Bridging station*. Once at the *Top cover application station*, the sensors record if the product needs a top cover; if it does, a top cover is placed on the PCB board. The next station is the *Muscle press station*, whose role is to press the upper and lower parts of the product together. Using the Human-

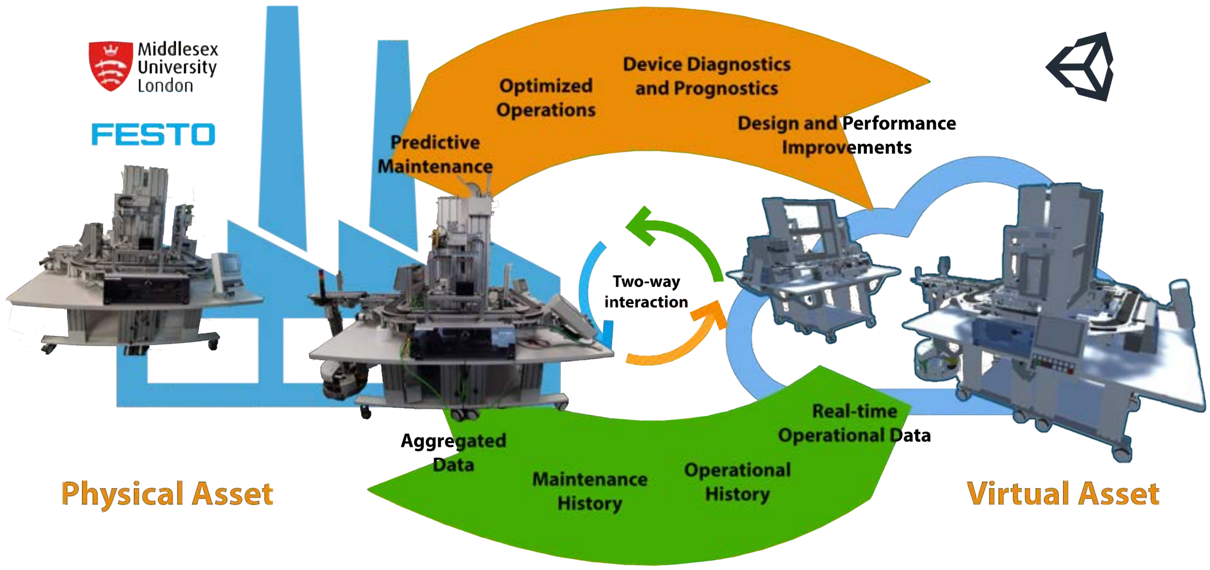

Fig. 2: The Digital Twin of the Festo Cyber Physical Factory.

Machine Interface (HMI) available, the operator can set the force that would be used to press the covers together, and the time interval for which that force is continuously applied. The last station is the *Tunnel furnace station* (or Heating station), where the product is kept for a user-defined period of time in the pre-heated furnace. The available HMI can be used to modify the time and target temperature parameters and monitor in real-time the actual temperature registered by the PT100 sensor inside the module, as well as the total time the carrier has spent inside the oven. If the current temperature is lower than the target temperature, the heating element and a cross-flow blower are turned on until the target temperature is reached, after which a count down clock will start from the user-picked time interval. When the time runs out, the carrier is released and the order is ready.

It is noteworthy that the CP Lab supports the simultaneous production of several orders, since the stations on each island are technically independent of each other. Additionally, several sensors are installed on the Cyber Physical Factory: each conveyor belt is equipped with six capacitive sensors that detect every passing carrier at all times, RFID readers and IR sensors at every station to assess the state of the order and its carrier, power consumption sensors for each island, and process-specific sensors, like temperature sensor, camera, etc. Furthermore, each station, except for the bridging stations, presents a HMI that offers monitoring, control, and configuration capabilities in real time, and is controlled via a Programmable Logic Controller (PLC).

IV. A Digital Twin framework for Predictive Maintenance

In order to build the digital twin for the Cyber Physical Factory, we first had to build its *digital shadow*. The digital shadow is a virtual replica of the system that imitates the real asset's processes, functioning, and behaviour, in a synchronous manner. Figure 2 illustrates the proposed digital twin concept of the Cyber Physical Factory, with the real and virtual asset synchronised via the two-way communication. Thus, enabling several services like storing maintenance and operational history, monitoring real-time data, and providing analytics capabilities to achieve predictive maintenance, optimised operations, and design improvements.

The first step was to build the 3D model of the Cyber Physical Factory using the CAD files of the system provided by Festo. The advantages of this approach and the format used is the the fact that the model can be broken down into individual components, therefore allowing custom characteristics to be applied to individual parts and even sensors.

The environment that has been chosen to create this digital shadow is Unity, a game engine with a C# based Application Programming Interface (API). This selection was motivated by the flexibility offered by the game engine in terms of simulation capabilities, licensing, and the lack of other developing environment-specific constraints. Within Unity, with an accurate physics engine, it is possible to simulate how a design would behave in any real-world conditions. Additionally, using C# makes it easy to interface any libraries and does not constrain the user to only using proprietary functions within Unity. Lastly, Unity is free and is not specific to any manufacturer, so building a Digital Twin framework on it makes it more accessible and the overall preferable option.

Once the 3D model of the system had been created, the next goal was to leverage the two-way communication in order to harmonize the real and digital twins. As a first step, live tracking of the product carrier has been implemented. For each station on the Cyber-Physical Lab, the data coming from capacitive sensors installed on the conveyor, along with the optical encoder measuring motor speed, is continuously being sent to the Unity Engine. As such, carrier velocity can be calculated from the conveyor motor RPM, therefore

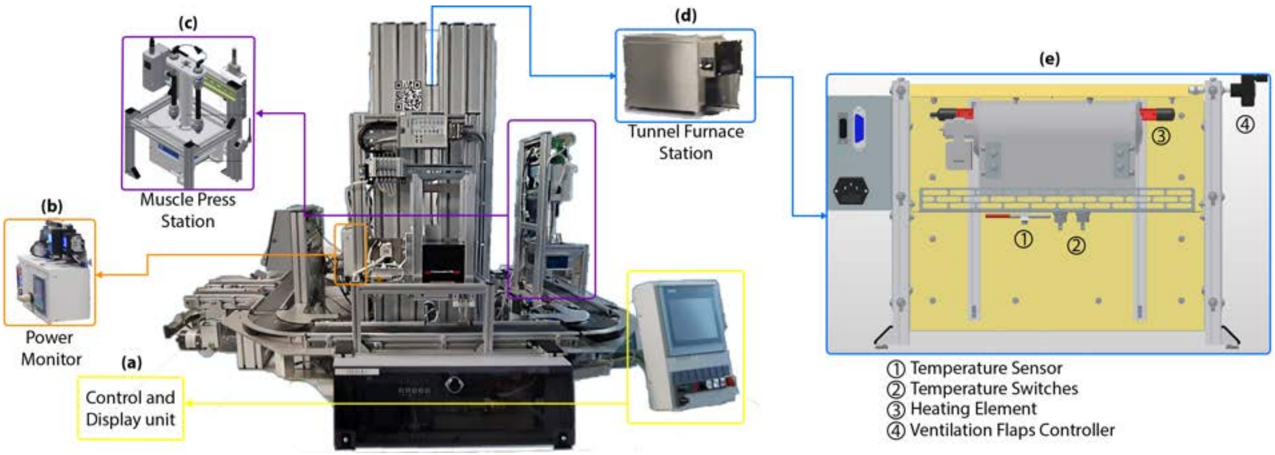

Fig. 3: The second island of the CP Lab and its Heating Station.

determining carrier position in real-time. Furthermore, since the carrier is sitting on a conveyor, slipping is probable; in this instance, the capacitive sensors data can be used to verify and update carrier position.

Orders can be placed on the Manufacturing Execution System (MES) software, the main control software for the CP Lab. The progress of the order can be viewed live via the Unity tracking control function. RFID data from the physical carriers are sent to Unity and added as a class property of the simulated carrier counterpart, creating digital twins of the carriers themselves. The data transfer between the physical asset's controllers and Unity is assured via TCP socket connections, to make the system accessible for most controllers on the market. Each PLC in the system has a server-client connection with a script running in Unity. Using TCP instead of UDP allows for certainty in the transfer of data, since it requires an acknowledgment from the recipient to guarantee the packet transmission. Data transfer frequency is approximately 50ms, depending on other system tasks. Each data packet is a delimited plain-text string that is parsed once a Unity script receives it.

It is also possible to run the software system independently of its physical counterpart. This is beneficial for testing new efficiency solutions and improvements and for generating an artificial historical data bank as a reference for predictive analytics. The data transfer protocol used works both ways, therefore Unity can send information to the physical system to change how the system functions, such as creating or editing an order, scheduling optimally-timed maintenance, and re-arranging orders to obtain maximum productivity. Consequently, the two-way interaction between the physical and the virtual asset defines our Digital Twin of the Cyber Physical Factory.

A. *Use-cases*

In order to leverage the Digital Twin for its Predictive Maintenance capabilities, an initial study has been conducted in the evaluation of the Festo Cyber Physical Factory for faults that could induce a critical failure, meaning a failure that could greatly reduce or nullify the productivity of the CP Lab.

As in any other complex system, there are several components that could malfunction, resulting in extended downtime, lower productivity, or just overall abnormal functioning. Furthermore, it is sometimes difficult to trace back the source of the failures without a proper data monitoring and logging system that can indicate when and why the error has begun developing. As such, in this work, intensive focus has been directed towards finding the failure modes that can be prevented and predicted via the data provided by the smart factory's multiple sensors, in a way that allows the detection and classification of faults leading to the effortless identification of the broken components. Amongst the potential critical failures, the one that has the most devastating impact is the erroneous triggering of the Tunnel furnace station's Safety Shutdown mechanism, which can induce immediate downtime of the entire second production cell. This mechanism is a protective measure that helps prevent fire hazards, as well as low-quality products, however it is imperative that it is triggered as rarely as possible, to reduce the inevitable hindering of production.

Figure 3 illustrates the subsystem of interest for the work conducted in this study, the second island of the Cyber Physical Factory, along with its stations, with extended focus on the Tunnel Furnace Station. A detailed functionality of the first island of the Cyber Physical Factory was introduced in [12]. Raza et al. [12] introduced the digital system for process replication of the first island of the assembly line.

The Tunnel furnace station uses a Resistance Temperature Detector (RTD) sensor, PT100, to sense the temperature emanated by the heating element in order to monitor and control the process effectively. As such, the quality of the product and the normal functioning of the station greatly depend on the health state of the heating station, in particular its heating element and temperature sensor. In case one of these components becomes unreliable, the software control mechanisms of the station will fail to stop the heating process, resulting in abnormally high temperatures being reached. For

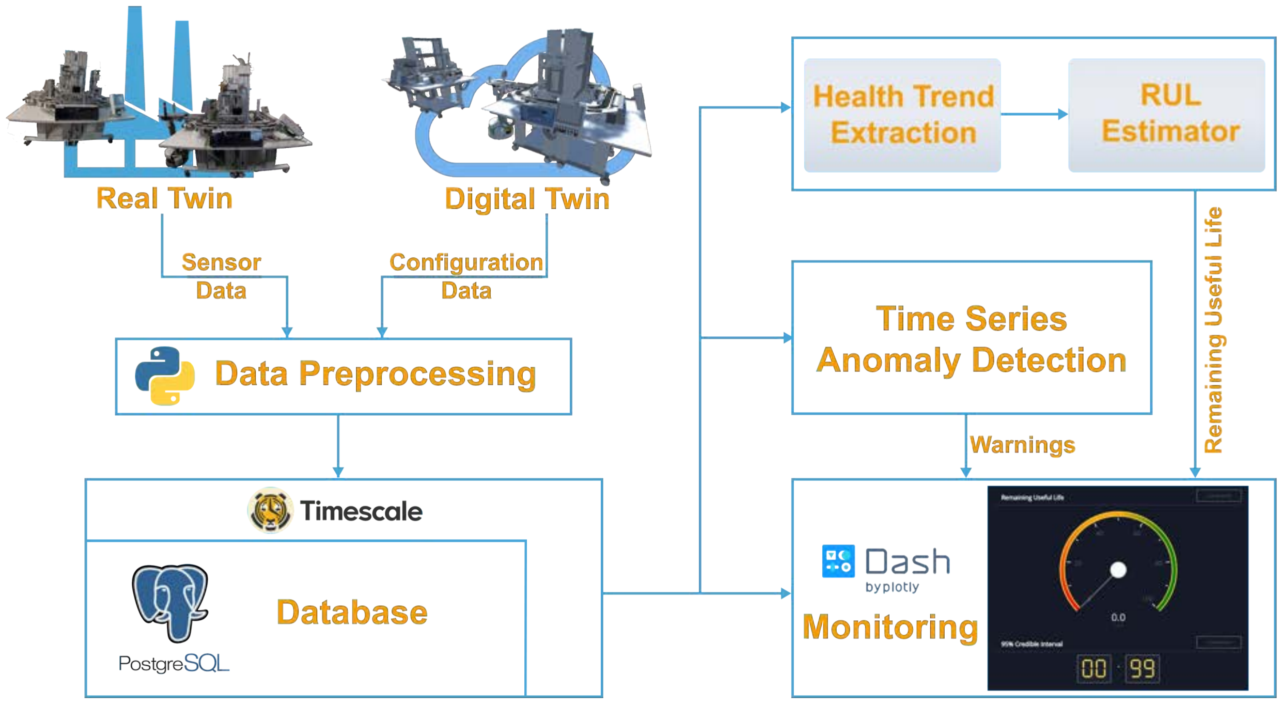

Fig. 4: A Digital Twin framework for Predictive Maintenance.

this reason, the machine is equipped with a hardware shutdown mechanism which is triggered by two temperature switches that get activated once the chamber temperature reaches 80°C. However, the Safety Shutdown mechanism can be erroneously triggered if:

- the *temperature sensor* is malfunctioning, failing to accurately gauge the temperature levels, thus allowing the heating element to increase the temperature past the user-imposed target temperature;
- the *heating element* is malfunctioning, ignoring control signals and increasing the temperature to 80°C.

In both scenarios, a fault in the temperature sensor or the heating element would manifest itself through anomalies in the power consumption patterns of the second production cell, a parameter which can be monitored via the power sensors installed on the island. Thus, the behaviour of the heating element is reflected in the power consumption signal; if the temperature of 80°C is reached, without it being intentionally set as a target temperature by the operator, then the whole second island will consume more power than what was expected for the order configurations that are active at the time.

Another potential failure mode is the degradation over time of the heating element, which could affect its capacity to properly heat the product in a time-efficient manner, such that the overall productivity of the whole Cyber Physical Factory does not fall below a certain pre-defined threshold. Such a failure could be caused by continuous usage in subpar working conditions, and it could manifest itself in either increasingly lower energy output (in which case it would take longer than usual to heat up the element to the target temperature), or, the opposite, which is gradually higher temperature output, leading to fire hazards risks, and lower heating times.

Once the potential critical errors have been defined, it is important to extract relevant data from the Cyber Physical Factory, which constitutes the real asset of the digital twin. However, for the successful identification and classification of the faults, sensor data is not enough. Additional meta-data can serve as indicators of the working regime of the smart factory, allowing the division of the other data into multiple categories, where each category is characterised by the same data distribution. For example, the power consumption patterns of the second island will differ depending on the number of active orders being carried out at its stations.

However, in order to train a data-driven model to recognise degradation patterns and accurately predict the station's Remaining Useful Life, an abundance of historical run-to-failure data would be needed. Given the absence of recordings of reactive maintenance having ever been applied to the station, faults can still be simulated through the manual ventilation flaps installed on the module, seen in Figure 3, to reflect the station's behaviour in the erroneous scenarios described above. Furthermore, the CP Lab's Digital Twin, as previously mentioned, can be used to emulate the physical asset to generate additional abnormal power consumption patterns, under various types and quantities of active product orders, to train a degradation model that can accurately and efficiently estimate the RUL.

B. Digital Twin Framework for Predictive Maintenance

With the use-cases for Predictive Maintenance powered by Digital Twin defined, this paper introduces a framework that aims to achieve optimised Predictive Maintenance by leveraging predominantly time-indexed streaming sensor data, along with configuration data coming from the digital twin of the Cyber Physical Factory. The proposed framework is illustrated in Figure 4 and consists of: the data acquisition block, the pre-processing block, the database, the time-series anomaly detection block, the RUL predictor block, and the monitoring dashboard.

1) Data Acquisition Block: The data acquisition block extracts information from both the real twin, and the digital twin. The types of data needed for the purpose of Predictive Maintenance are:

- *Sensor data* - acquired from the Cyber Physical Factory, the sensor data consists of time-indexed sensor readings from the PT100 temperature sensor installed in the baking station, the power monitor attached to the second island, as well as boolean flags readings from the capacitive position indicating sensors. The proposed framework communicates to the physical asset via the OPC UA protocol for timely machine-to-machine communication of the streaming data.
- *Configuration data* - acquired from the digital twin of the CP-Lab, it consists of meta-data that describes the configuration of the order, as well as the current load of the second island. It is important for the later stages of data analytics that the working regime of the smart factory be monitored in real time, so that the power sensor data, which describes the power consumption of all the four stations installed on the second island, can be scaled and normalised to reflect the power consumption of the heating station.

2) Data preprocessing Block: The data preprocessing block normalises the data if necessary, then it inserts it into the database.

3) Database Block: The database of the system consists of the relational PostgreSQL, which was optimised for time-series analysis through the TimescaleDB extension, allowing fast inserts of high-rate streaming data.

4) Time series Anomaly Detection: This block deals with identifying odd behaviour in the time-series sensor data, with respect to the expected behaviour based on the configuration data coming from the digital twin. The anomaly detection block's responsibilities are also to detect outliers, change-points, and transmit warnings with various levels of urgency based on the found anomalies.

5) RUL Predictor Block: The Remaining Useful Life Predictor block is functionally split into two subsystems:

- *Health Trend Extraction* - deals with analyzing the sensor and configuration data to extract a health trend. In order to qualify for the use Predictive Maintenance, the extracted health trend must fulfill (as much as possible) certain conditions: it must be a function of time, it has to be monotonically increasing (or decreasing) over time, and, preferably, it has to be an injection. Popular methods that combine features, depending on whether they can be modelled linearly or not, are Principal Component Analysis (PCA) and Isomap.
- *RUL Estimator* - extracts the Remaining Useful Life of the component, based on the current health status indicated by the health trend. A Machine Learning model is fit to the Health Trend, and it is updated for every coming sample. The Machine Learning algorithm that predicts the RUL will be chosen based on the characteristics of the health trend. In many cases, the RUL extraction can boil down to a simple curve fitting and regression technique, if the health trend respects the aforementioned conditions of monotonicity and dependence on time. However, if it does not, there are multiple other options which can be applied, like Recurrent Neural Networks (RNNs), Long Short Term Memory networks (LSTMs), Support Vector Regression (SVR), etc., depending on the characteristics of the pre-processed data. It is worth mentioning that, even though it is preferred that the RUL be a function of time, it can also be expressed in ”cycles” or ”orders”. For example, the output of the RUL Estimator block can indicate that the machine would fail in a certain number of cycles, if a given sequence of order configurations is to be serviced after the moment of prediction.

6) Monitoring Block: The Monitoring Block is a web application built with Dash, which displays continuous comprehensive information about the current state of the real asset. It also presents controls that can be communicated back to the physical twin for engineered process optimisation.

C. Challenges

One of the most pressing challenges of building a digital twin for the Festo Cyber Physical Factory, which is a sturdy and reliable system, is the lack of historical failure data to be exploited for predictive maintenance purposes. The smart factory's components have long lifespans, and, unlike in a real manufacturing factory, the working environment is not necessarily harsh enough to have a significant impact on the degradation of the system. For this system and others like it, it is important to find a digital twin solution that can be used to generate synthetic data sets to describe faulty behaviour. In this context, another significant challenge is creating artificial time series data that accurately reflects the real twin's operation in naturally occurring erroneous scenarios. While various intentional errors can be easily inserted into the data set used to train the RUL Estimator, the resulting RUL may not be a credible reference for the actual remaining useful life of the system. As in any other Machine Learning application, the output is only as good as the data used for training and testing.

On the same note, another important challenge is the validation of the results. In order to verify that the proposed framework and, ultimately, the Predictive Maintenance algorithm are accurate, meaning that the real remaining useful life of the system is indeed correctly given by the digital twin-provided RUL parameter, the real twin must fail naturally at least once.

It is equally important to consider the limitations of this framework. One of the most challenging ones is the fact that this framework will be able to predict only failure modes that it has been made aware of beforehand through training data, real or simulated. The Anomaly Detection block aims to compensate for this issue by providing warnings to the operator when the machine behaves in a new, odd pattern that it has never seen before. Of course, this implies that the Anomaly Detection block will be trained to recognise only normal functioning behaviour and known failure modes, treating any other behaviour as anomalies. However, it also needs to account for change point detection of a known pattern, and advise the operator accordingly.

V. Conclusions and Future Work

In this work, a framework for Predictive Maintenance using Digital Twin of a Festo Cyber Physical Factory has been proposed as a tool that can help drive manufacturing processes in the Industry 4.0. First, the real smart factory has been described with an emphasis on the processes involved in the creation and completion of an order. Then, the initial stages of digital twin development have been presented, namely the 3D modeling of the real asset in a game engine environment such as Unity, detailing also the synchronization of the twins with the purpose of live tracking of orders via the digital twin. Next, a framework based on the digital twin was proposed for achieving predictive maintenance. The introduced architecture uses real data coming from the CP Lab for monitoring and training purposes, along with configuration data provided by the digital twin to establish its working regime, so that the RUL of the target system, the Tunnel Furnace Station, can be effectively singled out. Lastly, the active challenges that were encountered in this study have been presented.

Future work will look into the validation and the performance evaluation of the proposed solution. More specifically, extensive focus will be placed on building simulation models that can mimic the signals coming from the temperature and power sensors installed on the Tunnel Furnace Station. The simulators' outputs will then be compared with the real data coming from the Heating Station under different order configurations. Once accurate synthetic data is obtained, we will look into using it together with real data to build the Predictive Maintenance model by following the framework presented in this work. This model's performance will be validated by the real twin, in the case of an eventual, naturally occurring failure. Once these feats have been accomplished, the Digital Twin will be completely integrated within the lifecycle of the Festo Cyber Physical Factory, and it will be able to provide insights into its failure modes.

VI. Acknowledgements

This work was supported by the UK-India Education and Research Initiative (UKIERI) and the Department of Science and Technology (India) under Grant DST UKIERI-2018-19-011.

References

[1] "Simplifying 5G with the Network Digital Twin", Spirent. Accessed March 16th, 2020. [Online]. Available: https://www.spirent.com/assets/wp/wp_simplifying-5g-with-the-network-digital-twin

[2] S. Ferreiro, E, Konde, and A. Prado, "Industry 4.0: Predictive Intelligent Maintenance for Production Equipment", *European Conference of the Prognostics and Health Management Society*, pp. 1–8, 2016.

[3] S. Wang, Y.C. Liang, W.D. Li, and X.T. Cai, "Big Data enabled Intelligent Immune System for energy efficient manufacturing management", *Journal of cleaner production*, vol. 195, pp. 507-520, September 2018.

[4] Y.C. Liang, S. Wang, W.D. Li, and X. Lu, "Data-Driven Anomaly Diagnosis for Machining Processes", *Engineering*, vol. 5, issue 4, pp. 646-652, August 2019.

[5] Y.C. Liang, W.D. Li, X. Lu, and S. Wang, "Fog computing and convolutional neural network enabled prognosis for machining process optimization", *Journal of Manufacturing Systems*, vol. 52, pp. 32-42, July 2019.

[6] Y. Hsieh, F. Cheng, H. Huang, C. Wang, S. Wang and H. Yang, "VM-Based Baseline Predictive Maintenance Scheme", *IEEE Transactions on Semiconductor Manufacturing,* vol. 26, no.1, pp. 132-144, Feb. 2013.

[7] I.V. Ortega, "Development of a Digital Twin to support machine prognostics with low availability of run-to-failure data", MSc. Thesis, Sch. of Industrial and Information Engineering, Politecnico di Milano, Milan, Italy, 2019.

[8] P. Aivaliotis, K. Georgoulias and G. Chryssolouris, "A RUL Calculation Approach Based on Physical-based Simulation Models for Predictive Maintenance", 2017 *International Conference on Engineering Technology and Innovation (ICE/ITMC)*, Funchal, pp. 1243-1246, 2017.

[9] A. Barthelmey, E. Lee, R. Hana and J. Deuse, "Dynamic digital twin for predictive maintenance in flexible production systems", *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, Lisbon, Portugal, pp. 4209-4214, 2019.

[10] P. Aivaliotis, K. Georgoulias and K. Alexopoulos, "Using digital twin for maintenance applications in manufacturing: State of the Art and Gap analysis", *2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pp. 1-5, 2019.

[11] Y. Xu, Y. Sun, X. Liu and Y. Zheng, "A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning," *IEEE Access*, vol. 7, pp. 19990-19999, 2019.

[12] M. Raza, P. M. Kumar, D. V. Hung, W. Davis, H. Nguyen and R. Trestian, "A Digital Twin Framework for Industry 4.0 Enabling Next-Gen Manufacturing," *2020 9th International Conference on Industrial Technology and Management (ICITM)*, Oxford, United Kingdom, pp.73-77, 2020.