O. Serradilla, E. Zugasti, J. R. de Okariz, J. Rodriguez, eta U. Zurutuza, «Methodology for data-driven predictive maintenance models design, development and implementation on manufacturing guided by domain knowledge», *Int. J. Comput. Integr. Manuf.*, 2022, doi: 10.1080/0951192X.2022.2043562

Methodology for data-driven predictive maintenance models design, development and implementation on manufacturing guided by domain knowledge

Submitted to: International Journal of Computer Integrated Manufacturing

Oscar Serradilla[a]*, Ekhi Zugasti[a], Julian Ramirez de Okariz[b], Jon Rodriguez[b] and Urko Zurutuza[a]

[a]*Mondragon Unibertsitatea, electronics and computing department, Loramendi 4, Mondragon (20500), Spain;* [b]*Koniker, San Andres auzoa 20, Mondragon (20500), Spain.*

* Corresponding author: Oscar Serradilla[a]; email: oserradilla@mondragon.edu; orcid: 0000-0003-2557-9859

Ekhi Zugasti[a]; email: ezugasti@mondragon.edu; orcid: 0000-0001-8506-5695

Julian Ramirez de Okariz[b]; email: j.rokariz@koniker.coop

Jon Rodriguez[b]; email: j.rodriguez@koniker.coop

Urko Zurutuza[a]; email: uzurutuza@mondragon.edu; orcid: 0000-0003-3720-6048

Word count (including captions and new modifications): 14749

Methodology for data-driven predictive maintenance models design, development and implementation on manufacturing guided by domain knowledge

The 4th industrial revolution has connected machines and industrial plants, facilitating process monitoring and the implementation of predictive maintenance systems (PdM) that can save up to 60% of maintenance costs. Nowadays, most PdM research is carried out with expert systems and data-driven algorithms, but it is mainly focused on improving the results of reference simulation datasets. Hence, industrial requirements are not commonly addressed, and there is no guiding methodology for their implementation in real PdM use-cases. The objective of this work is to present a methodology for PdM application in industrial companies by combining data-driven techniques with domain knowledge. It defines sequentially ordered stages, steps and tasks to facilitate the design, development and implementation of PdM systems according to business and process characteristics. It also facilitates the collaboration among the required working profiles and defines deliverables. It is designed in a flexible and iterative way, combining standards, state-of-the-art methodologies and referent works of the field. Finally, the proposed methodology is validated on two use-cases: a bushing testbed and a press machine of the production line. These use-cases aim to facilitate, guide, and speed up the implementation of the methodology on other PdM use-cases.

Keywords: predictive maintenance; methodology; data-driven; domain knowledge; manufacturing

1. Introduction

The industrial sector has been in constant development since the creation of the first machine, being classified into different revolutions according to the technologies that drove relevant changes. The first brought mechanisation with water power and steam power, the second brought mass manufacturing by product lines and electric energy and the third, which was related to digital revolution based on Information Technology (IT) and electronics, brought production automation Lukac (2016). All of them have brought benefits to the society by creating more affordable and accessible quality products that improved average quality life and boosted companies' production while optimising costs.

Nowadays, there is a transition towards the fourth revolution denominated as industry 4.0, which is driven by Cyber Physical Systems (CPS), Industrial Internet of Things (IIoT), Big Data, Cloud Computing and Remote Sensing Frank et al. (2019). Its objective is to improve industrial processes and adapt to their requirements using software, sensors and intelligent control units as described by Lukac (2016). The integration of these components forms infrastructures that collect much data in data lakes, which companies believe it can enhance their competitivity. This process-related data can be combined with big data and data analysis techniques to solve numerous problems; common applications in industry are listed below: Predictive Maintenance, product quality inspection, manufacturing parameters optimisation and energy consumption optimisation Frank et al. (2019).

Studies like Dhillon (2002) present that effective maintenance can reduce industrial company costs up to 50% by correcting failures of machines, systems and people. Currently, most industrial companies use the following maintenance approaches: either preventive/time-based maintenance that apply periodical interventions to avoid failures, or corrective maintenance that waits until failures occur to apply interventions. However, these maintenance strategies have a big optimisation potential. On the one hand, component working life can be extended by taking advantage of their unexploited correct working time before failure, which reduces downtime and replacement costs. On the other hand, replacing components before failure would avoid expensive breakdowns whose reparation is much more expensive than components cost given maintenance costs such as repairing tasks and production time loss. These are the gaps that Predictive Maintenance aims to address Selcuk (2017).

Concretely, Predictive Maintenance has a huge potential for manufacturing companies, which can achieve an Overall Equipment Effectiveness (OEE) over 90% by increasing asset availability and performance as stated by Coleman et al. (2017), and a 1000% increase in Return on Investment (ROI), indicated by Lavi (2018). It is the most advanced level of maintenance, and often the most efficient, allowing strategic decision making. These systems have been applied in many fields such as automotive, manufacturing, service and other industry domains (Prajapati et al. 2012; Selcuk 2017) by anticipating and attending failures to ensure smooth operation, enhancing total productive maintenance, improving OEE, safety and protecting the environment. The development of data-driven Predictive Maintenance systems is based on the data collected from CPS.

To date, many companies have optimised their costs using data-driven and computational models or increased sales using recommender systems, but according to the survey conducted by, only around 20% are deployed to production. Nonetheless, even though there are many publications of data-driven model application for maintenance, many manufacturing companies have not implemented Predictive Maintenance systems yet given they do not know their potential business benefit and cannot see a clear ROI. Nowadays, the application of data-driven techniques for manufacturing process improvement is easier and cheaper than ever before, given the accessibility to machine data monitoring systems and online platforms. Even studies like Ransbotham et al. (2017) expect artificial intelligence to have noticeable impact in manufacturing by 2022. Predictive Maintenance is expected to boost industrial companies' competitiveness by enabling data-driven process understanding and Predictive Maintenance strategies for decision making.

Predictive Maintenance (PdM) is a widely researched topic in academia. Conversely, the implementation of these systems in industrial companies is slower and resource and time costly. Some important reasons for these facts are the absence of knowledge and baggage of industrial companies in this field, data variability and unstructured development. Many state-of-the-art Predictive Maintenance works are developed based on experience and use it to justify undertaken decisions. However, no systematic forms or methodologies have been found to guide the development of data-driven PdM models supported on domain knowledge from scratch, that can be used in the implementation of other use-cases or systems to reduce uncertainty or arbitrary decisions.

1.1 Contribution

The main objective of this article is to propose a methodology to implement Predictive Maintenance data-driven models guided by domain knowledge. This will boost industrial and manufacturing companies by systemising model development steps adapted to their requirements, enabling maintenance optimisation and knowledge discovery in industrial machinery. In addition, this paper presents three secondary objectives. (1) summarise the most common problems industrial companies face in the application of Predictive Maintenance systems application (2) present the resources, profiles and strategies that enable structured and agile PdM implementation (3) define the methodology, specified into simple and flexible steps, containing necessary steps for model development. The hypothesis of this work is that despite each company and PdM system has its own characteristics, a general flexible enough methodology can be defined as reference, so that projects and companies can follow it to address their requirements. Hence, the presented methodology will systematise PdM system development, resulting into projects costs saving by planification and agile model development.

This work is aimed at different working profiles such as data scientists, business managers and domain technicians. They will learn about which are the resources, general steps and methodology to deploy a PdM model in their company successfully. This can help them understand project requirements like which profiles are necessary in each part of the process and thus facilitate project planning.

This work is organised as follows. Section 2 reviews Predictive Maintenance background and manufacturing companies' requirements. Section 3 reviews and categorises state-of-the-art models for Predictive Maintenance and analyses current Predictive Maintenance standards and methodologies, presenting their gaps. Section 4 explains the main contribution of this work: the proposed PdM methodology, complemented with step-by-step indications to guide its implementation. Section 5 presents the adoption of the methodology and adaptation to successfully meet requirements of a real use-case. Finally, Section 6 concludes this work by highlighting its most relevant aspects and gaps addressed by the proposed methodology, together with and prospect on how it could impact future PdM models development.

2. Background

This section summarises the basic context and background of manufacturing companies and Predictive Maintenance. It is useful to understand the basis and problems that the proposed methodology aims to address.

2.1. Manufacturing companies: context and requirements

Currently, many manufacturing companies are transitioning their data collection systems to cloud platforms with the objective to create machines that are self-aware, increasing their overall performance and facilitating maintenance management Vaidya et al. (2018). This transition by the adoption of data analysis and intelligent maintenance application enables data collection of relevant assets that facilitates implementing Predictive Maintenance and proactive maintenance strategies. Nonetheless, some challenges must be commonly faced by companies in this process, some of which are highlighted below.

According to Guzman and Rodriguez (2014), there are three failure types classified by the stage of asset life. In the initial working period, the failures can be caused by improper mounting or component defects. Once machine working and operation has stabilised, random failures arise due to: inadequate working Environmental and Operational Conditions (EOC); accidents; and change of necessary parts, among other reasons. In final machine or asset working period, the majority of failures are caused by natural wear; arising clearances and imbalances.

Usually, an industrial company that worked hard to achieve a good and competitive share of the market, produces machinery that rarely fails; given that the engineering work behind the design, and mounting is usually optimal. This type of companies tends to have very little historical failure data; which makes difficult for data scientists to gather the data. Even so, when problems or failures arise, delays and stops happen in machines and production lines, resulting in companies' loss. To prevent failures, these companies commonly apply periodical maintenance. Nonetheless, failure anticipation and prevention can save maintenance costs caused by many reasons like production time loss, component replacement, repair tasks or production delays.

Industrial companies mainly use two data collection methods, classified into real machine and simulation. On the one hand, real machine data is retrieved either manually by an operator or automatically from a CPS-based platform; like the one proposed in the Mantis project by Albano et al. (2018). On the other hand, simulations create data aimed to simulate real machine behaviour under different conditions; which are based on physical testbeds such as the ones presented by Borodulin et al. (2017), or Digital Twins (Aivaliotis et al. 2019; Alexopoulos et al. 2020), which are software simulations based on techniques like theoretical domain knowledge, finite element method, data mining or statistics. These simulation techniques can generate artificial failure data in industrial scenarios, where failure data is usually scarce. However, the results obtained from simulation platforms may differ from real machine behaviour. For this reason, some models are first validated with simulated data and then applied to real machines, using real data to train and test the system.

In any case, production machine data acquisition and visualisation are highly recommended, this can be done in several ways: firstly, online collection permits interactions with acquisition and online modifications, despite being more expensive. Secondly, offline collection offers the best trade-off regarding costs and implementation difficulty. Finally, manual collection is intractable when dealing with big amount of data, although it can be useful for small datasets. Another important concern regarding data collection is assuring customer privacy, since these seek to protect their data and knowledge from competence. Therefore, customer data must be transmitted, stored and analysed securely, avoiding data leakage to third parties and thus ensuring data analysis is beneficial for both, manufacturers and their customers.

2.2. Predictive Maintenance context, opportunities and challenges

Predictive Maintenance aims to keep assets working correctly and only apply maintenance actions when these stop working properly. This maintenance strategy extends components' working life with respect to periodical maintenance, while preventing damages by scheduling interventions before failures occur; concretely before corrective maintenance, as stated by UESystems (2019). The main benefits of PdM over other maintenance strategies are: maintenance cost reduction by avoiding unplanned maintenance; increasing the working life of assets by anticipating to failures and performing maintenance before they occur; and improving operational safety by preventing failures. In addition, OEE can be enhanced by downtime reduction and productivity increase.

Predictive Maintenance consists of monitoring assets behaviour and comparing it with previously analysed working conditions. This historical data is used to search for similar patterns and predict condition trend evolution, anticipating maintenance requirements in early stages by detecting failures and degradation. The concept of PdM is a step further from the Condition Based Monitoring (CBM), the application of prognosis and health management in the maintenance field that arose in 1940s Wiseman (2006). However, nowadays PdM is more accessible than ever before given that current technology permits its automatising.

PdM enables performing maintenance in a proactive way, based on anomaly detection, diagnosis and prognosis techniques. Concretely, its power lies on deploying a system that continuously monitors assets health to give data-driven maintenance information and recommendations to maintenance technicians. Historical failure and maintenance data is useful to learn from previous experiences and prospect on future events. Many works present component degradation patterns in a plot denominated P-F curve like UESystems (2019), where health decreases from healthy working condition until failure as time or machine cycles go by.

Predictive Maintenance architectures must deal with many factors, peculiarities, and challenges of industrial data; the most relevant ones are discussed below. One principal challenge of industrial data is its behaviour and data variability among assets. Even two identically produced machines vary given mechanical tolerances, mount adjustments, variations in EOC and many other factors. These difficult the creation of robust models, which hinders the reusability of PdM models among machines and assets. In addition, gathering high quality data is difficult given that sensors can be misplaced or damaged, the installed sensor types might not be suitable for the use-case, an adequate sampling frequency may not be adopted, additional context data is missing etc. In this scenario, building a training database that enables robust data analysis is one of the main challenges. After ensuring high quality data is collected, an adequate preprocessing and filtering must be performed to remove distorting noise that environmental factors create, with the objective to prepare the data for the data-driven model. Moreover, with the increase of collected data dimensionality, the complexity and resources required to process it grow. Another challenge of PdM is that industrial companies have difficulty on obtaining failure data as they take care of the machines to ensure their right working condition. Moreover, sometimes it is troublesome to identify, define and prioritise the use-case to solve.

Regarding monitored EOC data, environmental conditions refer to external conditions that influence the assets, like ambient temperature or surrounding vibration perturbations; operational conditions are technical specifications assigned to working processes, such as desired speed, force and positions. Additionally, sensor data comes from measurements taken by machine sensors, and derived variables are calculated by combination of the aforementioned ones or by converting them to other units of measurement.

Several components are more prone to be studied and applied PdM than the rest given they play a key role in industrial processes and suffer from higher failure rates. The article by Zhang et al. (2019) study some of these components like bearings, blades, engines, valves, gears and cutting tools . In addition, several common target failure types for condition monitoring are imbalance cracks, fatigue, abrasive and corrosion wear, rubbing, defects and leak detection. The publication by Li and Gao (2010) classifies the possible failure types of systems by its cause: component failure, environmental impact, human mistakes and procedure handling. Besides, there are many different condition monitoring techniques that enable data acquisition from the components, like vibration analysis, ultrasound, temperature and many more, like stated by Serradilla, Zugasti, and Zurutuza (2020).

PdM based on condition monitoring is possible given that the data follows patterns such as: trends, seasonality and noise, as indicated by Brownlee (2018). Analysing more than one variable in the same dataset is known as multivariate problem. In this situation, the variables are commonly analysed together, adding more context and complementary information, but at the same time, complexity.

The paper by Venkatasubramanian et al. (2003) presents 10 desirable characteristics a PdM system should have to address industrial requirements: “quick detection and diagnosis, isolability (distinguish among different failure types), robustness, novelty identifiability, classification error estimation, adaptability, explanation facility, minimal modelling requirements, real-time computation and storage handling, multiple fault identifiability”. It can be useful to give perspective when choosing PdM architectures and models adapted to address use-case requirements.

3. Related works

This section presents state-of-the-art and related PdM works containing techniques and methodologies that facilitate their development based on data-driven models. It helps to relate contributions of this work with widely reviewed current publications.

3.1. Predictive Maintenance modelling

The implementation of Predictive Maintenance models is a widely researched area. These can be classified by the amount of knowledge and data they require to be created (Liao and Köttig 2016) as presented in Table 1.

Table 1 Summary of PdM state-of-the-art works according to their technique

PdM approach	Characteristics	Techniques and references
Physical and knowledge-based methods	• Require domain knowledge • Mathematical models that are developed by domain experts • Linked to physical models • Require in-depth understanding of how the system works • White box, easy to understand • Have difficulties to model complex systems	First principle modelling Zhao and Magoulès (2012), parameter estimation Okoh et al. (2017), IF-ELSEs Vlasov et al. (2018), fuzzy-logic techniques Alvares and Gudwin (2019).
Data-driven methods	• Predict the systems state by monitoring their condition with methods that learned from historical data • Suitable for complex systems given they do not require to understand how physical processes work • They result in white, grey or black box approaches according to the used techniques • The most complex methods, which are usually the most accurate ones, are more difficult to interpret • They require a higher amount of data	Statistical methods Able et al. (2016), reliability functions Zhou et al. (2007), artificial intelligence methods Yuan et al. (2016). For instance, fuzzy rule-based Diez-Olivan et al. (2017), Bayesian Networks Ansari et al. (2020), deep learning Malhotra et al. (2016), and tree ensemble based like XGBoost Cerqueira et al. (2016) or Random Forest Dos Santos et al. (2017).
Hybrid methods	• Combine domain knowledge-based models with data-driven techniques, sharing characteristics of both • They aim to achieve accurate models whose estimations are linkable to physical meaning • Merging different models increases complexity	Combine physical/knowledge-based and data-driven methods for RUL estimation Liao and Köttig (2016), hybrid model that qualifies uncertainty for RUL estimation Zhao et al. (2013), domain technicians create model for failure detection and prediction with water pipe data Li and Wang (2018), and combine knowledge-based rules systems with data-driven methods to refine them Cao et al. (2020).

Physical/knowledge-based and data-driven systems both have proven to work well for many different use-cases, but choosing one over the other and even selecting a specific architecture should be relative to use-case requirements. Nonetheless, lately data-driven techniques have gained popularity given their easiness to automatise and higher accuracy in modelling complex systems. Therefore, creating accurate data-driven systems guided by domain technicians to embed expert knowledge could combine virtues of both approaches. This approach can simplify the lifecycle of PdM architectures, facilitating their design, development, deployment, and monitoring.

3.2. Methodologies and standards for Predictive Maintenance

There are a great number of ways to create data-driven Predictive Maintenance systems. Nonetheless, most of them generally address one or many of the incremental stages presented in the following reference architecture. The first step is anomaly detection, the second is diagnosing the anomaly, the third is prognosis of anomalies evolution and the fourth and last step is mitigation, like presented by Welz (2017) and is defined in the PdM roadmap of Figure 1.

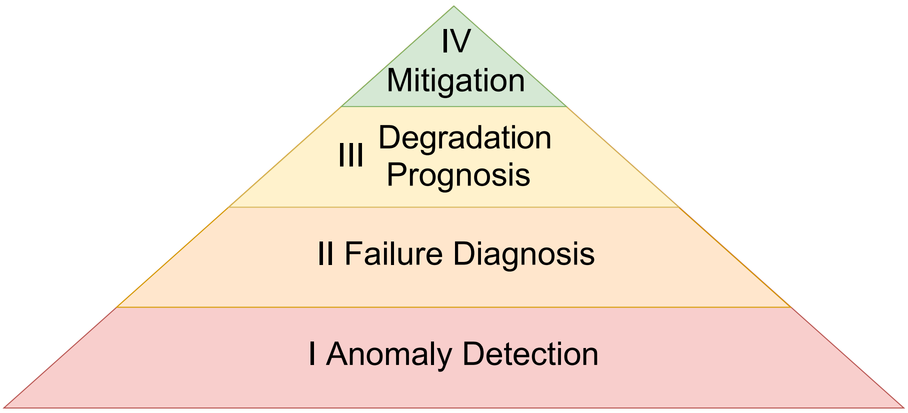

Figure 1 PdM roadmap by Serradilla, Zugasti, and Zurutuza (2020).

The standardisation, specifications and guidelines on manufacturing and PdM are gathered in norms and standards. The norm UNE-EN 13306 (2018) defines maintenance and maintenance management terms, which enable the understanding of key PdM concepts. It also classifies existing maintenance strategies in a tree scheme regarding their underlying technique, summarises the operating time distribution in down state and the sub-states of up state, and defines how to address intelligent maintenance by prioritising the most critical failures. The Open System Architecture for

Condition Based Maintenance (OSA-CBM) specification by Lebold et al. (2002) is a standard information flow architecture for CBM systems, on which the aforementioned reference architecture is based. It proposes XML schemes to facilitate CBM implementation and 7 layers of information flow. The international standard ISO 17359 (2018) is based on other standards, and provides a guideline for condition monitoring and machine diagnosis based on sensor data. It presents a procedure for implementing PdM on a schematic flowchart, divided in key steps that complement the stated reference architecture. Moreover, O'Donovan et al. (2015) presents a set of data and system requirements for implementing equipment maintenance applications of smart manufacturing in industrial environments. It also presents an information system model that provides a scalable and fault tolerant big data pipeline for integrating, processing, and analysing industrial equipment data.

Many publications present their own variations and frameworks that complement the reference architecture and stated standards, some relevant ones are presented below. Rana, Kumar and Srivdya (2016) present a guideline to help companies in the analysis of the most suitable maintenance strategy for each use-case. As they state, two relevant aspects are the analysis of data collection method and failure mode effects and criticality analysis (FMECA) Jordan (1972). If the Predictive Maintenance turns out to be most optimal, there exist other works that recommend the stages or suitable steps to facilitate its implementation, as the ones presented below.

The article published by Deloitte Institute (2017) presents a PdM journey in levels. Level 0 provides a maintenance based on reactive resolution of failures, level 1 is based on visualisation, level 2 depends on rules created by expert-knowledge, level 3 is supported on data-driven anomaly detection, level 4 is based on prognosis models and the 5th and last level is based on the identification and mitigation of the anomaly using Root Cause Analysis (RCA).

The review by Khan and Yairi (2018) proposes a data-driven flowchart methodology to successfully apply PdM. It is supported on OSA-CBM, PdM reference guidelines, architectures and previously reviewed articles. Moreover, it states that the difference between traditional data-driven, statistical and traditional Machine Learning (ML) in comparison with deep-learning methods is that the latter one feeds the model with preprocessed data given it can extract features directly whereas the former ones also need feature selection and extraction. Conversely, this article is focused on the capability of deep learning models' application in system health management.

Moreover, Nuñez and Borsato (2018) presents an ontological model to guide the implementation of expert systems for prognosis and health management. They demonstrate its applicability by implementing an expert system that models the RUL of a mechanical machine before entering into a functional failure.

Furthermore, the publication by Bousdekis et al. (2015) proposes how diagnosis, prognosis and decision support is influenced by company management and the relation of these tasks with maintenance management. It also explains how each PdM stage can influence or result into maintenance actions. This is the way PdM systems impact company maintenance by giving recommendations and prospect.

In addition, there are other methodologies designed to handle general machine learning life-cycle that can also be used for PdM model development. One of principal ones is CRISP-DM by Chapman et al. (2000), even though it is not specifically designed for maintenance and therefore it does not consider how to handle industrial requirements.

All in all, related methodological publications are either: about CBM, focused on details of data-driven model types, roadmaps that show trends and highlight future directions, focused on technical aspects of one or two PdM steps, overlook the importance of domain expertise, not specific for PdM and therefore not adapted to its characteristics, or are not developed to address industrial companies' requirements. For these reasons, many publications end up following their own steps to address their use-cases.

A research on electronic databases including Scopus, Engineering Village, Springer Link, Science Direct and the search engine google scholar was performed in the time period between 2011 and 2021, to search for articles with the terms "predictive maintenance" AND "methodology" AND ("data-driven" OR "life-cycle" OR "development"). This research reported no methodology for data-driven Predictive Maintenance systems application that details: their design, development and implementation, defining the required steps and resources, while specifying how to combine data-driven models with domain knowledge adapted to industrial use-case requirements.

4. MEDADEK-PdM

The main contribution of this paper is the *MEthodology for DAta-Driven techniques and Expert Knowledge combination for PreDictive Maintenance* (MEDADEK-PdM) presented in Figure 2. It contains the general stages and main steps to guide manufacturing companies in the design, development and deployment of data-driven PdM systems. It is open and modular, being flexible and adaptable to address different industrial use-case requirements iteratively while keeping simple to facilitate its implementation and understanding. Therefore, it enables the addition of new steps adapted to each use-case requirements while permitting the omission of the steps without asterisk, which are advisable yet not indispensable.

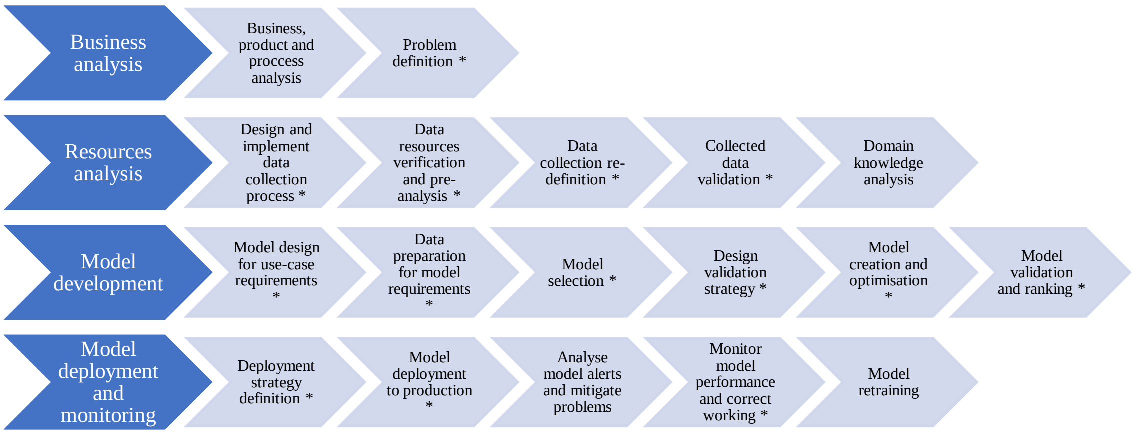

Figure 2 Scheme of proposed data-driven PdM methodology. Consists of four stages and their steps.

Moreover, this methodology presents the tasks required to complete the steps, the required worker profiles to succeed in its application, and specifies which deliverables are generated by the end of each stage. Concretely, three main working profiles collaborate in the implementation of this methodology: business profile contributes with business vision and leads problem definition, domain technician contributes with domain expertise supervising the project and collaborating in tasks, and data-scientist guides the project and handles the tasks related to data-driven model development, deployment, and monitoring. The complete step-by-step version of the methodology is presented in APPENDIX A.

This diagram is inspired by CRISP-DM data analysis methodology as described by Chapman et al. (2000) and PdM standards and methodologies like the ones presented in the article by Reliasoft Corporation (2007), and the ones presented in Section 3.2.

Methodologies and standards for Predictive Maintenance. It contains four main sequential groups, denominated as stages: *business analysis*, *resources analysis*, *model development* and *model deployment and monitoring*. Its iterative design facilitates and fastens models going into production, and then promotes incremental iterations to enhance their performance.

4.1. Business analysis

Business analysis is the first stage of the methodology, which is composed by two steps.

- *Business, product and proccess analysis*
  The first step is to perform *business, product and proccess analysis* adopting a business vision by understanding the products and services the company offers, its business model and how they are related to the manufacturing proccess. It enables to prioritise among problems that address business requirements, such as production parameters optimisation, quality control and choosing the suitable maintenance strategy.
- *Problem definition*
  If previous step's analysis results in prioritising maintenance optimisation, the assets with highest impact have to be evaluated, starting with the most common and critical failure types identified with tools like Failure Mode Effects Analysis (FMEA) and FMECA by Jordan (1972). After that, the most suitable maintenance strategy for each asset has to be evaluated: whether Predictive Maintenance, periodical maintenance or corrective maintenance; thus completing *problem definition*. A good resource for this analysis is the guideline by Rana, Kumar and Srivdya (2016). Finally, if Predictive Maintenance is the maintenance strategy to be implemented, then the methodology advances to the second stage; which focuses on analysing use-case resources.

Business analysis is the only stage where the three required profiles for the methodology work together for its completion: the business manager provides business perspective and helps to define business requirements; the domain technician provides technical and operating expertise; and the data scientist helps to guide this stage while learning background from other profiles. As a result of completing this stage, two documents can be created: Document 1.1 resources characteristics contains a summary of how business works, business model, products and services, and how manufacturing process works; and Document 1.2 problem definition, which explains manufacturing critical assets ordered by impact, which are maintenance requirements for these assets and components, and analysis of maintenance strategy suitability for them.

4.2. Resources analysis

After understanding the business and defining use-cases, the second stage is *resources analysis* for the problem.

- *Design and implement data collection process*
  The first step is to *design and implement data collection process*. The data should be collected initially under similar EOCs that want to be modelled from the same machines and assets through time. Comparisons among different machines/assets of the same type, even if being built under the same specifications, is difficult given differences in EOCs, tolerances, adjustments, etc. EOC information is essential to give data-driven algorithms working context of the monitored assets, like environmental perturbations or working conditions that can affect its performance and result in anomalies or component degradation. Whereas monitored EOCs can boost model accuracy, there exist other relevant variables that are not controlled or even monitored like environmental noise and disturbance, lack of sensors or their misplacement. Missing information like this that influences the monitored process adds uncertainty, and a result, the created PdM system will be less accurate.
  In addition, sampling rate of variables is of utter importance.
  According to the Nyquist-Shannon sampling theorem, a signal of unknown frequency locations has to be sampled at least at 2 times its frequency in order to enable signal reconstruction, thus maintaining enough information to avoid non reversible information loss by the aliasing effect, presented by Mishali and Eldar (2009). Anyway, collecting more data than needed is preferable to collecting less than the stated, given in oversampled data downscaling is possible, but under-sampled data cannot reconstruct original data correctly. However, big data collection and storage results into higher costs, so the collection strategy should be correctly designed to fit use-cases requirements to reduce costs and computational

time. The use of signal processing techniques is encouraged to design a suitable data collection strategy that addresses the use-case's PdM characteristics. Signal processing techniques can help to determine a suitable sampling frequency. Moreover, signal processing techniques include filters such as IIR Filters, Chebysev, Butterworth or Bessel as stated by Almaged and Hale (2019), which can be used to reduce the bandwidth of a signal that has a higher sampling frequency than required. When the sampling rate of the variables is different, in order to enable data analysis in any timestep for all available variables, timestep by imputation such as repeating last value or interpolation can be useful.

- *Data resources verification and pre-analysis*

  The second step is *data resources verification and pre-analysis*. It must be retrieved from the storing device, usually the online platform or hard disk on PLC. After that, a preliminary data analysis has to be performed using tools such as time-series or bi-variate plot visualisation, correlation analysis, feature description and distribution plots. It is extremely important to take into consideration that correlation does not mean causality and therefore, avoid believing that two or more variables are related whilst this relation is caused by external factors. For instance, these could be collected under same circumstances or were only modified at the same time by coincidence. This problem can be avoided with the integration of theoretical and domain expertise to validate aforementioned relations.

- *Data collection re-definition*

  The third step is *data collection re-definition*, where the integration on similar machines in aspects such as sensor type, models or placement is standardised. Moreover, the problem has to be re-evaluated considering the gathered data to check its suitability. This stage may reveal that collected data is not adequate for the defined use-case requirements and company characteristics, so iteration between current and previous steps is necessary until these are addressed; performing tasks like adding data sources or even revaluating the problem. Despite the initial machine monitoring process may not collect the most representative data for the designed task, this step helps to identify its gaps for further improvement. In addition, this methodology enables model contribution maximisation whilst minimising development effort; thus facilitating improvement by iterating steps.

- *Collected data validation*

The fourth step is *collected data validation* from data sources in general, and sensors in particular. The procedure consists of asserting it is in the expected range given sensor placement, sampling frequency, sensor type and related aspects. If any deviation is detected, iterations between current and previous steps are necessary to fix it. In addition, this step must be validated from domain technician perspective, thus iterating between current stage and following one to enable validation based on domain knowledge.

- *Domain knowledge analysis*

  The fifth and last step of resources analysis stage is *domain knowledge analysis* to gain insight about the use-case, its variable types, their physical meaning, how they are collected and the relation among them, both theoretically and in real process. This knowledge is essential for many steps such as data analysis, data preparation, model selection or ranking, and facilitates the creation of simpler, more accurate and easier to understand models. Nonetheless, even if domain knowledge is a key resource to understand data, there are usually differences between theoretical knowledge and collected data given the physical process is affected by many factors of collection procedure and component variability. For that reason, it important to analyse if data behaves as theoretically expected and if does not, be able to reason why. This can help discover problems in data acquisition process.

Data scientist and domain technician profiles have to work together in this stage in order to define use-case requirements and design the data collection process. Moreover, the data scientist will learn domain knowledge to understand better relations among variables and therefore facilitate data analysis and interpretation, combining domain knowledge with data-driven techniques. Domain technicians will also help to validate collected data and refine it until the obtained data is representative for the use-case.

By the end of this stage, two deliverable documents can be generated that can help in the analysis of resources suitability for the use-case and detect gaps, thus guiding the implementation of the methodology. Use-case requirements are collected in Document 2.1, whereas Document 2.2 contains information about resources to address them like: description of available data and how data collection process works; information of data pre-analysis with documentation of data and its characteristics, data visualisation plots, sensor placement or normal working range of variables; and data relation to physical meaning and domain knowledge.

These documents can facilitate the understanding of use-case requirements linked to its objectives, physical process, and data. Moreover, presentations composed by visual plots and concise text descriptions are interesting to convey the results of data analysis and resulting models to domain technicians and thus facilitate the communication to acquire knowledge. Therefore, domain knowledge can be used to complement data and resources analysis.

4.3. Model development

This is the stage where data-driven model of PdM system is designed, data is prepared according to its requirements, it is created and validated, obtaining as a result a version ready for deployment. It assumes that, after performing a preliminary data analysis and validation with domain expertise, the selected data subset contains predictor variables that are related to target variable. Moreover, relationship among variables is unknown given the complexity of physical systems. Therefore, the model is created under the basis that predictor variables have the power to predict target variables using an unknown function that the created model aims to represent in this stage. Several examples of target variables for PdM are: anomaly detection, diagnosis by RCA, health index calculation and prognosis, and RUL. Another assumption is that the observed data is big enough and of sufficient quality to represent those relations that can be generalised beyond the training process.

This stage has a more in-depth technical background than the rest to facilitate its implementation, given a high number of questions and difficulties arise when dealing with tasks related to data adaptation for the model, model selection and validation. To facilitate the understanding of this section, it is divided in the following subsections: *Model development flow* explains the flow of this stage's steps, while *Data preparation extension* and *Model selection extension* complement the information of corresponding steps with technical information. The reader not interested in technical details can skip the last two subsections and continue reading on *Model deployment and monitoring* stage.

4.3.1. Model development flow

- *Model design for use-case requirements*

  The first step of this stage is *model design for use-case requirements*. The addressed machine learning task must be chosen to accordingly, defining model requirements that could better fit that task with current data characteristics, such as classification, regression, clustering, one-class classification, etc. Then, state-of-the-art research on data-driven Predictive Maintenance models should be carried out to find commonly used data-driven architectures and models that could fit the problem and selected task, supported on articles like (Carvalho et al. 2019; Serradilla, Zugasti, and Zurutuza 2020).
- *Data preparation for model requirements*

  The second step is *data preparation for model requirements*, based on these processes: cleaning, preprocessing, feature engineering and split into train and test datasets. Commonly used data preprocessing techniques are: incorrect values cleaning, encoding and discretisation, segmentation, feature scaling (including normalisation and standardisation), noise reduction (reduce random variations of sensor output that are not related to sensor input) and imbalance data handling. Feature engineering can be done either extracting hand-crafted/traditional features that are relevant for the problem, or by using algorithms like Principal Component Analysis (PCA) or deep learning to extract features automatically from preprocessed features. The first type of features are easier to understand but require domain knowledge and are not specifically designed for the problem. In contrast, the second type of features are more difficult to understand but are trained directly from the data for the problem, so these do not require manual design of features. The extracted features should always be adapted to problem requirements and characteristics such as time-series, for instance extracting them in time-windows or cycles, to create features in new space where data context is easier to identify. When there is less information or data available, domain expertise and theory can help to gain additional insight or learn knowledge beyond the data. Data preparation is an essential step to achieve meaningful model results, and therefore, more in-depth information is presented in the subsection *Data preparation extension*.
- *Model selection*

The third step performs *model selection*, analysing state-of-the-art models and evaluating which could better fit the characteristics defined in step one of current stage. Moreover, the set of target variables must be chosen for the model and think of how these are related to the PdM stage it will perform, adapted to data characteristics like information level and additional resources like the available domain knowledge. More than one type of model can be combined to create a more robust model, thus complementing the gaps that only one model can have. Furthermore, the training strategy for the model must be defined to assert the model is robust to noise or changes in EOC by selecting the appropriate data train/test partition strategy or using cross-validation. In order to facilitate the selection of a model that addresses the desired PdM stages, the subsection *Model selection extension* explains how to create them, and which type of models are most suitable for use-case data characteristics and requirements.

- *Design validation strategy*

  The fourth step is the *design validation strategy* that will be used to compare and rank models in training and testing phases. It consists of choosing the most suitable metrics according to use-case characteristics, considering which are the target variables and how the model is designed to fit them. In addition to the quantitative approach the validation metrics offer, additional qualitative comparison strategies can be defined with domain technicians to integrate domain validation; these strategies assert that models also address use-case peculiarities from technical perspective.

- *Model creation and optimisation*

  Hereafter, the *model creation and optimisation* take place in step five, based on the defined use-case requirements, prepared data, selected model and designed validation strategy in current stage. The model is trained to map predictor variables to target variables with the objective to minimise the error of its estimations, thus learning to relate them based on data. However, it must be constrained to generalise from data beyond the training set; this way it will work with novel data belonging to the same distribution.

- *Model validation and ranking*

  The sixth and last step of this stage is called *model validation and ranking*, where validation metrics together with domain knowledge are used to evaluate, compare and finally rank the generated models. This step enables to prioritise and validate models in a systematised way, asserting that the chosen ones are the most suitable for the data and are aligned with use-case requirements; commonly, iterating between the two last steps of this stage is necessary to achieve this suitability. The Monte Carlo simulations technique can be used validate data-driven models for PdM, as implemented in Ley and Orchard (2021). Once the model meets the desired characteristics, it is ready for deployment.

This *model development* stage is guided by data scientists, who use different data-driven techniques to clean and prepare the data to create the chosen data-driven model. Nonetheless, constant interactions with domain technicians are necessary to assert the developed data-driven model addresses use-case requirements and ensure its estimations are related to physical meaning. This facilitates diagnosis, increases trust of stakeholders in the model, and ensures the model is created in a robust way; avoiding data biased relations by validation with domain knowledge.

In this stage two deliverables are generated: Document 3.1 contains the decisions and steps performed for data preparation and model development, gathering the following aspects: definition of suitable model characteristics to address use-case requirements; data preparation steps for the model; research on state-of-the-art models and reasons for prioritising some models over others; definition of model validation strategy; and definition of how domain knowledge is integrated into models, specifying how it guides their development and validation. In addition, the deliverable 3.2 is the data-driven Model trained and validated in training data, guided by domain technicians to integrate domain knowledge.

4.3.2. Data preparation extension

Model should be chosen to address use-case and data requirements. Accordingly, its performance, interpretability, processing time and many other characteristics vary among use-cases and are tied to their limitations and decisions undertaken during the preparation. For instance, linear models are usually faster and easier to interpret, but they have limitations when modelling non-linear data relations.

The most challenging task of PdM system development is to obtain a dataset that is representative for the problem, preprocessed, and containing only features that are relevant yet interpretable if possible. It is better to focus efforts on collecting better data when the collected one is not good enough or little for the designed task rather than optimising a specific model to achieve a slightly higher accuracy. The reason behind this statement is that even the most complex models capable of modelling any kind of relationship are useless in a dataset that lacks of information on target variables, or when these are not useful for the previously defined business problem.

Commonly, process data contains time-series signals, which can show characteristics like seasonality, stationarity and trend. Thus, observations of one variable are related to observations of the same in different timesteps and cycles, which can be useful to detect trends. This data is typically analysed together by taking chunks of specific size of continuous observations, technique denominated as sliding window. After analysing a specific time or cycles frame, the window advances to next chunk. Data can also be divided and loaded into smaller datasets when it does not fit into computer or server RAM, enabling to load and free memory at will. Some libraries that are specifically designed for that task facilitate this implementation. In addition, many factors influence performance of developed algorithms; the main ones are discussed below.

When the modelled data belongs to different working conditions, these can be grouped by similarity to be analysed together or even create one model for each working mode. The latter can improve accuracy while simplifying the model, but does not generate common relations among data of different EOC. This issue can be solved by using only one model instead of many and feeding it with working condition data.

With the objective to create a model that adapts to use-case requirements, data peculiarities, and in further stages be able to interpret its predictions or trust it, it is necessary for data scientists to gain domain knowledge. However, many times they will need the assistance of expert technicians for model interpretation or optimisation on any step performed by following the Predictive Maintenance stages presented in the methodology.

4.3.3. Model selection extension

Selecting the most adequate machine learning task to solve for each use-case is not trivial. This subsection aims to facilitate the analysis and choice of ML architectures given data characteristics, and recommended ML type to solve the corresponding task presented in Figure 3; these are described in incremental levels according to the information companies have with regard to data, in the paragraph below. The roadmap presented in this Figure can be used to select the correct model type to address target steps of PdM roadmap (Figure 1). Moreover, this section describes the characteristics of ML tasks for each PdM stage: anomaly detection, diagnosis, prognosis, and mitigation. Data acquisition and preprocessing are two additional stages that prepare data for PdM which are often overlooked. Despite this fact and being resource demanding, these are necessary to obtain high-quality data and, as a result, accurate models.

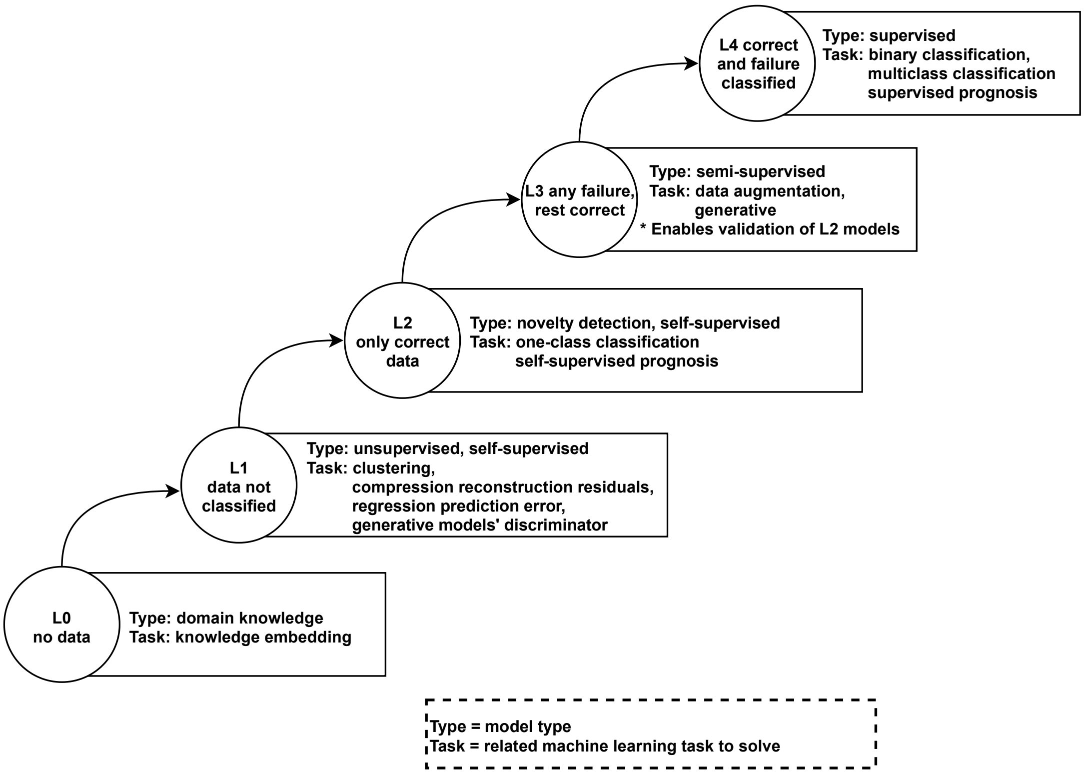

Figure 3 Roadmap to assess in data-driven task and machine learning model selection according to available data levels. Higher levels indicate higher information on collected data, which enable more accurate results and possibilities to address PdM roadmap of Figure 1.

Information about the capabilities of each data level in Figure 3, and which PdM steps can address is explained in the following list:

- L0) the company lacks working historical data. Therefore, the only possible approach is using domain knowledge-based systems that embed prior theoretical and expertise into the system.
- L1) unclassified data can only be treated in unsupervised way by clustering data-driven systems, and self-learning by tracking reconstruction residuals using compression techniques like autoencoder, calculating prediction error of regressors or by using discriminators of generative models.
- L2) domain technicians confirm that collected data belongs to normal working condition, so novelty detection techniques, commonly one-class classification algorithms, are used to detect instances different or far from the already known. This level also enables self-supervised prognosis, which can be used to identify changing trends on machine condition by integrating domain expertise.
- L3) there are few failure observations classified and the rest are classified as correct data, so semi-supervised algorithms are used like data-augmentation for imbalance class data handling, and synthetic failure data generators. Monitoring few failures of one failure type may not be enough to train a supervised model, and therefore more observations may be required. However, any failure data can be used to validate and calibrate models trained on level two, obtaining more accurate and robust models.
- L4) there is enough failure and non-failure data, therefore the problem is supervised. In this scenario, binary classification algorithms are used to classify failure and correct observations, and when there are different types of failures classified in training data, multi-class classification algorithms are used to differentiate them. This supervised historical data has applications on supervised machine condition prognosis to estimate how it will evolve, and therefore detect trends like degradation.

Each data level can use tasks of lower levels in addition to the ones of itself. The higher level in the roadmap, the more complete and accurate predictions the model can achieve. Machine learning tasks of data levels are used to solve Predictive Maintenance steps:

- *Anomaly detection* is commonly performed by classification models that classify asset condition into faulty or correct, and some can even classify different types of failures. However, these models can only be used when there are enough observations of the target failure types. In many cases, there is little or no failure data so the common strategy in these scenarios is modelling asset normal behaviour and detect anomalies when the data is different. There are three types of anomalies regarding the number of observations involved: individual observation denominated as local, global which formed by several observations, and context is not an anomaly by itself, but it becomes one when additional information is given. The type of anomaly to be detected should be chosen according to the use-case.

- *Diagnosis*: once an anomaly has been detected, diagnosis should be performed to analyse which components have been affected, in which way and to what extent. Some possible common factors are measurement errors, changes in EOC, component degradation while keeping correct working mode, failures and conditions that can lead to them. A useful technique to detect failure causes is root cause analysis, which is defined by Andersen and Fagerhaug (2006) as "structured investigation that aims to identify the true cause of a problem and the actions necessary to eliminate it". It also and defines three levels of causes: symptoms, the first-level causes that lead to the problem, and the higher-level causes that lead to first-level causes, where root cause is the origin.
- The used diagnosis techniques have to adapt to use-case, its data requirements and the implemented anomaly detection model. When correct and failure data or even different type of failures is available, performing a classification of failure types is straightforward. Conversely, when there is no failure data classified, this step has to be performed in unsupervised or self-supervised ways, by combining these types of machine learning models with domain expertise. These techniques can make use of health indexes, which represent the percentage of deviation the assets suffer with respect to past correct working data that could be related with damage.
- *Prognosis* in PdM is based on remaining useful life models which estimate the remaining time to failure of a component or asset in the working conditions of that moment based on its state and the detected and diagnosed anomalies. These models can also provide a confidence bound of their condition. Conversely, when there is no historical run-to-failure data, data-driven prognosis models can

only perform health degradation monitoring and estimation by tracking health index in a self-supervised way.

- *Mitigation*: the last step consists of providing operators with data-driven notifications and recommendations to speed up and optimise maintenance. These should be based on alarms and information gathered from previous PdM stages in a simple yet effective way to understand by domain technicians.

Table 2 presents how the integration of domain knowledge with data-driven systems can help in each step of Predictive Maintenance methodology application, supported on the scheme presented in the review by Khan and Yairi (2018).

All in all, no algorithm is better than the rest, their suitability depend on use-case and data requirements. Moreover, more than one model can be adequate for the same use-case. It is useful to analyse their specifications, and choose based on guidelines, related reviews and state-of-the-art works like (Selcuk 2017; Serradilla, Zugasti, and Zurutuza 2020). Nonetheless, nowadays it is common to combine different techniques

Table 2 Contributions of data-driven and domain knowledge in each PdM stage individually and combined.

Stage	**Domain knowledge (domain technician)**	**Explainable data-driven techniques**	**Combined**
Data adquisition and database creation	Define relevant features to monitor according to experience and theoretical background. Direct relation to physical meaning. Assert collection is correct given knowledge.	Feature relevance with respect to target feature according to data.	Most relevant features selection. Knowledge gain on process and understand it better combining data and domain knowledge.
Data preparation	Extract and select the most relevant features to monitor according to domain knowledge and relation to physical meaning. Validate and clean data.	Preprocessing: automatic machine learning based techniques to preprocess data: encoding, data cleaning, scalling, noise reduction, imbalance data handling and not available data handling. Automatic feature extraction, selection and fusion.	Select and extract the most representative features for the use-case and target variable that could be understood by domain technicians or linked to physical proccess. It may be automated by data-driven techniques, previously guided and afterwards validated by domain technicians.
Anomaly detection	Help create the data-driven system. Then, assert that anomaly detector works well in preproduction, production and detect changes in trends, helping to decide when to retrain. They also enrich models' output with expertise.	Automatically detect anomalies based on process and related data, comparing it with historical database and embedded knowledge.	Automatic data-driven anomaly detection and verification based on experience and theoretical background.
Diagnosis	Validate the diagnosis performed by data-driven system and complement it with expertise, theoretical background and add relevant external information, i.e. collecting additional EOC information not gathered in the data.	Diagnose anomaly type if anomaly detector is a multiclass classifier. Conversely, when the model is binary classifier or unsupervised or self-learning, it can outline the reasons or variable values that made it be anomaly, helping technicians perform diagnosis.	ML extracts data and models correlations to help in diagnosis, and domain technicians use context and knowledge to validate and complement algorithm predictions and gain additional knowledge.
Prognosis	Prognose degradation based on assets properties like materials, designed lifecycle and working experience.	Prognose asset degradation by tracking their health based on data, monitoring how it changes with time.	Combine data and knowledge to perform more accurate prognosis and gain knowledge.
Decision making and mitigation	Plan and coordinate maintenance actions supported on maintenance management and manufacturing operation management processes, using PdM system information to address process requirements. This enables moving towards a more optimised maintenance.	Raise alarms, notify strange working conditions and give recommendations to prevent failures. In addition, advice how to perform more optimal maintenance by comparing current condition with historical data.	Domain technicians investigate data-driven alerts, recommendations and highlighted data by comparison with previous events and knowledge to interpret, understand and validate their predictions. Based on these resources, technicians create and execute maintenance plan.

to create a more complete architecture that overcomes the gaps of containing algorithms, thus better addressing use-case challenges.

4.4. Model deployment and monitoring

The final stage of the methodology consists of preparing the model and taking the necessary steps for its deployment to production.

- *Deployment strategy definition*
  The first step is *deployment strategy definition* to systematise and speed up the action of putting models into production. Firstly, the most suitable location for model must be selected, choosing whether it will be executed in cloud or a Programmable Logic Unit (PLC) in the production plant or edge. Another relevant aspect is the execution periodicity, which can be offline, on streaming, or periodical after performing a certain number of cycles or working time. This strategy should also contain detailed steps to follow for model deployment to production, which will provide with a framework that facilitates model testing in production environment while avoiding mistakes.
- *Model deployment to production*
  The second step is *model deployment to production*. For that, first the developed model must be tested in a preproduction environment that shares the characteristics of the production environment. This enables to detect incompatibilities and possible problems of the model in production without interrupting or damaging that system, facilitating the deployment of the pretested model. Once the model is running in production, it has to be tested by monitoring its performance with new production data and maybe generating synthetic failures or degradation data to check that it works correctly and raises alert messages.
- *Analyse model alerts and mitigate problems*
  The third step of this stage is to *analyse model alerts and mitigate problems*. It aims to assert correct model performance in production, so the model has to work correctly with novel production data and should also be tested with different already tested data of the training phase to assert the alarms are raised correctly. In case any problem or abnormality arises in the process, it should be addressed until the model works properly according to the defined requirements.

- *Monitor model performance and correct working*

  The fourth step is to *monitor model performance and correct working* by tracking its evolution and adaptability to production data with techniques like prediction uncertainty, detection of changes in data such as EOC, and machine degradation that should be reflected on data. When any of these indicators suggests that the model is not working correctly with collected data, a more robust analysis should be performed before accepting this conclusion. This analysis must combine data-driven techniques and domain expertise to perform a complete evaluation from both perspectives, gaining more insight and facilitating the cause detection of incorrect working.
- *Model retraining*

  If the analysis concludes that the model is not working correctly, then the fifth step of this stage and last step of the methodology must be performed: *model retraining*. This step is supported on the conclusions of the abnormal working analysis performed in the previous step, which will be used to define the tasks and resources required to retrain the model. After the resources are collected and tasks are performed to prepare the retraining, the process switches to the model creation and optimisation step of the model preparation stage, where the model will be adapted to new requirements and its development and deployment will continue from this step.

In this last stage, domain technician and data scientist profiles have to work together to: define the most suitable deployment strategy for the model in the use-case according to its requirements and resources; deploy the model and validate it combining data-driven metrics and domain knowledge; monitor its performance and go back to retraining when it stops working correctly to adapt to new working conditions.

During this stage, two deliverables are generated: deliverable 4.1 contains the Model working in production that analyses production process data either on streaming or periodically. This model has been tested to assert it works correctly in production, and protocols to handle its alerts and retraining are collected in the deliverable 4.2. The last deliverable is 4.2 Document, which defines the deployment steps necessary to take the model generated in previous stage to production, containing: a protocol to test whether the model is working correctly in production; a protocol to analyse and mitigate PdM alerts the model generates, thus facilitating domain technicians the application of

PdM in the system; definition of how to monitor model performance in production using data-driven signals and domain knowledge to analyse and define when retraining is required; and guidelines to define how the retraining process should take place.

As explained throughout this section, the methodology is iterative, which means that backward steps are recommended to create a model that fits better use-case requirements. According to the previous argument, even when the model development has finished and it is into production by the implementation of the last step, it can be further improved and should be monitored to adapt to industrial evolving requirements. However, even if there is no need for the model to be continuously under development, it could have different versions as time goes by to adapt to new circumstances and integrate novel knowledge.

5. Methodology application in fatigue tests remaining useful life

This section explains how the methodology presented in current work has been applied to successfully implement a data-driven model that estimates the remaining useful life on bushing testbed experiments, as explained in the article by Serradilla, Zugasti, Cernuda, et al. (2020). That work enabled the validation of the methodology, demonstrating its application feasibility in an industrial environment where business, domain technician and data scientist roles worked together to obtain a model that met use-case requirements.

To sum up, the objective of that article is to apply interpretable yet accurate data-driven models to predict the remaining life of fatigue experiments, which belongs to the third stage of PdM roadmap of Figure 1. In these experiments, bushings are subjected to proportional EOC to real machines, which enables to extrapolate the knowledge learnt in controlled environment to real machines in the future. The dataset of this use-case contains 576 experiments where 97 EOC variables are monitored in a sampling rate of 1 sample per second. Given interpretability is targeted, the obtention of easy to understand features and models is essential, which means that domain knowledge plays a key role in the design, development and validation of models.

The complete methodology MEDADEK presented in APPENDIX A has been adapted to address current use-case requirements, simplifying and adapting its steps to facilitate its understanding and implementation. These steps are grouped by stages and summarised in diagram of Figure 4, where asterisks indicate required steps of the methodology, and the step-by-step process is explained in current section.

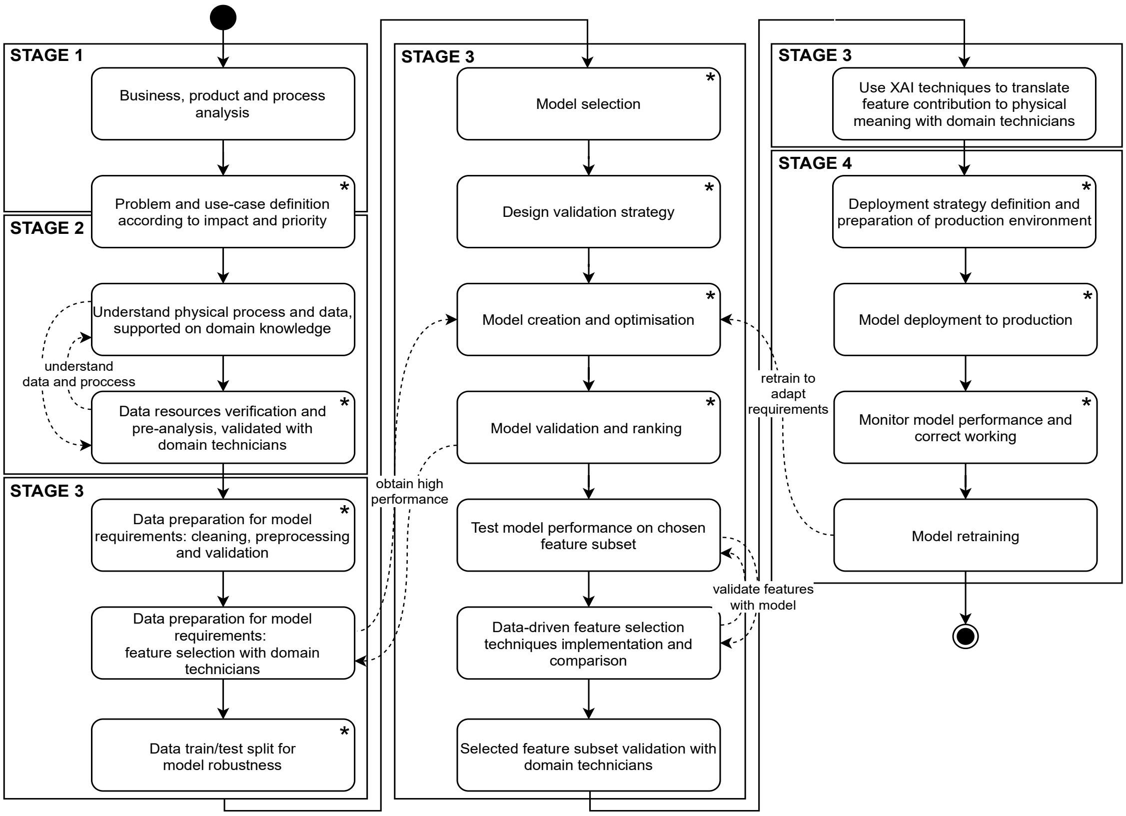

Figure 4 Diagram of MEDADEK adapted for explainable bushings' remaining life estimation use-case.

The first step was *business, product and process analysis*, which forms the first stage together with the second step. It focused on learning about the company of this use-case, which designs, manufactures and supplies metal forming machine tools that carry out different forming operations to manufacture parts. The second step was *problem and use-case definition according to impact and priority*, thus analysing and prioritising the maintenance issues that can have more impact for the business. Most manufacturing companies consider critical the unexpected failures given their downtime and repair costs. Therefore, the decision to study bushings was taken given they play a key role in machines. With the objective to perform experiments and gather data, a bushing testbed was used.

The second stage started in this step continued on the next step of *understand physical process and data, supported on domain knowledge*, first in real machines and then how the testbed worked. This facilitated the iteration with the following step denominated as *data resources verification and pre-analysis, validated with domain* *technicians* and its labour, which used plots, correlations and distributions to further increase the process understanding. As a result, a summary report was created and presented to domain technicians to contrast gained knowledge and information and answer questions arisen in the process.

The third stage started on the fifth step of the diagram called *data preparation for model requirements: cleaning, preprocessing and validation* of the collected data. Accordingly, it is directly related with the next step *data preparation for model requirements: feature selection with domain technicians*; this separation helps to highlight the high relevance feature engineering has in this use-case. This step aims to obtain a subset of representative features for the problem as the original features where highly correlated among them. Henceforth, domain knowledge and correlation analysis where combined to select a subset of features. Then, the following step *data train/test split for model robustness* was selected to be representative for the problem and consequently chosen to be analysed afterwards with the chosen model: 10-fold cross validation and 80/20 train/test split. The k-fold cross validation technique is a robust way to evaluate the performance of a model with new data. The training dataset is divided into k pairs of training and validation data subsets, to train and evaluate one model for each fold.

After preparing the data for the model, the *model selection* step was performed. In this step, the PdM task was delimited to a regression problem where the target was to estimate the remaining time for experiments at any time according to process data. Different state-of-the-art data-driven models were tested in the problem, concretely gaussian naive bayes, linear support vector regressor, k-neighbors regressor, linear discriminant analysis, xgboost regressor and random forest regressor. Deep learning models were discarded given these are harder to interpret and therefore collide with the main objective of the use-case.

At this point, a validation strategy to compare the models was required, so the life-cycle advanced to *design validation strategy* step, where the model ranking strategy was defined using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) metrics. Moreover, the model validation based on knowledge was defined by interacting with domain technicians in the following way: presentations on model results with aforementioned metrics and explainability methods to infer models' overall behaviour and check it fits domain knowledge. Then, the *model creation and optimisation* step for the selected state-of-the-art models was performed, and the *model validation and ranking* step was performed according to the previously defined model validation strategy. The testing procedure used 10-fold cross-validation for robustness, and the metrics, showed that tree ensemble models were much more accurate than the rest and their similar performance made us focus on random forest given it provides a naive explainability method of feature importance. The result of 10-fold cross-validation was compared with random train-test splits and was similar, so the later was chosen to speed up to testing models in training, and then use again cross-validation to test the final model; this is the procedure for the step *test model performance on chosen feature subset*.

Afterwards, the selected random forest regressor was used as baseline to improve the previously selected feature subset by applying the step *data-driven feature selection techniques implementation and comparison,* using the already experimented 80/20 train test split by iterating with the previous step. Its result showed that the best error results was feature importance, concluding that this explainability method was useful for this use-case. This feature selection analysis was also useful for the next step: *selected feature subset validation with domain technicians*. Here the feature importance results and model error deviations when changing selected features were analysed. The selection of correlated features was done according to sensor placement, selecting the sensors that were physically more important for the prediction. Therefore, the feature selection took place combining data-driven techniques with domain knowledge. After choosing the final feature subset, all the models were retrained and their performance was analysed, which did not significantly change while allowing model simplification by input dimension reduction. The last step of the third stage was to *use XAI techniques to translate feature contribution to physical meaning with domain technicians*, which uses model explainability to infer knowledge on feature prediction capability on the model in both ways: global indicating overall performance and local indicating feature importance evolution throughout the experiment, both enabling comparison among models.

The fourth and final stage is dedicated to connecting the developed model with the testbed data collection module so that the model can analyse testbed data on streaming. The first step of this stage was *deployment strategy definition and preparation of production environment*, where the deployment platform that replicated preproduction environment characteristics was designed and prepared, and it was connected to the testbed to enable the integration. After creating the production environment, the second stage was *model deployment to production*, where connection between testbed data and model is tested, and the model reads testbed data, use it to make predictions of the remaining experiment life, and returns the prediction to the testbed. Then, the prediction can be shown to the machine operator or be used to generate alerts that can predict current and future working failures of the system, and as a result, prevent failures. The penultimate step and currently the step being performed in this use-case was *monitor model performance and correct working*, which is a relevant task to keep the model up-to-date and adapted to changes in EOC, machine degradation and related factors. The final step of the methodology was *model retraining*, which defines and collects the resources, and performs the tasks necessary to adapt the model for the new environment and requirements, and then going back to the step model creation and optimisation to complete the retraining when required.

6. Annotations on methodology application

6.1. Applicability of the methodology

The methodology presented in current paper is divided into stages and steps to provide manufacturing companies with an easy-to-follow guideline to develop data-driven models guided by domain technicians for Predictive Maintenance. It is adaptable to company requirements, agile and iterative to minimise the time to production, delivering value since early stages. It is based on proof of concepts, going through the whole model creation process and ending in production for simplified models and once having a model deployed, iterating to improve it.

Its application has shown that data-driven models influenced by expert knowledge enable the application of more accurate yet complex data-driven models in industry by the physical meaning integration and understanding of domain knowledge. Moreover, even though business vision is an aspect commonly overlooked in state-of-the-art works, it is essential for project success and its alignment with company requirements and business characteristics. Therefore, data scientists need to know about companies and collaborate with domain technicians and business profiles, learning to communicate their problems in a simpler way and adapting to their language.

Manufacturing companies that are starting to implement Predictive Maintenance models may not have the most optimal resources or expertise to develop them. For that reason, is important to adapt to company requirements and try to generate value with available resources. Regardless, data acquisition process should be continuously improved given that model performance relies on data quality. Data can be improved in many ways such as adding metadata that enriches it with context, augmenting the sampling rate to a sufficient to detect the aimed patterns, adding new sensors or annotating data.

6.2. Remarks to address challenges

Most data scientists would benefit from learning technical knowledge about the use-cases, like how processes and machines work, and basic mechanical background; these skills will facilitate their communication with other profiles and open their vision to approaches that complement their mathematical and programming skills. The data-driven PdM techniques must be suitable for use-case requirements, which makes necessary model state-of-the-art research, background acquisition, implementation and adaptation of domain knowledge.

Data quality is an aspect often overlooked and taken for granted, despite being one of the most important aspects in data-driven model lifetime, having high influence on model performance, and being the base where model development steps are supported. Therefore, data preprocessing, cleaning, feature engineering, and data validation tasks are relevant to obtain a dataset of predictor variables related to the target; these will remove missing data or redundant variables that increase model complexity and training resources, while preventing the model from learning biased relations from the dataset that are not physically or theoretically related. Even the most advanced data-driven systems are unable to achieve good enough performance in bad datasets, but more basic algorithms can model well designed and meaningful datasets, at least to some extent. Hence, the data collection process optimisation must be a priority. Furthermore, data scientists spend much of their working hours in data cleaning, so the optimisation and systematisation of this process can speed up the implementation of the whole methodology.

Collaboration of at least three profiles is essential for successful PdM models development: business profile, domain technician and data scientist. Additionally, given the data acquisition process is key for these projects, domain technicians must research and analyse the adequate sensors for each use-case and computer science roles to automate data collection and storage.

Each use-case has to define its requirements and the developed system must address them. Commonly, the implemented systems have to process the data in streaming to detected the anomalies that are happening in the moment, and raise alarms for operators to stop the machine and check its health, thus preventing failures that can damage components, Stated in the article by Rieger et al. (2019). Nonetheless, if the aim is to detect periodical degradation, the accuracy of analysis is more important than prediction speed, where periodical execution of more precise algorithms is preferable over faster yet less accurate ones. All in all, use-case definition is essential to design data-driven characteristics.

A difficulty encountered when modelling machine data is that the components deteriorate with time but also maintenance is applied to them to restore their components' health. The work by Bergquist and Söderholm (2016) explains that models must be updated with new information about the current process dynamics. It is important to define a strategy to retrain models so they can adapt to monitored component degradation, but at the same time they are capable of detecting when assets are not working properly or are degrading.

The systematisation of design, creation and deployment of models by the presented methodology facilitates estimation of project cost, necessary resources and required time. These projects are focused on business and process requirements and therefore they can add value to manufactured products, thus improving manufacturing company competitiveness. Despite client privacy could be affected when collecting data from them, if this process is handled securely, data can bring competitive advantages to industrial companies.

It is important to add business perspective to the problem and think about the tasks machine learning can address in industry and concretely their applications on

Predictive Maintenance. Although business perspective is out of the scope of this article, each company has to embrace its own strategy and think of ways to take advantage of this technology. It has applications on managing post-sales services, monitoring the aging of their machines in order to find the most fragile/critical components, finding failure causes that will enable components improvement by proactive maintenance or changing and evolving the business model to servitisation, and many other benefits.

7. Conclusions

This article presents the main challenges manufacturing companies face when implementing Predictive Maintenance, and highlights related works, standards and related methodologies as reference framework. Then, presents a four stages methodology that is afterwards validated on a real machine use-case, and finally the results and lessons learned from this application are exposed.

The main contribution of this article is the methodology for data-driven Predictive Maintenance models generation guided by domain technicians for industrial companies: MEDADEK-PdM. It is designed to be open and modular to facilitate its adaptation to diverse industrial use-cases in an iterative way. The systematisation of this process was necessary due to each company was implementing models according to their expertise and had to make significant efforts to understand and structure the whole process. In addition, many state-of-the-art works focus on model development under controlled environment. For instance they use reference simulated data like turbofan by Saxena and Goebel (2008) that do not reflect industrial companies requirements such as changing EOC, interpretability or real-time/online data processing. This methodology also covers the model deployment stage, which many state-of-the-art works lack, with the objective to facilitate companies deploying their models to production.

Business, domain and data perspectives are necessary to understand which problems are more critical to solve, and understand the physical process and data. All this context is necessary to create models that better align and fulfil company requirements. Data-science should not be an isolated process of the company limited to data analysis and computation, it should be integrated and contribute to the whole data acquisition process, using expertise and use-case data to ensure high-quality data collection, thus facilitating model creation.

The more data and the higher its quality, the more accurate models can be created and more information will be able to infer from them. One relevant task of the process is to obtain a representative dataset that contains both, defined target variables and predictor variables that are somehow correlated to them, giving context or related signals which have the power to estimate target variables. Once obtained this dataset, model selection, optimisation, ranking and validation will take place, but the quality of these relies on the quality of previous dataset. The data has to be representative for the use-case and accordingly the model has to be chosen to meet the requirements of both.

Moreover, physical, theoretical and domain knowledge are necessary to be able to interpret the data and understand the generated data-driven models. Without them, even the simplest models are difficult to understand and test whether they are working correctly or not, thus being handled as black-box. Therefore, the machines and component target operating working modes have to be understood. For this task, domain technicians are essential, who explain the relationships among collected variables either theoretically or by their expertise and enrich the process understanding, data analysis and model creation and validation.

The completion of this research work has opened up two future research lines. The first research line is the validation of the methodology in production line use-cases, using it to develop data-driven PdM models guided by domain knowledge. The first steps of this research line include the implementation of the methodology to develop a PdM system for a stamping production line, which has been disseminated in the work Serradilla et al. (2021). The second open research line is the application of the methodology to develop PdM systems in other use-cases beyond manufacturing, analysing its suitability and adaptability to other research fields.

Funding

Oscar Serradilla, Ekhi Zugasti and Urko Zurutuza are part of the Intelligent Systems for Industrial Systems research group of Mondragon Unibertsitatea (IT1357-19), supported by the Department of Education, Universities and Research of the Basque Country. This work was partially supported by the European Union's Horizon 2020 research and innovation programme's project QU4LITY under Grant agreement n.825030, and Provincial Council of Gipuzkoa's project MEANER under Grant agreement FA/OF326/2020(ES).

Declaration of interest

This manuscript is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere. All authors have checked the manuscript and have agreed to the submission, reporting no potential competing interests.

Availability of data and material

For the research of this work, only publicly available studies, works and references combined with authors' experience has been used.

References

Able, C. M., Baydush, A. H., Nguyen, C., Gersh, J., Ndlovu, A., Rebo, I., et al. (2016). A model for preemptive maintenance of medical linear accelerators-predictive maintenance. *Radiation Oncology*, *11*(1), 36. https://doi.org/10.1186/s13014-016-0602-1

Aguiar Guzman, Leonardo Rodriguez Borja, H. (2014). Análisis De Modos Y Efectos De Falla Para Mejorar La Disponibilidad Operacional En La Línea De Producción De Gaseosas, 81. https://repository.unilibre.edu.co/bitstream/handle/10901/7838/Doc Final Proyecto A-rmando y Leonardo sustentacion.pdf?sequence=1

Aivaliotis, P., Georgoulias, K., & Chryssolouris, G. (2019). The use of Digital Twin for predictive maintenance in manufacturing. *International Journal of Computer Integrated Manufacturing*, *32*(11), 1067–1080. https://doi.org/10.1080/0951192X.2019.1686173

Albano, M., Jantunen, E., Papa, G., & Zurutuza, U. (2018). *The MANTIS book: Cyber physical system based proactive collaborative maintenance*. *The MANTIS Book: Cyber Physical System Based Proactive Collaborative Maintenance*. River Publishers. https://doi.org/10.13052/rp-9788793609846

Alexopoulos, K., Nikolakis, N., & Chryssolouris, G. (2020). Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing. *https://doi.org/10.1080/0951192X.2020.1747642*, *33*(5), 429–439. https://doi.org/10.1080/0951192X.2020.1747642

Almaged, M., & Hale, J. (2019). Virtual Instruments Based Approach to Vibration Monitoring, Processing and Analysis. *International Journal of Instrumentation and Measurement*, *4*.

Alvares, A. J., & Gudwin, R. (2019). Integrated System of Predictive Maintenance and Operation of Eletronorte Based on Expert System. *IEEE Latin America Transactions*, *17*(01), 155–166. https://doi.org/10.1109/TLA.2019.8826707

Andersen, B., & Fagerhaug, T. (2006). *Root cause analysis: simplified tools and techniques*. Quality Press.

Ansari, F., Glawar, R., & Sihn, W. (2020). Prescriptive Maintenance of CPPS by Integrating Multimodal Data with Dynamic Bayesian Networks. In J. Beyerer, A. Maier, & O. Niggemann (Eds.), *Machine Learning for Cyber Physical Systems* (pp. 1–8). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-59084-3_1

Baptista, M., Sankararaman, S., de Medeiros, I. P., Nascimento, C., Prendinger, H., & Henriques, E. M. P. (2018). Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Computers and Industrial Engineering*, *115*, 41–53. https://doi.org/10.1016/j.cie.2017.10.033

Bergquist, B., & Söderholm, P. (2016). Predictive Modelling for Estimation of Railway Track Degradation. In U. Kumar, A. Ahmadi, A. K. Verma, & P. Varde (Eds.), *Current Trends in Reliability, Availability, Maintainability and Safety* (pp. 331–347). Cham: Springer International Publishing.

Borodulin, K., Sokolinsky, L., Radchenko, G., Tchernykh, A., Shestakov, A., & Prodan, R. (2017). Towards digital twins cloud platform: Microservices and computational workflows to rule a smart factory. In *UCC 2017 - Proceedings of the10th International Conference on Utility and Cloud Computing* (pp. 205–206). https://doi.org/10.1145/3147213.3149234

Bousdekis, A., Magoutas, B., Apostolou, D., & Mentzas, G. (2015). A proactive decision making framework for condition-based maintenance. *Industrial*

Management & Data Systems, *115*(7), 1225–1250.

Brownlee, J. (2018). *Introduction to time series forecasting with Python: how to prepare data and develop models to predict the future*. Machine Learning Mastery.

Cao, Q., Zanni-Merk, C., Samet, A., de Beuvron, F. de B., & Reich, C. (2020). Using Rule Quality Measures for Rule Base Refinement in Knowledge-Based Predictive Maintenance Systems. *Cybernetics and Systems*, *51*(2), 161–176. https://doi.org/10.1080/01969722.2019.1705550

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers and Industrial Engineering*, *137*, 106024. https://doi.org/10.1016/j.cie.2019.106024

Cerqueira, V., Pinto, F., Sá, C., & Soares, C. (2016). Combining boosted trees with metafeature engineering for predictive maintenance. In H. Boström, A. Knobbe, C. Soares, & P. Papapetrou (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9897 LNCS, pp. 393–397). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46349-0_35

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Cross Industry Standard Process for Data Mining. *CRISP-DM Consortium*.

Coleman, C., Damodaran, S., Chandramouli, M., & Deuel, E. (2017). Industry 4.0 and predictive technologies for asset maintenance. *Deloitte Insights*. https://www2.deloitte.com/us/en/insights/focus/industry-4-0/using-predictive-technologies-for-asset-maintenance.html. Accessed 16 September 2019

Dhillon, B. S. (2002). *Engineering maintenance: a modern approach*. cRc press.

Diez-Olivan, A., Pagan, J. A., Sanz, R., & Sierra, B. (2017). Data-driven prognostics using a combination of constrained K-means clustering, fuzzy modeling and LOF-based score. *Neurocomputing*, *241*, 97–107. https://doi.org/10.1016/j.neucom.2017.02.024

Dos Santos, T., Ferreira, F. J. T. E., Pires, J. M., & Damasio, C. (2017). Stator winding short-circuit fault diagnosis in induction motors using random forest. In *2017 IEEE International Electric Machines and Drives Conference, IEMDC 2017* (pp. 1–8). https://doi.org/10.1109/IEMDC.2017.8002350

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, *210*, 15–26. https://doi.org/10.1016/J.IJPE.2019.01.004

Institute', 'Deloitte Analytics. (2017). Predictive Maintenance: Taking pro-active measures based on advanced data analytics to predict and avoid machine failure. *Deloitte Consulting GmbH*.

ISO 17359. (2018). *Condition monitoring and diagnostics of machines — General guidelines* (Vol. 2018). pub-ISO:adr.

Jordan, W. E. (1972). Failure modes, effects and criticality analyses.

Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, *107*, 241–265. https://doi.org/https://doi.org/10.1016/j.ymssp.2017.11.024

Lavi, Y. (2018). The Rewards and Challenges of Predictive Maintenance. *InfoQ*. https://www.infoq.com/articles/predictive-maintenance-industrial-iot/

Lebold, M., Reichard, K., Byington, C. S., & Orsagh, R. (2002). OSA-CBM architecture development with emphasis on XML implementations. In *Maintenance and Reliability Conference (MARCON)* (pp. 6–8). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.4066&rep=rep1&type=pdf

Ley, C. P., & Orchard, M. E. (2021). Simultaneous inference of Lithium-Ion battery polarising impedance surface and capacity degradation using a Hybrid Neural Adaptive State Space Model. *Journal of Energy Storage*, *36*, 102370. https://doi.org/10.1016/J.EST.2021.102370

Li, D., & Gao, J. (2010). Study and application of Reliability-centered Maintenance considering Radical Maintenance. *Journal of Loss Prevention in the Process Industries*, *23*(5), 622–629. https://doi.org/10.1016/j.jlp.2010.06.008

Li, Z., & Wang, Y. (2018). Domain Knowledge in Predictive Maintenance for Water

Pipe Failures. In J. Zhou & F. Chen (Eds.), *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent* (pp. 437–457). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-90403-0_21

Liao, L., & Köttig, F. (2016). A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction. *Applied Soft Computing Journal*, *44*, 191–199. https://doi.org/10.1016/j.asoc.2016.03.013

Lukac, D. (2016). The fourth ICT-based industrial revolution “industry 4.0” - HMI and the case of CAE/CAD innovation with EPLAN P8. In *2015 23rd Telecommunications Forum, TELFOR 2015* (pp. 835–838). https://doi.org/10.1109/TELFOR.2015.7377595

Malhotra, P., TV, V., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). Multi-Sensor Prognostics using an Unsupervised Health Index based on LSTM Encoder-Decoder. *arXiv preprint arXiv:1608.06154*. http://arxiv.org/abs/1608.06154

Mishali, M., & Eldar, Y. C. (2009). Blind multiband signal reconstruction: Compressed sensing for analog signals. *IEEE Transactions on signal processing*, *57*(3), 993–1009.

Nuñez, D. L., & Borsato, M. (2018). OntoProg: An ontology-based model for implementing Prognostics Health Management in mechanical machines. *Advanced Engineering Informatics*, *38*, 746–759. https://doi.org/10.1016/j.aei.2018.10.006

O’Donovan, P., Leahy, K., Bruton, K., & O’Sullivan, D. T. J. (2015). An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. *Journal of Big Data 2015 2:1*, *2*(1), 1–26. https://doi.org/10.1186/S40537-015-0034-Z

Okoh, C., Roy, R., & Mehnen, J. (2017). Predictive Maintenance Modelling for Through-Life Engineering Services. *Procedia CIRP*, *59*, 196–201. https://doi.org/10.1016/j.procir.2016.09.033

Prajapati, A., Bechtel, J., & Ganesan, S. (2012). Condition based maintenance: A survey. *Journal of Quality in Maintenance Engineering*, *18*(4), 384–400. https://doi.org/10.1108/13552511211281552

Rana, A., Verma, A. K., & Srividya, A. (2016). Maintenance of Large Engineering Systems. In U. Kumar, A. Ahmadi, A. K. Verma, & P. Varde (Eds.), *Current Trends in Reliability, Availability, Maintainability and Safety* (pp. 599–609). Cham: Springer International Publishing.

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping Business With Artificial Intelligence. *MIT Sloan Mangement Review and The Boston Consulting Group*, (59181), 1–17. https://doi.org/10.3934/dcdsb.2008.10.925

Reliasoft. (2007). Basic Steps of Applying Reliability Centered Maintenance (RCM). *Reliability HotWire*. https://www.weibull.com/hotwire/issue76/relbasics76.htm. Accessed 29 September 2020

Rieger, T., Regier, S., Stengel, I., & Clarke, N. (2019). Fast predictive maintenance in Industrial Internet of Things (IIoT) with Deep Learning (DL): A review. In *CEUR Workshop Proceedings* (Vol. 2348, pp. 69–79).

Saxena, A., & Goebel, K. (2008). Turbofan engine degradation simulation data set. *NASA Ames Prognostics Data Repository*, 878–887.

Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *231*(9), 1670–1679. https://doi.org/10.1177/0954405415601640

Serradilla, O., Zugasti, E., Cernuda, C., Aranburu, A., de Okariz, J. R., & Zurutuza, U. (2020). Interpreting Remaining Useful Life estimations combining Explainable Artificial Intelligence and domain knowledge in industrial machinery. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1–8). IEEE.

Serradilla, O., Zugasti, E., Okariz, J. R. de, Rodriguez, J., & Zurutuza, U. (2021). Adaptable and Explainable Predictive Maintenance: Semi-Supervised Deep Learning for Anomaly Detection and Diagnosis in Press Machine Data. *Applied Sciences 2021, Vol. 11, Page 7376*, *11*(16), 7376. https://doi.org/10.3390/APP11167376

Serradilla, O., Zugasti, E., & Zurutuza, U. (2020). Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect. *arXiv preprint arXiv:2010.03207*.

UESystems. (2019). Understanding the P-F curve and its impact on reliability centered maintenance. *July 2019*. http://www.uesystems.com/news/understanding-the-p-f-curve-and-its-impact-on-reliability-centered-maintenance. Accessed 18 September 2019

UNE-EN 13306. (2018). *Maintenance. Maintenance terminology* (Vol. 2018). Génova, Madrid.

Vaidya, S., Ambad, P., & Bhosle, S. (2018). Industry 4.0 – A Glimpse. *Procedia Manufacturing*, *20*, 233–238. https://doi.org/10.1016/J.PROMFG.2018.02.034

Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis part I: Quantitative model-based methods. *Computers and Chemical Engineering*, *27*(3), 293–311. https://doi.org/10.1016/S0098-1354(02)00160-6

Vlasov, A. I., Echeistov, V. V, Krivoshein, A. I., Shakhnov, V. A., Filin, S. S., & Migalin, V. S. (2018). An information system of predictive maintenance analytical support of industrial equipment. *Journal of applied engineering science*, *16*(4), 515–522.

Welz, Z. A. (2017). Integrating Disparate Nuclear Data Sources for Improved Predictive Maintenance Modeling : Maintenance-Based Prognostics for Long-Term Equipment Operation.

Wiseman, M. (2006). A History of CBM (condition based maintenance). *omdec.com*. http://www.omdec.com/moxie/Technical/Reliability/a-history-of-cbm.shtml. Accessed 17 September 2019

Yuan, M., Wu, Y., & Lin, L. (2016). Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network. In *AUS 2016 - 2016 IEEE/CSAA International Conference on Aircraft Utility Systems* (pp. 135–140). https://doi.org/10.1109/AUS.2016.7748035

Zadeh, L. A. (1973). Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC-3*(1), 28–44. https://doi.org/10.1109/TSMC.1973.5408575

Zhang, W., Yang, D., & Wang, H. (2019). Data-Driven Methods for Predictive

Maintenance of Industrial Equipment: A Survey. *IEEE Systems Journal*, *13*(3), 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565

Zhang, X., Xu, R., Kwan, C., Liang, S. Y., Xie, Q., & Haynes, L. (2005). An integrated approach to bearing fault diagnostics and prognostics. In *Proceedings of the American Control Conference* (Vol. 4, pp. 2750–2755). https://doi.org/10.1109/acc.2005.1470385

Zhao, F., Tian, Z., & Zeng, Y. (2013). Uncertainty quantification in gear remaining useful life prediction through an integrated prognostics method. *IEEE Transactions on Reliability*, *62*(1), 146–159. https://doi.org/10.1109/TR.2013.2241216

Zhao, H., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, *16*(6), 3586–3592. https://doi.org/https://doi.org/10.1016/j.rser.2012.02.049

Zhou, X., Xi, L., & Lee, J. (2007). Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation. *Reliability Engineering & System Safety*, *92*(4), 530–534. https://doi.org/https://doi.org/10.1016/j.ress.2006.01.006

APPENDIX A

The methodology presented in this article and its main steps have been summarised in Figure 5, containing a complete visual scheme with all the tasks, principal subtasks and relations that form steps of the methodology. Being an open and modular methodology, its implementation is flexible and does not require the adoption of all its steps, only the ones marked by an asterisk in the diagram. Moreover, it contains the principal profiles involved in each stage and the deliverables like documents and models produced at them.

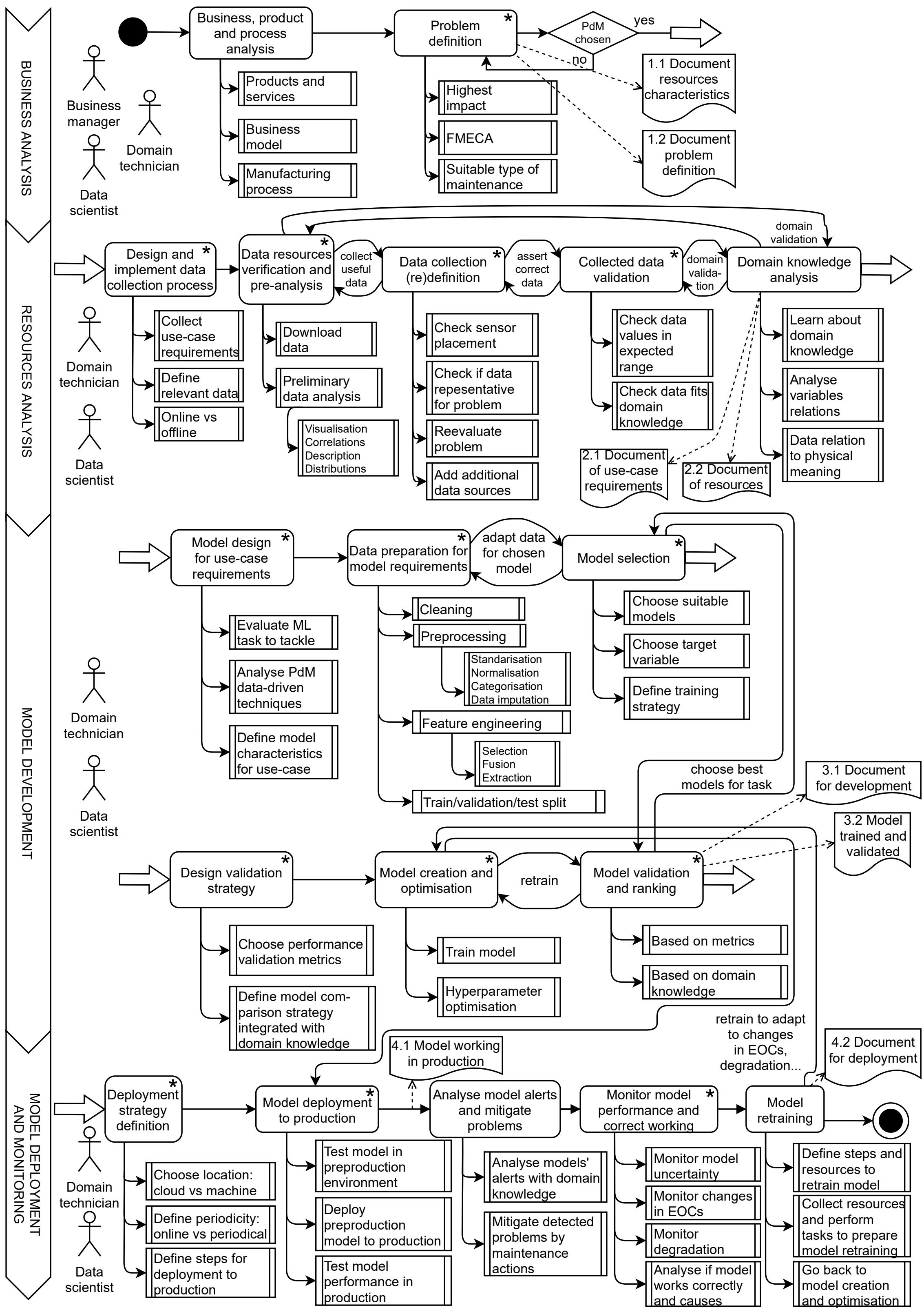

Figure 5 Detailed MEDADEK methodology, specifying its stages, steps and tasks in a flow diagram. It also contains the required profiles and indicates deliverables created at each stage.