

Quarterly Newsletter

CHAIR

iDANAE

INTELLIGENCE · DATA · ANALYSIS · STRATEGY

4Q22

**Industry 4.0:
predictive maintenance**



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

MSO *Management
Solutions*
Making things happen

Introduction

Throughout history, the emergence and development of new techniques, machines and inventions, or the transformation of industrial processes has generated disruptive advances that have transformed economies and society. This has occurred at different points in history: the first industrial revolution (18th century), with the application of water and steam to mechanical production; the second industrial revolution (19th century), with the transformation of assembly lines and mass production systems based on electricity; and in the 20th century, the appearance of computers and their incorporation into industrial processes, giving rise to the third industrial revolution. Finally, during the 21st century, the evolution of technology, the generation and storage of huge amounts of data, and the development of Machine Learning (ML) and Artificial Intelligence (AI) techniques have become transformational pillars in all sectors of the economy, giving rise to what is known as the Fourth Industrial Revolution¹.

In this context, the so-called Industry 4.0 arises, which is based on the principle of digitisation of industry in the different products and services of the value chain of companies² (incorporation of intelligent control units, robotisation, internet of things, application of infrastructures and big data technologies, use of cloud environments, interconnection of processes for sharing information based on the storage of large amounts of data, etc.). This digitisation makes it possible

to make decisions in factories without human intervention, and to adapt production to the means and resources available at any given time to maximise efficiency.

Industry 4.0 principles and techniques can be applied to a multitude of industrial fields (segmentation and understanding of customer preferences, product redesign, demand prediction, search for efficiencies, intelligent operation of machinery, remote monitoring, and traceability, etc.). Among them, one of the areas that has become more relevant is the so-called predictive maintenance: instead of performing periodic or corrective maintenance (replacement of broken parts or obsolete machinery), this approach aims to predict when a certain component or machine will fail or need to be replaced, based not only on pre-established criteria, but also on the monitoring of multiple elements associated with actual use, wear and tear, or the appearance of anomalous elements. This also makes it possible to plan maintenance operations in advance, reducing breakage or downtime of production lines, and unnecessary maintenance operations, thus minimising operating costs.

Indeed, predictive maintenance has proven to be one of the most cost-optimal maintenance type given its potential to achieve the best overall equipment effectiveness (OEE)³. Data from the U.S Department of energy⁴ suggest that implementing a predictive maintenance system yields remarkable results: ten times return on investment, 25-30% reduction in maintenance costs, 70-75% elimination of breakdowns, 35-45% reduction in downtime, and 20-25% increase in production.

This paper aims to provide an overview of predictive maintenance, the different AI and ML techniques that can be used to develop such mechanisms, as well as the challenges that need to be addressed for their effective implementation.



¹K. Schwab, 2016.

²D. Lukac, 2015.

³ivctechnologies, 2022.

⁴U.S. Department of Energy, 2010.

Predictive maintenance

Definition

Maintenance is defined as the combination of all technical, administrative, and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function⁵. Although some authors have classified up to 18 different types of maintenance⁶, the more widely used in the industry are the following:

- ▶ Reactive maintenance: actively repairing parts or components once they break down (i.e. “run it until it breaks”). This method yields high break down production times since it is not capable of predict or avoid a failure.
- ▶ Periodic maintenance: run on periodic time intervals for a certain number of components. This maintenance does not depend on the condition or state of the component. It avoids down production times by avoiding failures, but it increases maintenance costs.

In contrast to the traditional maintenance, the predictive maintenance aims to combine the advantages of the different approaches: avoiding down production times due to failures and lower maintenance costs by scheduling maintenance operations only when needed. In order to get this objectives, an ideal predictive maintenance system should have 10 desired properties⁷: be able to perform a quick detection and diagnosis,

be able to distinguish among different failure types (the so-called isolability feature), be robust, be able to identify novel or unusual behaviour, classify the error estimation, be able to adapt to changing environment, be easy to explain, count on minimal modelling requirements, incorporate real-time computation and storage handling, and be able to identify multiple faults.

To construct such a maintenance system, the types of failures are usually classified as component failure, environmental impact, human mistakes, and procedure handling⁸. To predict these failures, many components have to be usually monitored, applying different techniques (e.g. vibration analysis, particle testing, thermography, motor signal analysis, acoustic emission, pressure or temperature monitoring, etc.)⁹. All together these techniques commonly detected failures like imbalance cracks, fatigue, abrasive and corrosion wear, rubbing, defects and leak detection, among others.

⁵UNE-EN 13306, “Maintenance. Maintenance terminology,” Madrid, 2018.

⁶Idem.

⁷V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, 2003.

⁸D. Li and J. Gao, 2010.

⁹F. P. García Márquez, A. M. Tobias, J. M. Pinar Pérez and M. Papaelias, 2012.



Box 1. Physical model-based maintenance methods

There are many different approaches for the physical model-based maintenance methods. Two of the most used physical models for maintenance are the Arrhenius model and the Coffin-Manson mechanical crack growth model, whose theoretical background is mentioned below (these models can be developed and applied in multiple variations).

The Arrhenius model

The Arrhenius model has been applied to a variety of failure mechanisms although originally it was designed for those that depend of chemical reactions, diffusion processes or mitigation processes^a. The goal is to estimate the time to failure.

The operative equation is:

$$t_f = Ae^{(\Delta H/kT)}$$

where t_f is the time to failure, A is a scaling factor, ΔH is the activation energy, k is the Boltzmann constant, and T is the temperature at the point where the failure process takes place.

The Coffin-Manson mechanical crack growth model

This model is typically applied to mechanical failure, material fatigue, or material deformation^{b,c}. The goal is to estimate the number of cycles to failure.

The operative equation is:

$$N_f = Af^{-\alpha} \Delta T^{-\beta} G(T_M)$$

where N_f is the number of cycles to failure, A is a scaling factor, f is the cycling frequency, ΔT is the temperature range during a cycle, $G(T_M)$ is an Arrhenius term evaluated at the maximum temperature reached in each cycle, α is the cycling frequency exponent, and β is the temperature range exponent.

^aS. Arrhenius, 1889.

^bEngineering Statistic Handbook.

^cH. Cui, 2003.

Two main different approaches of predictive maintenance are defined: physical model-based and data-driven predictive maintenance.

- Physical model-based methods use expert knowledge to build a mathematical description of the degradation of the system monitored. For example, the resistance of the materials, the indication of engineers or producers, ex-ante test-driven information, etc. Although real data can be used for calibrating the parameters, the results are based on a *priori* theoretical models. This approach can be difficult to apply in case of large, complex systems.
- Data-driven methods predict the current state of the system by monitoring their condition during real-time use and capitalize historical data on actual use to learn from past failures. To be able to apply this method it is relevant to collect and store real-time data (e.g. incorporating sensors), and apply advance machine learning or data science algorithms.

A hybrid model could be developed, where real data would be used to periodically calibrate the parameters of the physical model. The implementation of such an approach would allow the two perspectives to be contrasted, so that one perspective could serve as a contrast for the other, thus improving the quality of the predictive maintenance system. In any case, since ML is incorporated in data-driven methods, this document focuses on said methods (for more information on physical model-based methods, see Box 1).

Data Driven methods

The design and implementation of a predictive maintenance system can leverage on specific methods to extract information from the actual industrial processes. When these methods are based on data and on the application of AI techniques, the methods are commonly referred to as “data driven methods”, and the predictive maintenance system can also be called “data driven system”.

The data driven methods are increasingly being developed due to the incorporation of techniques that allow to capture and store data directly from the productive processes. The data are collected by sensors or measuring tools across time, and form a time-series dataset. Therefore, data-driven methods designed to be applied to time-series must be incorporated. In addition, when the asset works correctly, different training datasets could be constructed, given different mechanical tolerances,

mount adjustments, or other factors. These elements add dimensionality and complexity to the problem.

To develop a data driven predictive maintenance system, two sequential stages have to be implemented^{10,11} (see figure 1):

1. The use of AI techniques to produce models and algorithms. It requires the development of three phases: two phases common in almost all data science projects, important for the preparation of the dataset (preprocessing and feature engineering), and a specific third phase for failure detection (the application of anomaly detection techniques).
2. The application of the results of said models to implement the process of maintenance, which is usually structured in three phases: diagnosis, prognosis and mitigation.

Use of AI Techniques

- ▶ **Preprocessing:** this phase consists in preparing the collected data for the following steps. Each ML model or architecture (including deep learning models) has different data structure requirements that need to be satisfied. A proper preprocessing step may also boost the model performance. In addition, since the data collected have usually a time-series structure, it is important to ensure that the data gathered at different timestamps becomes a time-series easy to handle (the process called data synchronization). This process is characteristic of predictive maintenance. There are some other techniques that can be applied in this step: data validation (to review the correctness of data), data cleansing (to remove or interpolate missing values), oversampling (to deal with highly imbalanced datasets), data augmentation (to deal with small datasets), encoding of categorical variables (to make it easier to handle by different models), scaling (to ensure the comparability across variables), or noise treatment (to model how noise affects the sensors), among others¹².

- ▶ **Feature engineering:** this phase consists in the extraction of information and the creation of new features that will be used as model input in later stages. Some of the most common techniques usually applied are the extraction of statistical features over time, the extraction of time/frequency relation of features, the application of dimensionality reduction techniques (such as PCA or feature selection) to reduce the feature space and the complexity of the model, the combination of existing features, or the analysis of existing correlations between variables.
- ▶ **Anomaly detection:** the objective of this phase is to detect whether an asset is working under normal conditions. The approach to tackle this problem depend on different conditions: the existence of labelling for the available data, the dimensionality of the labels (i.e. multiple or binary classification), the need for including the time-dependency, using for example ML algorithms such as RNN (e.g. LSTM or GRU), etc. In case of unlabelled data, multiple unsupervised methods can be applied, such as outlier detection or clustering algorithms, together with other traditional statistical methods.

The objective of the techniques of an anomaly detection framework is to study the probability distribution of the data and find outliers as points which do not fit the usual distribution. Among others, there are some techniques that can be useful:

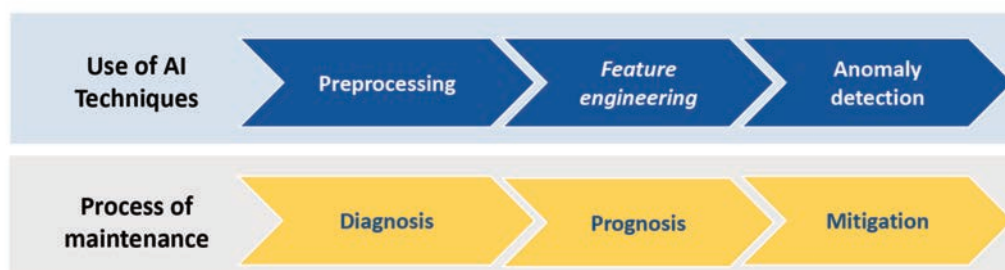
- Unsupervised learning methods based on density computations, such as k-NN, local outlier factor (LOF) or local correlation integral (LOCI). The underlying principle is that normal behaviour corresponds to high density regions, and anomalies correspond to low density regions.

¹⁰A. Prajapati, J. Bechtel and S. Ganesan, 2012.

¹¹O. Serradilla, E. Zugasti, J. Rodríguez and U. Zurutuza, 2022.

¹²Z. Zhang and P. Zhang, 2022-

Figure 1: stages and phases of predictive maintenance



- Clustering methods such as K-means or hierarchical clustering. For these methods, normal data belong to large clusters and anomalies correspond to small or even no clusters.
- Deep learning techniques, such as self-organizing maps (SOM) or autoencoders (both with vanilla neural networks or RNN to incorporate time-dependency). In this case outliers are detected via a reconstruction error, assuming that most of the training data represent the normal working conditions of the devices or system monitored.
- Other methods can also be considered, such as gaussian mixture models (GMM), kernel density estimation (KDM), KL-divergence, histogram-based outlier detection, or box-plot analysis, among others.
- Supervised modeling: to use time-series models (e.g. ARIMA, RNN) to directly determine the time to failure. This approach is only viable when labelled data of previous failures is available.
- Unsupervised analysis: to compute and monitor the HI over time.

► **Mitigation:** the last step consists of planning maintenance operations to minimize costs and losses, given all the information provided in previous steps. This step does not involve any kind of advance analytics methods, but it is key to present the information from data-driven models in a concise and understandable manner, so that responsible people involved in the maintenance, or in the making-decision process, can take appropriate business decisions. In this step, the interpretability and the explainability of the different methods applied in previous steps could be crucial from a business perspective¹⁵.

Process of maintenance

- **Diagnosis:** this phase aims to quantify how dangerous an anomaly is, i.e., it measures if the anomaly will evolve in a failure or not. An anomaly that does not evolve in failure does not necessarily mean the anomaly detection model is wrong (if all anomalies were failures there would be no way to predict when one might occur, other than giving a simple probability estimator). It is also plausible to have small anomalies that can evolve in future failures.

There are multiple ways to tackle this step, which will depend on the technique used in the anomaly detection step. It is common to define a metric called Health Index (HI) that measures how close/far is the current behaviour of the system compared to normal working conditions. It can either be a percentage of deviation compared to normal data or a measure of degradation in a numerical scale. The exact definition of the HI will depend on the technique used (e.g., if a clustering method has been used, the distance to the closest cluster is a good candidate, whereas under a density-based outlier detection algorithm, the local density deviation is a good candidate).

- **Prognosis:** This step is intended to compute or estimate the remaining useful life of the system, or the time or number of cycles until next failure. Data to compute it with a high level of accuracy is normally not available. However, monitoring the health index and comparing it to normal working conditions is a possible approach. To do this, there are different methods to tackle this step¹³:

- Similarity-based method: to compare current behaviour with past run-to-failure data.
- Statistical method: to use statistical measures like the distribution of time-to-failure or survival models¹⁴.

Main challenges

The implementation of a data-driven predictive maintenance system is a complex process, since there are many elements that must be combined together. Besides financial or economic considerations, there are two main types of challenges that need to be considered related to data and models.

- Challenges related to Data:

- It is key to collect up-to-date data from the many aspects of the functioning of the industrial device, machine or process. These data are usually produced in real time, automated way, and therefore sensors that measure and capture the data must be developed and installed.
- These data from sensors sometimes are of different types (recording of physical variables, sounds, pictures, etc.), and therefore have to be initially processed to be used for the ML algorithms.
- The data have to be stored in an efficient way, and have to be available (even in real time) to be sent to other parts of the maintenance process.

¹³O. Serradilla, E. Zugasti, J. Rodríguez and U. Zurutuza, 2022.

¹⁴Z. Zhang and P. Zhang, 2015.

¹⁵X. T. Liu, F. Z. Feng and A. W. Si, 2014.

- The very large amount of data imposes another challenge in the storing and the processing of the information.
 - A specific consideration has also to be done regarding the data governance and data quality, as for many other systems that relies on information.
 - Finally, is important to ensure that the data are protected against cyberattacks, and that the regulation on data protection is also complied with.
- Challenges related to models:
- One of the key challenges is to ensure that there is a proper analytics function, which includes the existence of areas within the organization dedicated to the development and the validation of the models, counting with highly skilled people.
 - In addition, given the highly demanding performance requested to the modelling area, it is important to count on a comprehensive modelling framework that ensures the correctness and efficiency of the modelling process and monitoring, as well as a model governance framework established to manage all models that can be developed.
- Both elements must rely on technological frameworks that allow a proper and continuous integration of the different processes (from model planning, development, testing, implementation, use, monitoring and deprecation), being a MLOps framework a common practice.
 - It is important to ensure that the models are constructed to build prediction systems (and not only detection), to ensure that the anticipated management can be successfully implemented.
 - One of the major challenges is the real-time feature, that requires highly demanding systems in terms of accuracy, timeliness, coordination, and technical devices to ensure a proper implementation of a predictive modelling system.



Conclusions

New technologies, the Internet of Things (IoT), big data, availability of data, the application of ML techniques, and the increase in computing power and storage capacity in recent years have resulted in a new industrial revolution, the so-called Industry 4.0. One of the areas that have attracted attention is the transforming of the maintenance processes to a data-driven approach, capitalizing the production of new information on the industrial processes and the application of the machine learning techniques to make the maintenance of the machinery and devices more efficient, in terms of cost and time saving.

This new type of maintenance aims to predict when and where a future failure will happen to schedule maintenance operations when appropriate. This has multiple advantages compared to traditional maintenance strategies (such as reactive or periodic maintenance), which show higher costs.

The machine learning techniques can help in the design and development of a data driven predictive maintenance system. However, there are several challenges that have to be addressed in order to implement such a successful system.



Authors

Ernestina Menasalvas (UPM)

Manuel Ángel Guzmán (Management Solutions)

Segismundo Jiménez (Management Solutions)

Sergio Ruiz (Management Solutions)

Bibliography

- K. Schwab, "The Fourth Industrial Revolution: what it means, how to respond," World Economic Forum., 14 Jan 2016. [Online].
- D. Lukac, "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 (TELFOR), 2015.
- ivctechnologies, [Online]. Available: <https://ivctechnologies.com/2017/08/29/reactive-preventive-predictive-maintenance/>. [Accessed 14 11 2022].
- U.S. Department of Energy, "Operations & Maintenance Best Practices Release 3.0," 2010.
- UNE-EN 13306, "Maintenance. Maintenance terminology," Madrid, 2018.
- V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, "A review of process fault detection and diagnosis," Computers & Chemical Engineering, vol. 27, no. 3, pp. 293-311, 2003.
- D. Li and J. Gao, "Study and application of Reliability-centered Maintenance considering Radical Maintenance," Journal of Loss Prevention in the Process Industries, vol. 23, no. 5, pp. 622-629, 2010.
- F. P. García Márquez, A. M. Tobias, J. M. Pinar Pérez and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," Renewable Energy, vol. 46, pp. 169-178, 2012.
- S. Arrhenius, "Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte," Zeitschrift für Physikalische Chemie, vol. 4U, no. 1, pp. 96-116, 1889.
- Engineering Statistic Handbook.
- H. Cui, "Accelerated temperature cycle test and Coffin-Manson model for electronic packaging," in Annual Reliability and Maintainability Symposium, 2005. Proceedings., 2005.
- A. Prajapati, J. Bechtel and S. Ganesan, "Condition based maintenance: a survey," Journal of Quality in Maintenance Engineering, vol. 18, no. 4, pp. 384-400, 2012.
- O. Serradilla, E. Zugasti, J. Rodriguez and U. Zurutuza, "Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects," Applied Intelligence, vol. 52, no. 10, pp. 10934-10964, 2022.
- Z. Zhang and P. Zhang, "Seeing around the corner: an analytic approach for predictive maintenance using sensor data," Journal of Management Analytics, vol. 2, no. 4, pp. 333-350, 2015.
- X. T. Liu, F. Z. Feng and A. W. Si, "Condition based monitoring, diagnosis and maintenance on operating equipments of a hydraulic generator unit," IOP Conference Series: Earth and Environmental Science, vol. 15, no. 4, p. 042014, 2012.



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

The Universidad Politécnica de Madrid is a multisector and multi-disciplinary Public Law Entity, which carries out teaching, research and scientific and technological development activities.

www.upm.es



Management Solutions is an international consulting firm, focused on business, finance, risk, organization, technology and process consulting, operating in more than 40 countries and with a team of over 3,200 professionals working for more than 1,500 clients worldwide.

www.managementsolutions.com

For more information visit

blogs.upm.es/catedra-idanae/