

A data-driven framework for supporting the total productive maintenance strategy

Laura Lucantoni^{*} , Sara Antomarioni , Filippo Emanuele Ciarapica, Maurizio Bevilacqua

Department of Industrial Engineering and Mathematical Science, Università Politecnica delle Marche, Via Brecce Bianche, 12, Ancona 60131, AN, Italy

ARTICLE INFO

Keywords:

Total productive maintenance
Lean 4.0
Lean automation
Overall equipment effectiveness
Association rules
Network analysis

ABSTRACT

Maintaining machinery efficiency is a challenge in modern manufacturing, with companies adopting Total Productive Maintenance (TPM) strategies and using Overall Equipment Effectiveness (OEE) as a key metric for operational improvements. Industry 4.0 features have provided advanced data processing tools to assess machine performance, but effectively leveraging this data to enhance OEE remains challenging due to their multisource nature.

To fully leverage the vast amount of data, new analytical methodologies are essential, not only to identify production events negatively impacting OEE but also to enhance positive events, thereby supporting strategic decision-making for TPM implementation.

This study addresses this need by introducing a new data-driven methodology that integrates lean TPM principles with I4.0 analytical tools. Specifically, the proposed framework provides company managers with a structured approach, employing a dual application of Association Rule Mining (ARM) and Network Analysis (NA) to discover hidden relationships within OEE factors—availability, performance, and quality—and between key factors and operational variables. By assessing the impacts on equipment effectiveness, this methodology aims to prevent negative domino effects from adverse events while promoting positive events. Finally, a data-driven decision support system provides a long-term roadmap for determining which TPM pillars and associated actions should be prioritized for improvement.

A case study demonstrates the framework's effectiveness, leading to key outcomes: implementation of 24 kaizen actions, achievement of 7 out of the 8 TPM pillars, and a 14% improvement in OEE.

1. Introduction

The development of Industry 4.0 (I4.0) has significantly transformed production processes through advanced data analytics, with Total Productive Maintenance (TPM) showing substantial potential for convergence with I4.0 technologies from a Lean Production (LP) perspective (Wolska et al., 2023). TPM has long been a crucial methodology for improving asset availability, reducing downtimes, and optimizing maintenance practices (Chan et al., 2005). Focusing on various “pillars”—such as focused improvement, autonomous maintenance, quality-oriented maintenance, planned maintenance, early equipment management, training and education, safety, environment and health, and administrative TPM—TPM offers a comprehensive strategy for enhancing operational efficiency in manufacturing environments.

Despite TPM's importance, challenges emerged in its application to

dynamic, data-intensive environments, largely due to delays in decision-making and inefficient data integration (Djatna & Alitu, 2015; Lopes et al., 2016). To address these limitations, Lean Automation (LA), or Lean 4.0, integrated I4.0 technologies with LP methods, such as TPM, to ensure continuous flow and minimize failure rates (Tortorella et al., 2021; Valamede & Akkari, 2020). However, many companies struggled to leverage these technologies effectively due to the need for advanced data-driven methodologies capable of providing actionable insights and supporting decision-making (Alenizi et al., 2023; Dogana & Birant, 2021; Sajid et al., 2021).

Nowadays, while integrating I4.0 technologies with TPM is still an emerging strategy (Samadhiya et al., 2024; Mouhib et al., 2024), empirical research on advanced data science applications within TPM is still underexplored (Pozzi et al., 2024).

The proposed framework aims to fill these existing gaps by

^{*} Corresponding author.

E-mail addresses: l.lucantoni@staff.univpm.it (L. Lucantoni), s.antomarioni@univpm.it (S. Antomarioni), f.e.ciarapica@univpm.it (F.E. Ciarapica), m.bevilacqua@univpm.it (M. Bevilacqua).

<https://doi.org/10.1016/j.eswa.2024.126283>

Received 23 February 2024; Received in revised form 9 December 2024; Accepted 23 December 2024

Available online 28 December 2024

0957-4174/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1
Industrial applications on implementing TPM through DM techniques.

Ref.	Objective	Approach	Pros	Cons
(Djatna & Alitu, 2015)	OEE improvement and decision-making	Pareto chart, Fishbone diagram, ARM.	Full investigation of OEE factors and data mining.	Lack of TPM implementation phase, data manually collected and traditional data analytics techniques.
(Candra et al., 2017)	OEE improvement	Pareto chart, Regression analysis, Fishbone diagram.	OEE factors assessment and testing activity.	Implementation of the Focused Maintenance TPM pillar only. Lack of advanced data analytics techniques. Short-term research period.
(Braglia et al., 2020)	Equipment behavior prediction	k-NN, Random Forest, naïve Bayes, and Decision Tree.	Exploitation of advanced data analysis techniques for knowledge extraction from big data.	Implementation of the Preventive Maintenance TPM pillar only. Lack of OEE assessment.
(Antomarioni et al., 2023).	OEE improvement and decision-making	Failure Mode and Effect Analysis and ARM	Data mining and application evidence.	Investigation of the OEE Availability factor only. Implementation of just four of the eight TPM pillars.
(Ale et al., 2024)	Predicting the Remaining Useful Life of a specific component.	Additive Manufacturing, Automatic Train Control, smart sensors, IoT, cloud computing, AI	Adoption of AI	Implementation of the Preventive Maintenance TPM pillar only. No explicit reference to the OEE assessment.
(Paiva et al., 2024)	Maintenance policy decision-making	Sequential Pattern Mining	Co-occurrence analysis of events for decision-making.	Implementation of Autonomous Maintenance TPM pillar only. Unavailability of extended period of data.

developing a novel data-driven methodology that leverages advanced Data Mining (DM) techniques to enhance the implementation of most TPM pillars. The key contributions of the methodology can be summarized as follows:

- Comprehensive analytics applied to TPM environment: unlike traditional approaches or blended methodologies, this framework utilizes the dual application of Association Rule Mining (ARM) and Network Analysis (NA) for advanced decision-making. It allows a thorough investigation of hidden relationships between Overall Equipment Effectiveness (OEE) factors (availability, performance, and quality) and all operational variables. This holistic approach provides deeper insights into system performance and uncovers patterns that go beyond the scope of actual techniques such as focusing only on preventive maintenance and availability losses.
- OEE impact evaluation via different analytics techniques: a unique aspect of this methodology is its comprehensive evaluation of both positive and negative influences on OEE. By leveraging a detailed metrics-driven classification—incorporating support and confidence factors, out-degree, in-degree, and centrality degree—it effectively mitigates the negative domino effect from adverse events while fostering beneficial events through a reward system, offering a more nuanced approach compared to existing methods.
- Application to an industrial case study: the framework's effectiveness is validated through a practical, longitudinal case study, which not only demonstrates its real-world applicability but also aligns with long-term organizational goals. This case study provides tangible evidence of the framework's positive impact on OEE, offering a structured approach for managers to make informed decisions and implement TPM more effectively.

The remainder of the paper is organized as follows: [Section 2](#) presents a review of relevant literature on data-driven approaches for TPM implementation. The structure and methodology of the proposed framework are described in [Section 3](#), while [Section 4](#) details the industrial case study. Lastly, [Section 5](#) discusses the framework's implications and presents the conclusions.

2. Literature review

The importance of maintenance interventions in preserving machinery has grown alongside the increasing complexity of production processes and technologies. The exponential rise in data availability directed researchers' attention toward more precise data analytics. In this context, LA emerged for fostering a data-driven culture within LP by leveraging I4.0 technologies to enable a data-informed decision-making

process. Furthermore, the adoption of I4.0 and data-driven methodologies offers significant benefits for the sustainability of production processes. For example, performance improvements achieved through LA can enhance energy efficiency and promote environmental sustainability (Tortorella, Fogliatto, et al., 2021).

Despite these advantages, research on the impact of LA on production systems was limited up until 2021 (Tortorella, Saurin, et al., 2021), largely due to rigid automation in many manufacturing companies (Polge et al., 2020). To deepen the investigation, a literature search was conducted across Scopus, Web Of Science, IEEE Xplore, ScienceDirect, and OpenAlex, focusing specifically on TPM within the Industry 4.0 context. The search employed the string ("Total productive maintenance" AND "Industry 4.0" AND "case*study") restricted to titles, abstracts, and keywords. This search returned 16 papers addressing the industrial applications of TPM in the context of I4.0. These studies varied in focus, techniques, and application contexts. The majority of the studies focus on the Internet of Things (IoT) or other methods such as multicriteria decision-making, value stream mapping, or digital transformation for maintenance systems, but do not explicitly utilize DM for TPM.

While statistics analyze large datasets, DM differs by focusing on pattern identification and process classification. For example, (Pozzi et al., 2024) discussed the role of data science in supporting LP through interviews with manufacturing companies but did not apply DM techniques. Similarly, some studies focused on real-time monitoring systems for TPM but relied on IoT sensors without using advanced DM methods (Mendes et al., 2023b, 2023a). Specifically, only two industrial applications applied DM techniques for TPM implementation: Antomarioni et al. (2023) and Ale et al. (2024).

To enhance the literature review, the scope was broadened by searching for industrial applications of TPM using data mining techniques with the string ("Total productive maintenance" AND ("data mining" OR "data analy*" OR "data-driven") AND ("case study" OR "case-study" OR "application*" OR "compan*"). This search identified a total of 29 papers published between 2015 and 2024 across the previously mentioned databases. However, only seven of these were found to effectively align with the research objectives outlined, including the two already referenced studies. Details of current research on DM techniques applied to TPM are provided in [Table 1](#), with a particular focus on improving OEE. OEE is a critical performance indicator for TPM, as it integrates equipment availability, process efficiency, and quality rate into a single, comprehensive metric (dos Reis et al., 2019). Current research on data-driven methodologies supporting TPM strategy largely focused on applying traditional techniques, such as Pareto and Ishikawa (Candra et al., 2017). A few works in literature attempted to integrate DM with conventional methods. For instance, (Djatna & Alitu, 2015)

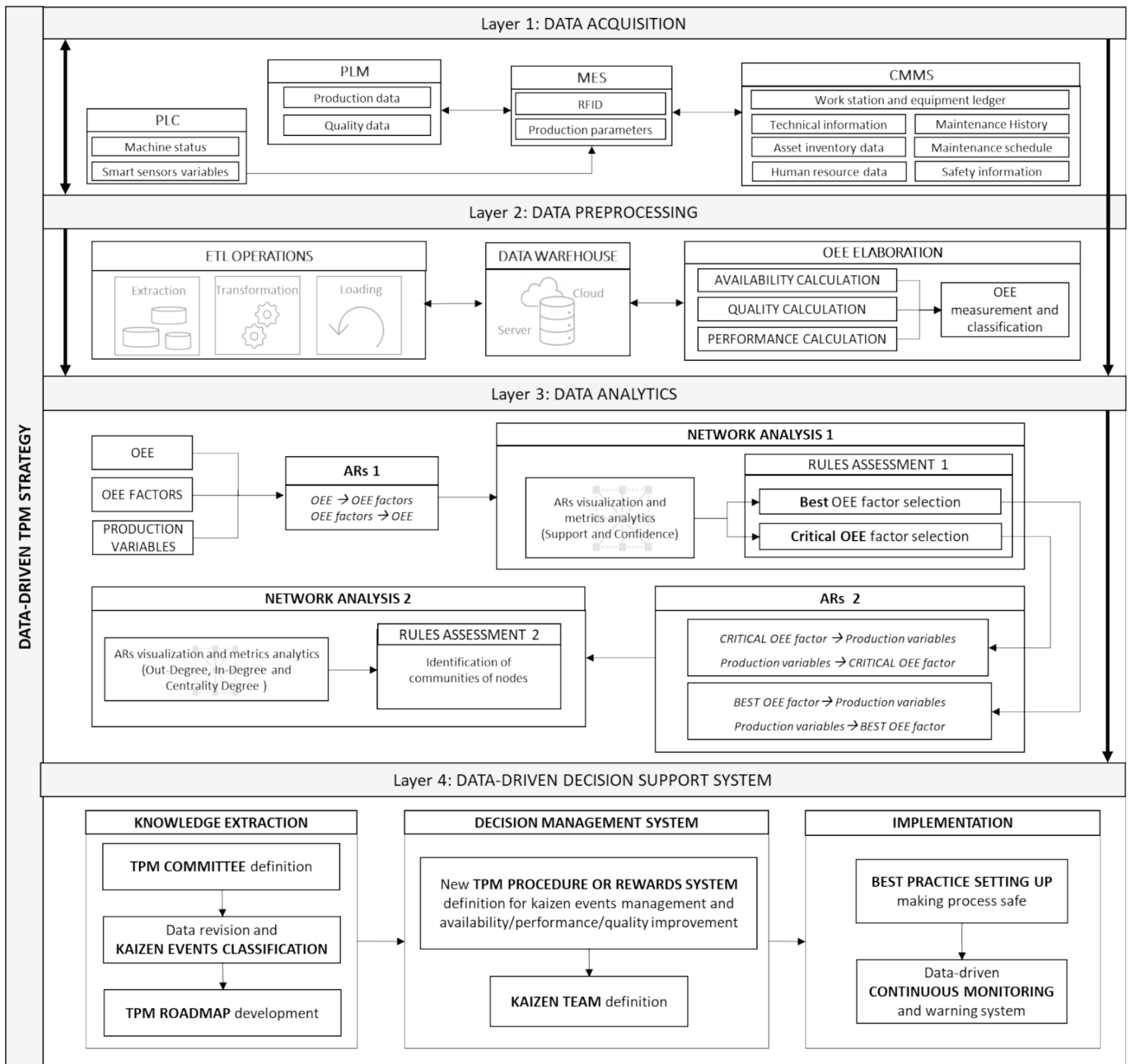


Fig. 1. Four-layer framework for data-driven TPM strategy.

combined ARM with Pareto charts and Fishbone diagrams to enhance OEE. While this approach added value to literature, it currently remains largely conventional, not fully exploring advanced DM techniques or Artificial Intelligence (AI) for deeper insights into OEE.

A few studies applied more advanced DM techniques, such as Random Forest and AI, to TPM. For example, Braglia et al. (2020) and Ale et al. (2024) utilized these methods to address specific TPM pillars, such as autonomous and preventive maintenance. However, their focus remained narrow, primarily addressing OEE factors such as availability and maintenance, while neglecting other critical OEE losses, such as performance and quality.

Paiva et al. (2024) employed Sequential Pattern Mining to optimize operation and maintenance policies for improving OEE. This approach analyzes temporal sequences of data, unlike ARM, which is an unsupervised technique. Although this study advances beyond conventional methods, it is limited by the lack of extended data periods and

emphasizes the need for more robust predictive and preventive maintenance practices. In this regard, ARM presented potential insight for maintenance prediction, though its application remains limited in this domain. Antomarioni et al. (2023) presented in fact a preliminary data-driven implementation of TPM, combining ARM with the traditional method Failure Mode and Effect Analysis. While it is among the few studies applying ARM for implementing TPM, along with Djatna & Alitu (2015), it is limited in scope, addressing only OEE availability losses and focusing on just four of the eight TPM pillars (Autonomous Maintenance, Planned Maintenance, Quality Maintenance, and Continuous Improvement). Moreover, combining ARM with traditional techniques like Failure Mode and Effect Analysis (FMEA) restricts the full potential of DM to uncover deeper insights into the TPM process.

Thus, while data-driven methodologies have been applied to TPM, much of the literature remains constrained to a limited subset of TPM pillars and availability losses, primarily using conventional or blended

data analytics techniques. A research gap exists in the development of advanced DM frameworks that leverage fully advanced DM techniques within I4.0 to implement most of the TPM pillars and improve OEE by analyzing all of its factors. Furthermore, advanced DM frameworks that simultaneously analyze both loss and gain production factors within TPM are still lacking in the literature, with no such framework fully explored as of 2022 (Lucantoni et al., 2022).

As previously mentioned, the LA concept is still relatively recent: longitudinal studies examining the impact of I4.0 on TPM were largely absent as of 2021 (Tortorella, Fogliatto, et al., 2021). Only two studies in the Scopus database focus on long-term monitoring of the integration of lean maintenance practices with I4.0, but neither of them directly addresses the implementation of all TPM pillars (Amjad et al., 2021; Kufre et al., 2024).

In conclusion, to the best of the author's knowledge, there is a significant gap in the literature regarding the application of fully advanced DM frameworks for comprehensive TPM implementation, OEE assessment, and decision-making, as well as the simultaneous analysis of both loss and best production viables. Specifically, there is no evidence of the joint application of ARM and NA as the driving tools for implementing TPM. Given the potential benefits of combining ARM and NA to improve OEE (Lucantoni et al., 2023), a four-layer data-driven approach using ARM and NA is proposed in the current paper to address these gaps and offers a guideline for long-term TPM implementation.

3. Research approach

A four-layer framework (Fig. 1) for TPM implementation using data-driven techniques, emphasizing continuous improvement in LA.

The proposed framework consists of the following four layers:

- 1) Data acquisition (layer 1): Focused on collecting data and defining problems.
- 2) Data preprocessing (layer 2): Dedicated to measuring and classifying OEE.
- 3) Data analytics (layer 3): Aimed at identifying key issues and event communities.
- 4) Data-driven decision support system (layer 4): Designed to implement a proactive TPM program.

Data acquisition and preprocessing from multiple logs provide insights into current and future TPM performance, including OEE, losses, and wastes. Data analytics identifies root causes and factor communities

positively and negatively impacting OEE. A data-driven Decision Support System then supports planning, improvement activities and reward systems. Lastly, a real case study demonstrates the implementation of TPM.

3.1. Layer 1: data acquisition

This fundamental layer focuses on collecting records through intelligent sensors and integrating primary data sources with a Computerized Maintenance Management System (CMMS) for managing maintenance data. Key functionalities include (a) Assets Management for equipment and historical maintenance records, (b) Maintenance Management for scheduling activities, and (c) Report Management for data processing and key performance indicators generation. The Data Acquisition level proposed in this work has been generalized to facilitate adaptability within different industrial realities, hence the following data sources can be integrated with CMMS depending on the typology of an information system:

- Manufacturing Execution System (MES) for the real-time monitoring activities, workers, machinery parameters, and downtimes.
- Product Lifecycle Management (PLM) for managing processes and product data.
- Programmable Logic Controller (PLC) for real-time data collection from intelligent sensors and integration with MES.

Quality information should also be captured by fitting the equipment with quality control devices such as poka-yoke.

In the next layer, data will be processed to be able to extract useful information: indeed, raw data collected and stored into different information systems may contain errors, inconsistencies, missing values, or irrelevant information that can distort analysis and lead to inaccurate conclusions. Through preprocessing, these issues are addressed by cleaning the data, standardizing formats, and removing redundancies.

3.2. Layer 2: Data preprocessing

This layer calculates and classifies OEE as a global indicator of production efficiency from a TPM perspective. Steps include: In particular, the data processing level manages the previously collected data by integrating all information servers in one maintenance management system. The implementation complexity of this level depends on the number of information systems available for integration and the quan-

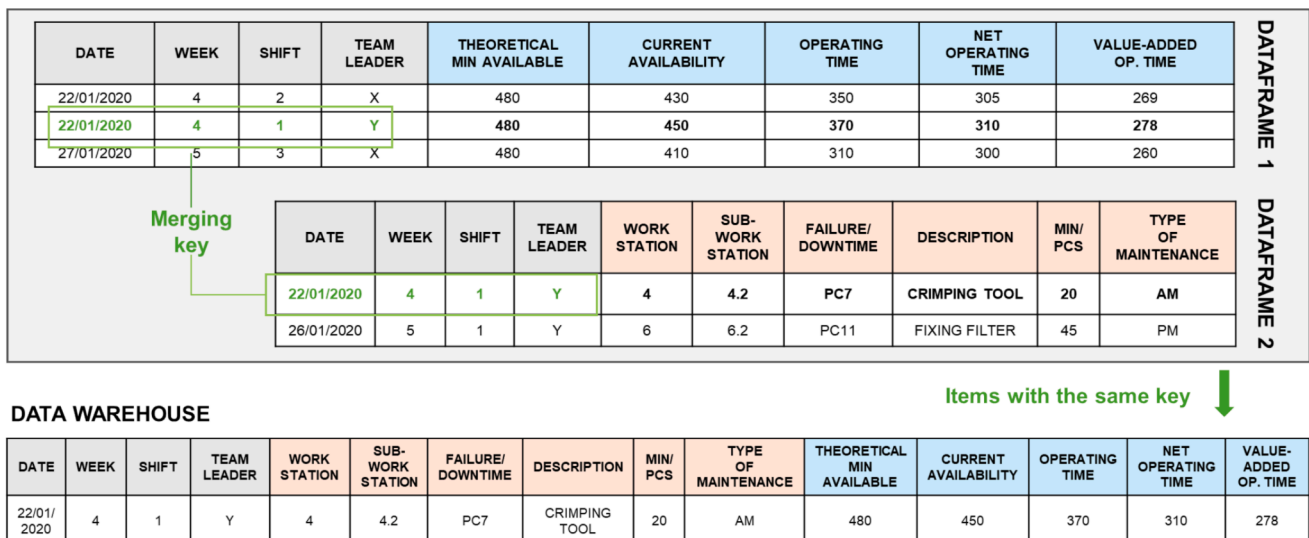


Fig. 2. Integration of operational and maintenance data in the data warehouse.

Table 2
Problem Codes details.

ID	Problem Code	Description	ID	Problem Code	Description
PC1	Problem Code 1	Calibration PRV1 + Clamp	PC28	Problem Code 28	Refilling Filter
PC2	Problem Code 2	Calibration PRV2 + Clamp	PC29	Problem Code 29	Refilling Inlet Valve
PC3	Problem Code 3	Conveyor + Loading Coil	PC30	Problem Code 30	Refilling Pin-Spring
PC4	Problem Code 4	Conveyor + Loading Inlet Valve	PC31	Problem Code 31	Refilling Pin-Spring 2
PC5	Problem Code 5	Conveyor + Loading Pin	PC32	Problem Code 32	Refilling PRV 1
PC6	Problem Code 6	Conveyor + Loading Sealing Cup	PC33	Problem Code 33	Refilling PRV 2
PC7	Problem Code 7	Crimping tool	PC34	Problem Code 34	Refilling Pump Body
PC8	Problem Code 8	DMC reader	PC35	Problem Code 35	Refilling Seat Cup
PC9	Problem Code 9	Elevation control	PC36	Problem Code 36	Refilling Semicons
PC10	Problem Code 10	Fixing Koenig Cap	PC37	Problem Code 37	Refilling Spacer
PC11	Problem Code 11	Fixing Filter	PC38	Problem Code 38	Refilling Sphere
PC12	Problem Code 12	Fixing Inlet Fitting	PC39	Problem Code 39	Refilling Spring
PC13	Problem Code 13	Fixing Inlet Valve	PC40	Problem Code 40	Refilling Spring – Stopper
PC14	Problem Code 14	Fixing Sandwich	PC41	Problem Code 41	Refilling Spring-Retainer
PC15	Problem Code 15	Fixing Seat	PC42	Problem Code 42	Refilling Washer Cup
PC16	Problem Code 16	Fixing Sphere	PC43	Problem Code 43	Replacing Oring
PC17	Problem Code 17	Fixing Spring – Stopper	PC44	Problem Code 44	Replacing Slide
PC18	Problem Code 18	Helium room 1	PC45	Problem Code 45	Replacing pump body
PC19	Problem Code 19	Helium room 2	PC46	Problem Code 46	Robot 1 replacement
PC20	Problem Code 20	Helium room 3	PC47	Problem Code 47	Robot 2 replacement
PC21	Problem Code 21	Helium room 4	PC48	Problem Code 48	Saracen
PC22	Problem Code 22	HP Anomaly	PC49	Problem Code 49	Welding Coil
PC23	Problem Code 23	Hydraulic control unit	PC50	Problem Code 50	Welding Damper Cup
PC24	Problem Code 24	Insertion Plunger	PC51	Problem Code 51	Welding Inlet Fitting
PC25	Problem Code 25	Plunger sealing	PC52	Problem Code 52	Welding Sealing Cup
PC26	Problem Code 26	Refilling Coils	PC53	Problem Code 53	Welding Valve Body
PC27	Problem Code 27	Refilling Damper			

tity and quality of their data. At this level, data processing and management are structured as follows:

- Data warehouse*, for centralizing data storage and integrating various sources. Both cloud applications and local servers can be utilized to support storage, preliminary analysis, and AI interactions. Continuous monitoring, such as through ETL operations, ensures data integrity, reliability, and continuous improvement, requiring seamless interaction between the warehouse, ETL, and analysis process.
- ETL operations involve data Extraction, Transformation, and Loading to enhance the reliability of findings. Data extraction integrates multiple sources, transformation focuses on cleaning and

standardizing relevant data, and loading consolidates this data into a consistent warehouse.

- OEE elaboration*, calculates the OEE as a key indicator of production line efficiency, defined as (Muchiri & Pintelon, 2008):

$$\text{PERFORMANCE} \times \text{AVAILABILITY} \times \text{QUALITY} = \text{OEE} \quad (1)$$

The accuracy of OEE values is highly dependent on the quality of the collected data. As a key KPI of the TPM strategy, OEE is classified into three ranges based on the complexity of the data warehouse. Effective OEE tracking requires thorough documentation of unplanned downtimes and fault events.

3.3. Layer 3: data analytics

The Data Analytics layer drives TPM by analyzing and assessing the interactions between OEE values and its loss factors through Association Rules (ARs) and Network Analysis for communities' identification. This process can generate thousands of rules, varying in the number and type of factors analyzed, however, reviewing these rules can be time-intensive for managers. To overcome this issue, the Data-Driven TPM framework outlines the following steps for piloting TPM implementation:

Identification of key OEE factors: ARs1 allows the selection of the factors most critical in lowering or raising OEE, helping managers prioritize their focus on the most impactful factors.

Network Analysis 1: Rules classification into two groups under RULES ASSESSMENT 1:

- OEE Factors → Critical factor
- OEE Factors → Best factor.

Production variables analysis: Variables identification significantly affecting the critical and best OEE factors using ARs 2. The objective of this step is to assist managers in saving time and prioritizing actions targeting these key variables.

Network Analysis 2: Expanding RULES ASSESSMENT 2 to provide a comprehensive analysis of ARs 2, focusing on identifying communities of nodes (operational variables) responsible for root causes of lower and higher OEE values. The goal is to assess the influence of production variables on other variables, factors and OEE, using deeper metric analytics. This analysis enables managers to identify and prioritize the best and worst production events, referred to as Kaizen events, for effective management and improvement.

3.3.1. Association rule mining

The Association Rule Mining aims to identify the relations among attributes and values stored in large datasets that frequently co-occur (Buddhakulsomsiri et al., 2006). The objective is to capitalize on the production process's data to extract information and knowledge about them from a TPM perspective.

A set of Boolean data (i.e., items) $I = \{i_1, i_2, \dots, i_n\}$ and transactions $T = \{\tau_1, \tau_2, \dots, \tau_m\}$ are given. Each element of T is composed of an itemset belonging to I .

$\alpha \rightarrow \beta$ is an Association Rule, that is an implication between α and β , which are itemsets taken from I ($\alpha, \beta \subseteq I$) and having null intersection ($\alpha \cap \beta = \emptyset$). In this approach, Support (Supp) and Confidence (Conf) are the metrics used for assessing rule's quality, since, according to the existing literature, they are the most widely used. As per formula (2), Supp determines if a rule is statistically relevant, being the ratio between the transactions containing both α and β over the cardinality of the transaction set T . The confidence of a rule (3) is found as the ratio between the support of the rule over the support of the antecedent itemset. Indeed, it represents the conditional probability of having β in a transaction containing α .

Table 3
Final database variables.

Group	Variable	Description	Type
Shift Details	Date	Date of data acquisition	Numerical
	Week	Week number within the year	Numerical
	Shift	Shift number	Numerical
	Team Leader	Name of the shift team leader	Textual
Failure Data	Workstation	Workstation number	Numerical
	Sub-workstation	Station number within the workstation	Numerical
	Failure/Downtime	Fault event code	Textual
	Description	Description of the fault type	Textual
	Duration (Min/Pcs)	Duration of the intervention	Minutes
	Maintenance Type	Type of maintenance (AM = Autonomous Maintenance, PM = Programmed Maintenance)	Textual
Shift Availability	Theoretical Min Available	Shift duration $[h] * 60 \left[\frac{min}{h} \right]$	Minutes
	Current Availability	Theoretical minutes available in the shift – [festivities + holiday + (plant not available > 15 days) + minutes not worked + strike + out of work + canteen + union meeting + blackout]	Minutes
Availability Losses	Operating Time	Current availability – $\left[\text{unprogramed tool renewal stops} + \text{programmed tool renewal stops} + \text{scheduled } \frac{AM}{PM} \text{ stops} + \text{cleaning} \right.$	Minutes
		$\left. + \text{resolved } \frac{AM}{PM} \text{ faults} + \text{PM intervention wait} + \text{stop for automation exclusion} + \text{setup} + \text{reboot} \right]$	
Performance Losses	Net Operating Time	Operating Time – [microstop + emergency + intentional cycle stop + cycle time different from standard + lack of personnel + lack of direct materials + lack of power supply from upstream location + lack of absorption from downstream location]	Minutes
Quality Losses	Value-added Operating Time	Net Operating Time – [stop quality problems + slow down quality problems + scrap production minutes + rework production minutes]	Minutes
OEE Metrics	Availability	$\frac{\text{Operating Time}}{\text{Current availability}}_{Low} = [0—0.333]$	% value
		Medium = (0.333 – 0.666]	
		High = (0.666—1]	
	Performance	$\frac{\text{Net Operating Time}}{\text{Operating Time}}_{Low} = [0—0.333]$	% value
		Medium = (0.333 – 0.666]	
		High = (0.666—1]	
	Quality	$\frac{\text{Value – added operating time}}{\text{Net Operating Time}}_{Low} = [0—0.333]$	% value
		Medium = (0.333 – 0.666]	
		High = (0.666—1]	
	OEE	AVAILABILITY*PERFORMANCE*QUALITY _{Low} = [0—0.333]	% value
		Medium = (0.333 – 0.666]	
		High = (0.666—1]	

$$Supp(\alpha \rightarrow \beta) = \frac{\#(\alpha, \beta)}{\#T} \quad (2)$$

$$Conf(\alpha \rightarrow \beta) = \frac{Supp(\alpha \rightarrow \beta)}{Supp(\alpha)} \quad (3)$$

FP-growth algorithm supports ARs mining in the proposed application (Han et al., 2007): it has been chosen due to its efficiency in both time and memory requirements. Therefore, only the itemsets meeting the support and confidence constraints are mined. The main steps followed for ARM can be summarized as follows.

1. Minimum *Supp* and *Conf* thresholds definition;
2. FP-growth running and Frequent Itemsets generation (i.e., itemset having support higher than the minimum *Supp* determined at the first step);
3. Items belonging to the Frequent Itemsets are combined to identify relevant ARs, i.e., rules whose *Conf* is higher than minimum *Conf*.

In the proposed application, ARM can be used to relate the events occurring in the production process with the metrics measuring the performance of the process itself. In this way, it is possible to identify hidden relationships and have a complete view of unexpected patterns.

3.3.2. Network analysis

In order to provide a graphical representation of the ARM extracted in the previous step, network analysis is proposed. Indeed, networks are used to display relations occurring in social structures; they are usually represented by ordered pairs of vertices (V) and edges (E): $N\{V, E\}$ (Otte & Rousseau, 2002). Their implementation refers to analyzing exchanges and interactions, i.e., edges or arcs, among actors, i.e., the network's vertices. The V set is constituted by OEE values, its sub-factors, or operational variables; the E set, instead, connects the nodes to represent the ARs determined as described in Section 3.2: for this reason, the resulting graph is directed. In order to provide the representation with a ranking of rules preferences and a measure of their importance, a weight is attributed to each arc: the rule's confidence of each ARs is used. Specifically, if a rule in the form $a \rightarrow b$ exists, it will be represented in the graph in the following form: vertices a and b will be connected by an arc, weighted by the confidence of $a \rightarrow b$. The meaning of such a relationship is that when condition a occurs, also b will verify with a probability equal to the confidence value ($Conf(a \rightarrow b)$). The resulting graph is, as mentioned before, directed and weighted; it can be cyclical or acyclical depending on the association rules hidden in the input data and on the specific case study.

NA is used to visualize the most important relationships between low and high values of the OEE in connection with its sub-factors and then these with the operational variables in order to verify their reciprocal influence. Indeed, from such relationships, it is interesting to notice which actions should be prioritized to avoid low-efficiency levels or, in parallel, which events should be rewarded because of their positive impact on the OEE. In particular, the NA rules assessment process is based only on the confidence factor analysis in NA 1. Similarly, a more in-depth study should then include other variables in NA 2 so that interesting relationships will guide the decision-making process in layer 4.

In particular, the importance of the nodes and consequent actions and follow-ups devoted to improving their performance can be assessed through some metrics related to the network itself. Specifically, three performance indicators are used for the analysis (Baghipur et al., 2022):

a) Out-Degree (OD): Defined as the weighted sum of the outgoing edges (n) from node j , as represented by equation:

$$OD_j = \sum_{i=1}^n w_i \quad (4)$$

A high OD value indicates that node j strongly influences subsequent nodes. Indeed, the weight (w_i) of each edge represents the rule's confidence or probability of occurrence for the consequent i ($1 < i < n$). In the current framework, OD serves as a risk indicator.

b) In-Degree (ID): Defined as the weighted sum of the incoming edges (m) to node j , as shown in equation:

$$ID_j = \sum_{i=1}^m w_i \quad (5)$$

A high ID value signifies a strong influence of the incoming nodes i ($1 < i < m$) on node j .

c) Centrality Degree (CD): Calculated as the sum of a node's Out-Degree and In-Degree, as specified in equation:

$$CD_j = OD_j + ID_j \quad (6)$$

The CD reflects the overall importance of a node within the network by combining its outgoing and incoming influences.

3.4. Layer 4: Data-driven decision support system

This level refers to the TPM implementation model proposed in (Tian Xiang & Jeng Feng, 2021) with the distinction that the input data derive from objective analysis rather than expert opinion. The Data-driven decision support system extracts knowledge about operational variables impacting OEE (positively and negatively) and integrates a decision management system aligned with TPM principles. Therefore, this layer provides actionable insights to production managers for corrective and improvement actions, supporting the TPM philosophy. Key functionalities include:

1. *Knowledge extraction*, facilitating the establishment of a TPM committee for data review, Kaizen events classification, and TPM roadmap development.
2. *Decision management system*, defining new TPM procedures and providing recommendations for managing loss events or rewarding best-performing events, focusing on the most impacting OEE factors among availability, performance, and quality.
3. *Implementation*, enabling the adoption of best TPM practices, supported by data-driven continuous monitoring and a warning system within an industrial case study.

4. Embedding the data-driven TPM framework in an industrial case study

The proposed framework has been applied in an automotive company, targeting one fully automated assembly line comprising twelve workstations. The analysis spans a reference period of three years, with the outcomes detailed in the current section.

4.1. Implementation of layers 1 and 2: Data acquisition and preprocessing

As an initial step, Fig. 2 shows how operational and maintenance data were integrated into a consolidated Data Warehouse to provide a complete view of production performance.

DATAFRAME 1 (top) and DATAFRAME 2 (middle) each contributed distinct types of data, which were combined to create a unified view (bottom) for analysis:

- DATAFRAME 1: Operational data were sourced from the MES (a cloud-based platform), providing insights into production availability, performance losses, and quality losses. Key metrics:
 - *Theoretical Min Available*: The total shift duration, typically 480 min for an eight-hour shift, was represented.
 - *Current Availability*: Theoretical minutes were adjusted to account for breaks and other interruptions.

- *Operating Time*: The actual time the plant operated at standard production rates was calculated by subtracting downtime due to resolved maintenance issues and setup time from current availability.
- *Net Operating Time*: Operating Time was further adjusted for performance losses, such as micro-stops and cycle time deviations.
- *Value-added Operating Time*: Quality-related losses, such as waste and rework, were subtracted from Net Operating Time.

This structured data enabled effective monitoring and calculation of production performance, providing insights into time utilization, performance deviations, and quality impacts of each work shift.

- **DATAFRAME 2**: Maintenance-related data were sourced from the CMMS (a server-based platform). Each row recorded maintenance activities linked to specific workstations. Key information includes:
 - *Workstation and Sub-Workstation*: The affected equipment or machine was identified.
 - *Failure/Downtime and Description*: A fault code and description were provided, enabling categorization and analysis of recurring issues.
 - *Min/Pcs*: The duration of each maintenance intervention was recorded in minutes, showing time spent on repairs.
 - *Type of Maintenance*: Maintenance types were classified as either Autonomous Maintenance (AM) or Programmed Maintenance (PM).

Essential details on downtime events, fault descriptions, and maintenance types were captured, supporting a detailed analysis of maintenance efficiency and reliability.

Aiming to aggregate the data frames, a merging key—consisting of Date, Week, Shift, and Team Leader—was used to uniquely identify each record (row), allowing accurate integration without duplication. The merged view (bottom of Fig. 2) combined entries with matching keys, resulting in a comprehensive Data Warehouse with 1,222 records for the first six months, periodically updated.

After integration, the data was cleaned and standardized. Inconsistent rows were removed, particularly those with negative values due to extended downtimes unrelated to production. The final Data Warehouse was then validated with checks for missing information such as maintenance duration and team leader names. Failure types and downtimes were also standardized (see Table 2).

The final data warehouse was created with OEE metrics, providing structured insights into production performance.

A detailed description of input variables is provided in Table 3, organized as follows:

- *Shift details*: Basic information about each shift (date, week, shift number, team leader).
- *Failure data*: Information on workstation faults, including codes, descriptions, durations, and maintenance type.
- *Shift availability*: Theoretical and actual minutes available per shift, accounting for non-operational times.
- *Availability losses*: Operating time adjusted for downtime.
- *Performance losses*: Net operating time adjusted for performance delays like micro-stops.
- *Quality losses*: Value-added operating time, excluding quality-related losses like scrap and rework.
- *OEE metrics*: Metrics for availability, performance, quality, and the composite OEE score, each categorized as low, medium, or high.

This structured approach allowed for comprehensive tracking and analysis of both production performance and maintenance effectiveness.

4.2. Implementation of layer 3: Data analytics

This layer consists of different steps. The following sub-sections are dedicated to explaining these steps for the case study under

consideration.

4.2.1. Identification of most relevant factors

The first analysis consists of identifying the factors that are most relevant in lowering or raising OEE values. The objective of this analysis is to save time for company managers and focus their attention first on the factors that most affect OEE.

Different Factors can be used as predictors among those collected in Layer 2. In this case study, the first step has been focused on the three main OEE factors, i.e., PERFORMANCE, AVAILABILITY and QUALITY.

The number of association rules extracted strictly depends on the setting of minimum support and confidence thresholds since such level limits the space of the solutions displayed. In other words, the higher the minimum threshold is, the lower is the number of ARs. Since there is no specific rule of thumb indicating the ideal value, this is related to the specific case study, considering the risk of losing information – by setting a high threshold – or the possibility of slowing down the algorithm by setting a low value. In the proposed application, since the cardinality of the dataset and the number of variables selected allowed it, both minimum support and confidence were set to 0. Table 4 details an excerpt of the 189 ARs extracted to relate OEE factors and ranges. Rules are ranked by descending confidence since the aim is to report the most impactful factors over the OEE value.

Observing the ARs, it can be noticed that availability is the most impactful on the OEE: the first rule, indeed, has a confidence of 100 %, indicating that whenever availability is low – i.e., lower than 33 % – then the OEE will be low as well. Similarly, when the availability is in the highest range, OEE is high in 84 % of cases, indicating a high relevance even in terms of good performance. On the other hand, the PERFORMANCE factor is the less impacting on OEE value: indeed, in rules *PERFORMANCE = High, AVAILABILITY = Low, QUALITY = Medium* → *OEE = Low* and *PERFORMANCE = High, QUALITY = High, AVAILABILITY = Low* → *OEE = Low*, it can be noticed that PERFORMANCE factor always belongs to the highest range, even when OEE is at its lowest. In parallel, AVAILABILITY factor always remains in the lowest range, highlighting the strongest influence of the latter factor in 100 % of cases since the confidence is 1.00 for both rules.

Hence, it can be stated that availability is a variable of interest both in terms of best events identification and loss events identification.

In order to get a complete overview of this analysis, NA is recommended by the proposed framework. Fig. 3 shows the network representation of 53 two-itemset rules relating the factors among them and to the values of the OEE. The network nodes represent parameters used in ARM analysis. The bigger the node, the higher the corresponding out-degree, i.e., the node appears as a consequent of several rules. In addition, the thickness of the edge represents the rule's confidence, hence it represents the strength of the rule itself. In order to coherently interpret the results displayed by the network, the following aspects should be taken into consideration:

1. The dimension: as previously stated, high OD indicates a high influence of the node on consequent nodes.
2. The thickness of the connection edges, which represents the strength of the rule displayed.
3. The practical relevance of the relationship, in order to identify appropriate corrective actions.

Observing Fig. 3, nodes *OEE = Low*, *QUALITY = High*, and *PERFORMANCE = High* capture the attention since they are the biggest and are highlighted in a different color, i.e., as previously explained, they have the highest OD. Considering the fact that the former node provides a negative impact on the performance of the production system, it is the ones that need prioritization. The connection between *AVAILABILITY = Low* and *OEE = Low* is specifically worthy of mention since among the OEE factors, availability is the only one influencing the OEE in a negative manner and with substantial confidence: rule *AVAILABILITY = Low*

Table 4

Excerpt of the results of ARs mining.

Premises	Conclusions	Support	Confidence
PERFORMANCE = High, QUALITY = High, AVAILABILITY = Medium	OEE = Low	0.04	0.15
PERFORMANCE = High, QUALITY = High, AVAILABILITY = High	OEE = Medium	0.22	0.32
PERFORMANCE = High, QUALITY = High, AVAILABILITY = High	OEE = High	0.48	0.68
PERFORMANCE = High, QUALITY = High, AVAILABILITY = Medium	OEE = Medium	0.20	0.85
PERFORMANCE = High, QUALITY = High, AVAILABILITY = Low	OEE = Low	0.03	1.00
PERFORMANCE = High, AVAILABILITY = Low, QUALITY = Medium	OEE = Low	0.01	1.00
....			

→ OEE = Low, indeed, has a confidence of 100 %, i.e., whenever the availability factor belongs to the lowest range, the OEE is below 33 % as well. The confidence of the opposite rule is still high, but it stands at 48.5 %. Hence, the OEE factors are classified in Table 5:

This classification highlighted AVAILABILITY as the most critical factor influencing both the lowest and highest OEE ranges, followed by QUALITY as the second priority. Lastly, PERFORMANCE consistently ranked in the highest range, as indicated by the previously analyzed ARs.

The subsequent step focuses on identifying operational variables associated with the most critical and best OEE factors. However, according to the previous results, only the AVAILABILITY root causes analysis is presented in this paper, identifying both negative and positive events. Layer 4 then proposes improvement actions and reward strategies to address these findings.

4.2.2. Availability root causes analysis

The company analysed in this paper recorded in the CMMS all production variables connected to AVAILABILITY, QUALITY and PERFORMANCE factors. In order to analyse the relationships between different operational variables related to Availability, three variables including general information – as disclosed in Table 3 (Shift, Team Leader and Problem Code) – and, for the statistics group, OEE, quality, performance and availability are taken into consideration in order not to lose valuable information. Specifically, 8247 association rules were extracted with Rapid Miner software, and a further selection of the most relevant ones was provided by following in order two constraints:

$$ARs = ARs \cap (AVAILABILITY = Low) \quad (7)$$

$$ARs = \{Shift \rightarrow PC, TeamLeader \rightarrow PC, PC \leftrightarrow PC\}_{PC=\{PC_i \in PC \mid PC_i \rightarrow OEE=Low\}} \quad (8)$$

The application of constraint (7) leads to the selection of 732 rules out of 8247 since only the ones referring to AVAILABILITY = Low are currently of interest. In an attempt to develop the NA, Fig. 4 was thus created weighting the connection by confidence. In all, 45 PCs out of 53 have been selected, as well as three work shifts and three team leaders.

Anyways, the network represented in Fig. 4 still does not allow easy interpretation, due to the high number of nodes and connections. Based only on such representation, would not ensure the extraction of meaningful insights. Hence, according to (8), the following filtering and selection is carried out:

1. Only the eight PCs causing low OEE values were considered, namely, PC10, PC18, PC19, PC34, PC45, PC46, PC47, and PC51.
2. All the relationships between the above PCs, Shifts, and Team Leaders were thus considered. Those with team leader and shift as a consequence of the rule and those between team leaders and shifts as not of interest to experts since they do not represent modifiable events.

As a result, 48 relevant rules were identified, and a matrix was constructed based on the confidence factors of each rule. These rules offered insights into the relationships between PCs, shifts, and team leaders under conditions where availability falls within the lowest range, leading to a corresponding reduction in OEE (Table 6).

The numbers in the matrix represent the weight of the edges between the nodes in each row and column, highlighting the strength of their relationships. These ARs provide valuable insights for company managers, particularly in identifying inefficiencies in maintenance operations. For instance, relationships between problem codes and shifts can reveal areas requiring intervention, such as the need for increased maintenance efforts or pre-shift inspections. As an example, Problem Code 45 occurs most frequently during the Shift 3 (77 % of cases, compared to 38 % for PC47 and 23 % for PC46). Notably, this issue is also strongly linked to a specific team leader, Leader Z, in 77 % of cases. This correlation suggests that the team leader's operations should be thoroughly reviewed, with potential additional training or corrective actions considered. In contrast, Problem Code 51 has a low impact on Shift 1, with a confidence level of 24 %. However, Problem Code 18 shows a strong dependency, with an 80 % probability of co-occurring with PC45, indicating a higher level of criticality.

To further enhance the visualization of these relationships, a network could be developed to display the data from Table 6, providing a clearer understanding of these interconnections. The nodes highlighted in Fig. 5 represent the ones having the higher ODs, hence the ones that should draw specific attention. More in details, PC18 and PC34 are characterized by higher ODs among problem codes; hence, they should be prioritized. Similarly, concerning the team leaders and shifts, Team Leaders X and Z and Shift 2 are the ones to be addressed with higher priority.

If the network representation provides an immediate snapshot of potential criticalities, a deeper classification of the operational variables and relationships displayed in Fig. 5 supports the analysis of the inter-node influence: company managers should be provided with additional information compared to Table 6 such as ODs, IDs and CDs values representing the strength of influence of each operational variable, namely each node, and calculated as in Section 3.3. In addition, the node's classification is also deepened in Table 7 based on the following:

- *Source node*: variable that affects other nodes but is influenced by none.
- *Target node*: variable that is affected by other nodes but does not affect (i.e., the opposite of the Source node).
- *Regular nodes*: variables that both affect and are affected.

As shown in Table 7, no production variable as classified as a Target node, which is a positive outcome. Target nodes are the most challenging to manage due to their complete dependence on other variables. Instead, Table 7 reveals a high concentration of Regular nodes (57.14 %) and a moderate presence of Source nodes (42.86 %). Source nodes are easier to handle due to their independence from others, such as "Team Leader" OR "Shift" → PROBLEM CODES. Additionally, Table 7 suggests that each problem code in the case study under investigation either influence or be influenced by other problem codes, shifts, and/or team leader variables. The correlations between different problem codes highlight the necessity of implementing the preventive maintenance pillar of TPM. This approach is essential because critical problem codes cannot be analyzed in isolation, requiring preventive measures and correlation analysis to effectively address interdependencies.

To further understand these relationships, a graphical representation was created to compare IDs and ODs values of each node, as in Fig. 6.

First, it is important to note that variables with ID ≈ 0 , such as shifts and team leaders, are unaffected by other variables. Conversely, variables with OD ≈ 0 exert the minimal influence. Referring to the graph in Fig. 6, Problem code 18 emerges as the variable with the highest OD value, closely followed by PC34 and Team Leader Z, respectively. These

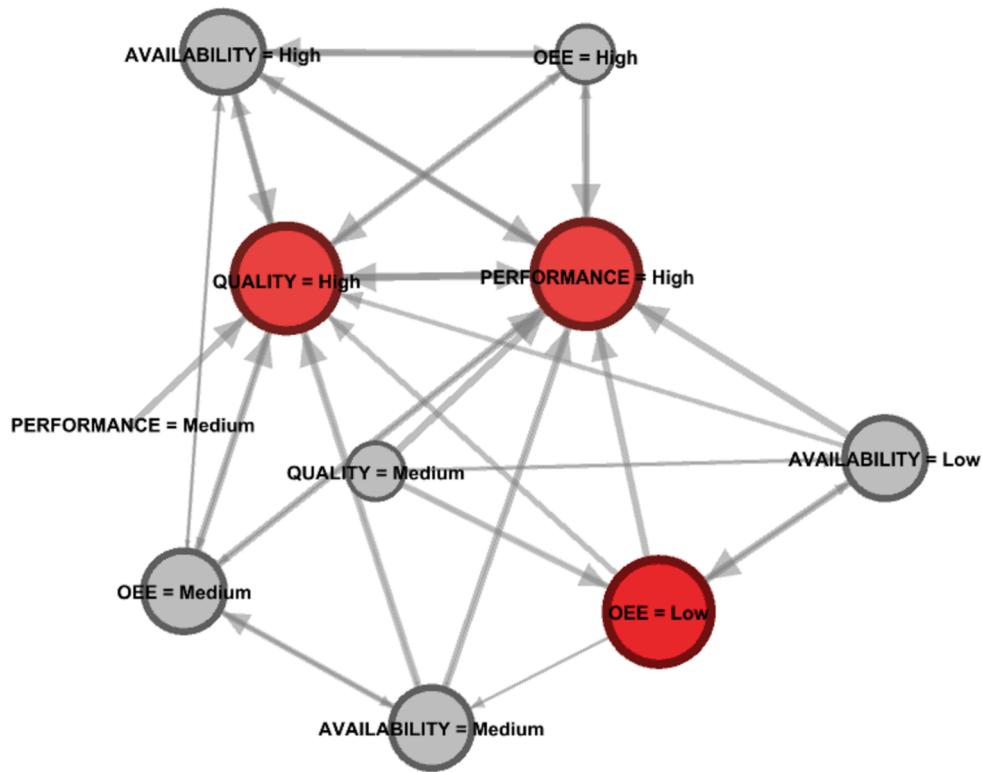


Fig. 3. Network representation of the ARs relating the factors among them and to OEE values.

variables are identified as critical variables affecting availability, and their prioritization in the roadmap is essential for improving availability. In contrast, PC45 has the lowest OD value, indicating minimal influence. However, when considering ID values, PC 45 has the highest input degree, demonstrating its sensibilities to being affected by the rest of variables. This characteristic makes improving availability by targeting these variables particularly challenging. Additionally, PC 45 stands out due to its high CD value, signifying its significant importance in cause-effect interactions. In such a case, as for the other variables, the data-driven decision support system supports managers in selecting the most effective improvement strategies aligned with TPM pillars to address these challenges.

4.3. Implementation of layer 4: Data-driven decision support system

Layer 4 focuses on knowledge extraction and decision-making to classify loss and best events and enhance OEE following the TPM philosophy. Specifically, it emphasizes developing new procedures, recommendations and reward strategies to address availability issues, with some focus on quality challenges, as suggested by Table 5.

Knowledge extraction

To plan the roadmap and guide decision-making, it is essential to analyze the results from Layer 3. This involves aggregating key information, such as:

- The influence of each variable, reflected in Fig. 6.

Table 5

Ranking of OEE factors from most impacting to least.

OEE factors classification
1° Availability
2° Quality
3° Performance

- The root causes of the most critical target nodes, highlighted in Fig. 5.

Using this aggregated data, the 14 operational variables were classified into four groups (G1, G2, G3, G4) based on their impact, as shown in Fig. 7, wherein 2,4 and 8 are the highest OD and ID values, respectively.

G1 refers to operational variables with high OD (indicating a strong capacity to improve availability) and low ID (indicating ease of management and control due to minimal influence from other variables). *Helium room 1* (PC 18), *Refilling pump body* (PC 34) and *Team Leader Z* fall into this category, making them ideal targets for improvement efforts.

Next, G2 generally represents the most critical variables from a management perspective, characterized by both high OD and ID values, as well as high centrality degrees. Fortunately, in the current case study, no operational variables fall into this category.

As opposed to G2, G3 refers to neutral variables with both low OD and ID values, indicating that they neither significantly impact availability nor are heavily influenced by other variables. As such, they are easy to manage and control. This group includes *Fixing Koenig Cap* (PC 10), *Helium room 2* (PC 19), *Robot 1 replacement* (PC 46), *Robot 2 replacement* (PC 47), *Welding Inlet Fitting* (PC 51), *shift 1*, *shift 2*, *shift 3*, *team leader X*, and *team leader W*.

Lastly, G4 consists solely to *Replacing pump body* (PC 45), the only operational variable with a low OD, (indicating a minimal impact on availability) but a high ID (indicating complex management). Despite having the highest centrality degree due to its high ID, PC 45 requires less attention as it does not significantly influence availability.

Knowledge extraction enables managers and experts to design a TPM roadmap (Fig. 8) to plan and prioritize improvement actions for TPM implementation. This roadmap follows a three-year implementation strategy aimed at achieving TPM objectives, enhancing OEE, and prioritizing actions based on the impact of operational variables. The focus is primarily on improving availability, while incorporating quality

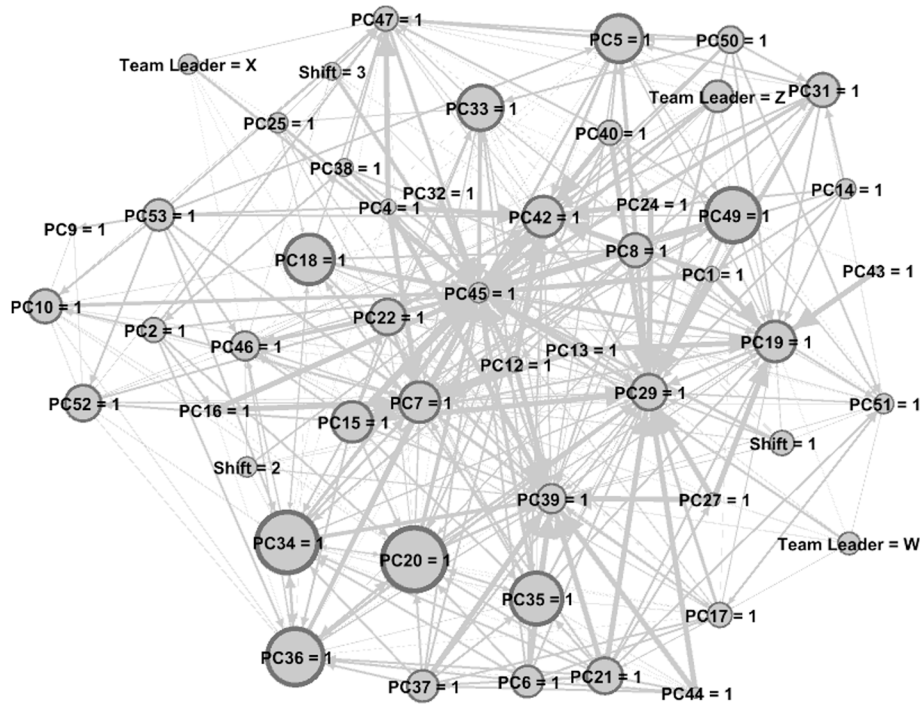


Fig. 4. Network to display the relationships among production variables with low availability.

as a secondary yet significant factor. Specifically, the roadmap consists of:

- Yearly targets: Each year is associated with specific OEE improvement goals (+5%, +10 %, and + 15 %), providing clear, measurable milestones.
- Priority-based approach: Actions are structured according to three levels of event's prioritization:
 - Priority 1: Focus on the best events (including a reward system) and G1 variables significantly impacting availability.
 - Priority 2: Address G4 variables, which have a low impact on availability but require complex management.
 - Priority 3: Include G3 variables, which are the most neutral in their impact and easiest to manage.
- Layered structure: The roadmap also shows Kaizen actions for each event, the typology of the action, TPM pillar achieved, and monitoring and warning systems to ensure a comprehensive approach.
- Progressive implementation: Each year builds on the previous year's achievements, starting with the construction of a TPM committee in Year 1 to oversee and guide the process.

- Visual design: The roadmap helps manager and operators in visualizing roles, timelines, priorities, actions, and objectives for continuous improvement.

As shown in the long-term TPM roadmap, the Kaizen team consisted of a diverse group of 4 team leaders, 5 production line operators, 3 maintainers, 1 production manager, and 2 administrative officers. By the end of the TPM process, 24 improvement actions were implemented, addressing 12 main loss events and 2 best events. Overall, the cumulative OEE values closely aligned with the targets set in the roadmap.

Decision management insights.

PC18, PC 34, and Team Leader Z were classified in the G1 class, making them critical priorities for addressing availability and OEE losses. A deeper analysis (Table 6 and Fig. 5) identified PC45 as the most critical problem code for Team Leader Z, with a confidence factor of 77 %, followed by PC19, while PC46 and PC47 were tied. This highlighted the need for Team Leader Z to receive additional training, particularly in managing PC45 faults.

In contrast, Team Leaders X and W, along with all shifts, were categorized under the G3 class, representing minimal criticalities. Further investigation (Fig. 6) revealed Team Leader W as the best

Table 6
Rule confidence matrix.

Production variables	PC10 = 1	PC18 = 1	PC19 = 1	PC34 = 1	PC45 = 1	PC46 = 1	PC47 = 1	PC51 = 1
PC10 = 1		0.25			0.75		0.25	
PC18 = 1	0.20		0.40	0.40	0.80		0.40	0.20
PC19 = 1		0.20			0.60		0.20	0.20
PC34 = 1		0.40			0.60	0.40	0.20	
PC45 = 1			0.21				0.24	
PC46 = 1				0.22	0.44			
PC47 = 1		0.20	0.20		0.70			
PC51 = 1			0.29		0.29			
Shift = 1			0.29		0.65			0.24
Shift = 2			0.24	0.24	0.47	0.29		
Shift = 3					0.77	0.23	0.38	
Team Leader = W			0.21		0.54			0.25
Team Leader = X		0.20			0.60	0.20	0.30	
Team Leader = Z			0.31		0.77	0.23	0.23	

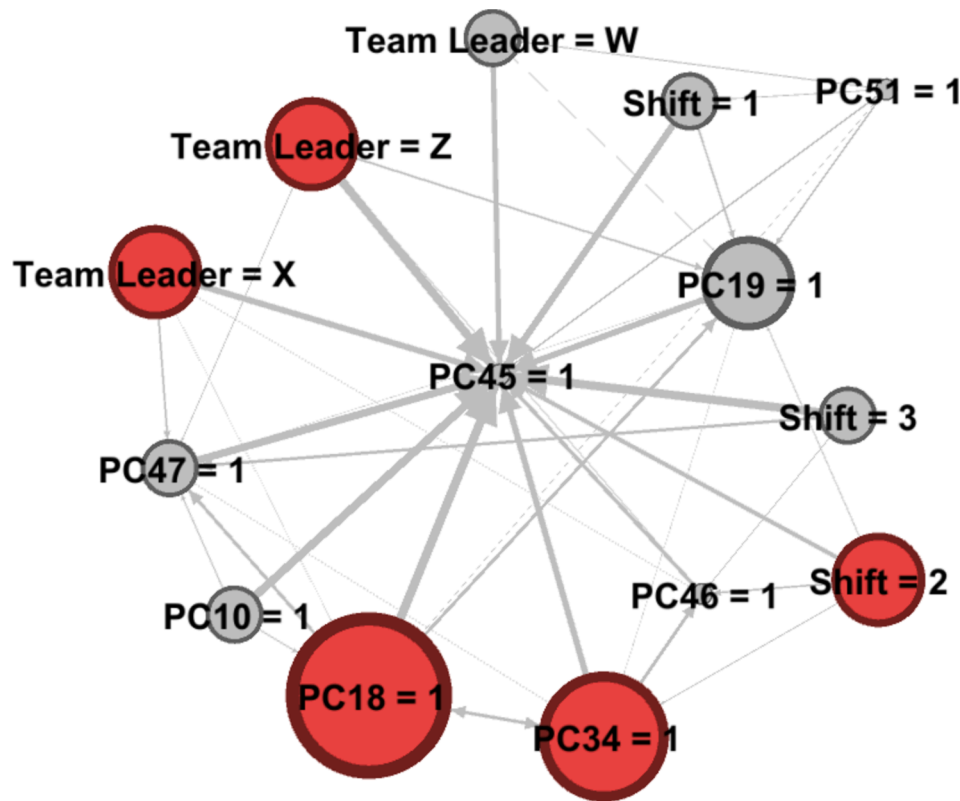


Fig. 5. Network to display the relationships reported in Table 6.

performer due to his lowest centrality degree, reflecting minimal cause-effect interactions. A reward system was thus recommended for W to acknowledge his performance.

For the shifts, the first shift was identified as the best-performing, while the third shift was the most problematic. Based on Table 6 and Fig. 5, recommendations for Shift 3 included improving PC45 management through pre-shift inspection and intensifying preventive maintenance for PC46 and PC47. PC45, belonging to G4, was the only event in the second-priority category and required focused attention. For Shift 2, fewer interventions were needed, as it was the second critical slot in a working day. On the other hand, monthly sessions were planned to improve operators' TPM mindset during Shift 1 (the best-performing slot), along with training for the production manager to standardize procedures and enhance plant awareness.

These findings validated the new procedures and recommendations, emphasizing collective responsibility over individualism, encouraging all employees to contribute to and improve TPM initiatives. Specifically,

Table 7
Classification of the nodes.

Nodes	IDs	ODs	CDs	Class
PC10 = 1	0.20	1.25	1.45	REGULAR
PC18 = 1	1.25	2.40	3.65	REGULAR
PC19 = 1	2.34	1.20	3.54	REGULAR
PC34 = 1	0.86	1.80	2.66	REGULAR
PC45 = 1	7.98	0.45	8.43	REGULAR
PC46 = 1	1.36	0.67	2.02	REGULAR
PC47 = 1	2.21	1.10	3.31	REGULAR
PC51 = 1	0.89	0.57	1.46	REGULAR
Shift = 1	0.00	1.18	1.18	SOURCE
Shift = 2	0.00	1.24	1.24	SOURCE
Shift = 3	0.00	1.38	1.38	SOURCE
Team Leader = W	0.00	1.00	1.00	SOURCE
Team Leader = X	0.00	1.30	1.30	SOURCE
Team Leader = Z	0.00	1.54	1.54	SOURCE

the TPM actions were reviewed every six months to evaluate the outcomes of previous improvements and reward activities. Real-time data were compared with pre-TPM data to monitor procedural accuracy and measure progress. The OEE was reassessed annually to validate the implementation of TPM pillars.

5. Results

The results of the three-year TPM framework implementation highlighted year-over-year progress in equipment effectiveness, consistent with the targeted goals, as detailed in Table 8:

- Baseline KPIs (Year 0): Initial benchmarks for availability, performance, quality, and OEE.
- Calculations for Availability, Performance, Quality, and OEE.
- Δ OEE: Annual improvement in OEE compared to the previous year.
- Cumulative Δ OEE actual vs target: Tracking actual vs. targeted OEE gains each year.

Key findings:

- Availability increased the most (+4%) in the first year, driven by targeted actions on high-priority areas. Performance remained stable, with slight gains in Years 1 and 2 due to 5S-based workstation reorganization and inspections. Quality showed its greatest improvement in Year 3, confirming the success of the Quality Maintenance Pillar.
- Overall, the implementation achieved a 14 % OEE improvement, closely aligning with the target: the largest improvement (+6%) occurred in Year 1, driven by Kaizen initiatives targeting high-priority production events. Gains of + 4 % and + 3 % in subsequent years resulted from lower-priority interventions, with category two requiring only one specific action. While slightly below target,

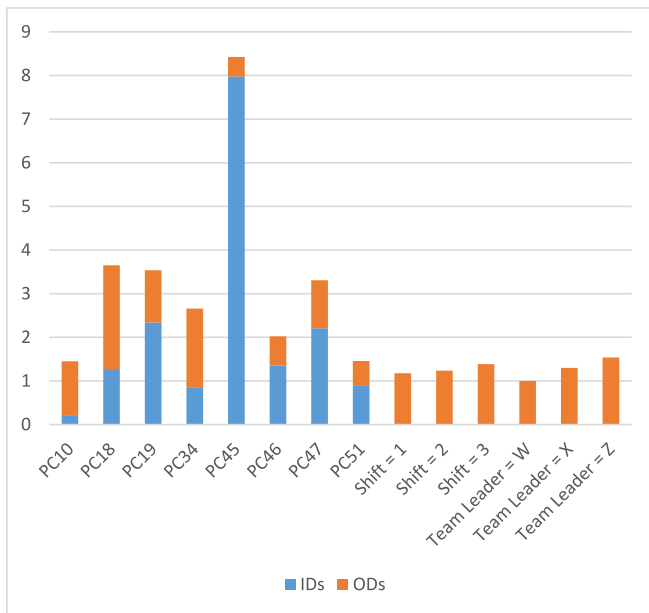


Fig. 6. Comparison of Input and Output Degrees of the operational variables most affecting availability.

the cumulative results demonstrated the sustained impact of prioritized actions over time.

6. Discussion

The present study focused on the progressive implementation of the TPM policy using data-driven techniques to identify the production events most significantly impacting OEE loss factors and best-performing factors. Specifically, the current paper addressed the following research question: How can ARM and NA be leveraged to

develop a long-term roadmap for TPM implementation? The study provides both practical implications for business managers and theoretical contributions for operations management researchers, detailed below.

6.1. Practical implications

Validated through a longitudinal industrial case study, the proposed framework demonstrated practical applicability and effectiveness in enhancing operational performance while aligning with I4.0. The TPM program was tailored to the specific needs and conditions of the company. Its application, which replaced individualism with teamwork, allowed for a progressive implementation of TPM in a context where the regular maintenance program was initially scheduled daily based on operator experience and machinery characteristics. Unlike previous studies that partially addressed TPM pillars (Correia Pinto et al., 2020), this framework achieved seven out of eight TPM pillars, omitting only administrative TPM due to the company's lack of interest in logistics losses. Key achievements for each pillar include:

- **Early Management:** implementing 24 minor improvements rather than a few considerable improvements to address technical and human failure causes, maximizing efficiency and service effectiveness to enhance production processes.
- **Autonomous Maintenance:** engaging operators in routine maintenance tasks (e.g., PC18 and PC19 in helium rooms) and implementing 5S principles to ensure clear standards for cleaning, inspection, and lubrication.
- **Preventive Maintenance:** establishing a plan based on MTTR, MTBF, and expert input to reduce unexpected downtime (e.g., PC10 and PC51), indirectly improving quality performance and being supported by autonomous maintenance by acting on the causes of such failure events.
- **Quality Maintenance:** detecting and preventing defects and waste for PC46 and PC47, automating robot performance checks with visual sensor signals, keeping both Robots 1 and 2 more performant. This

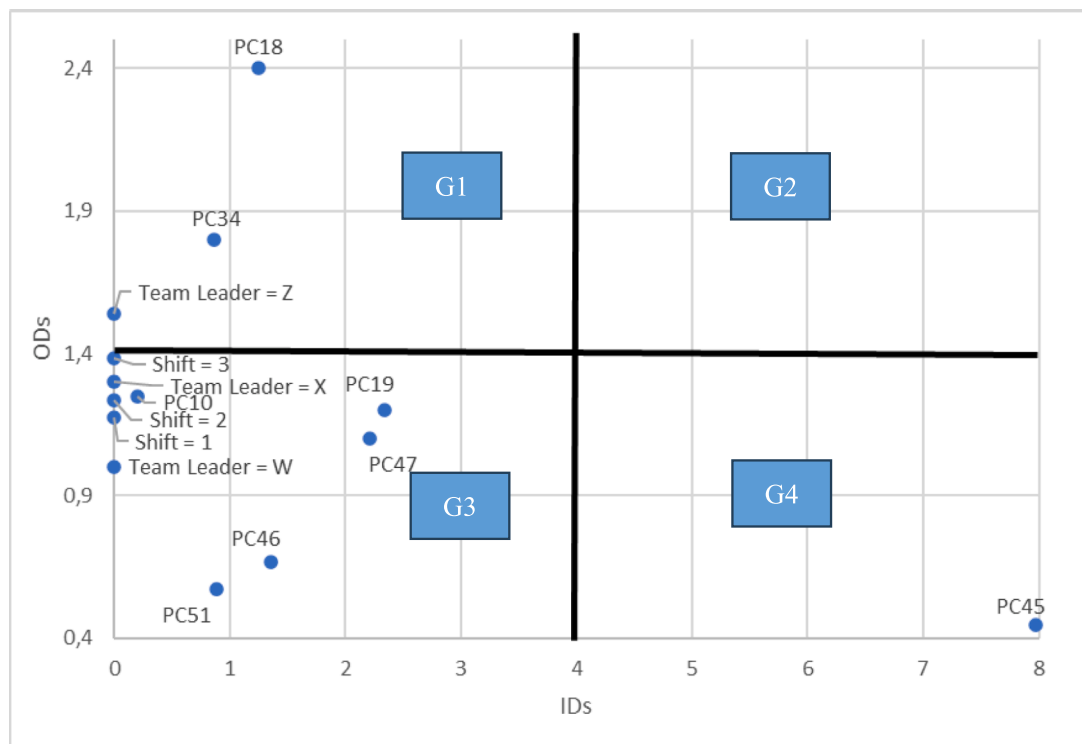


Fig. 7. Findings classification.

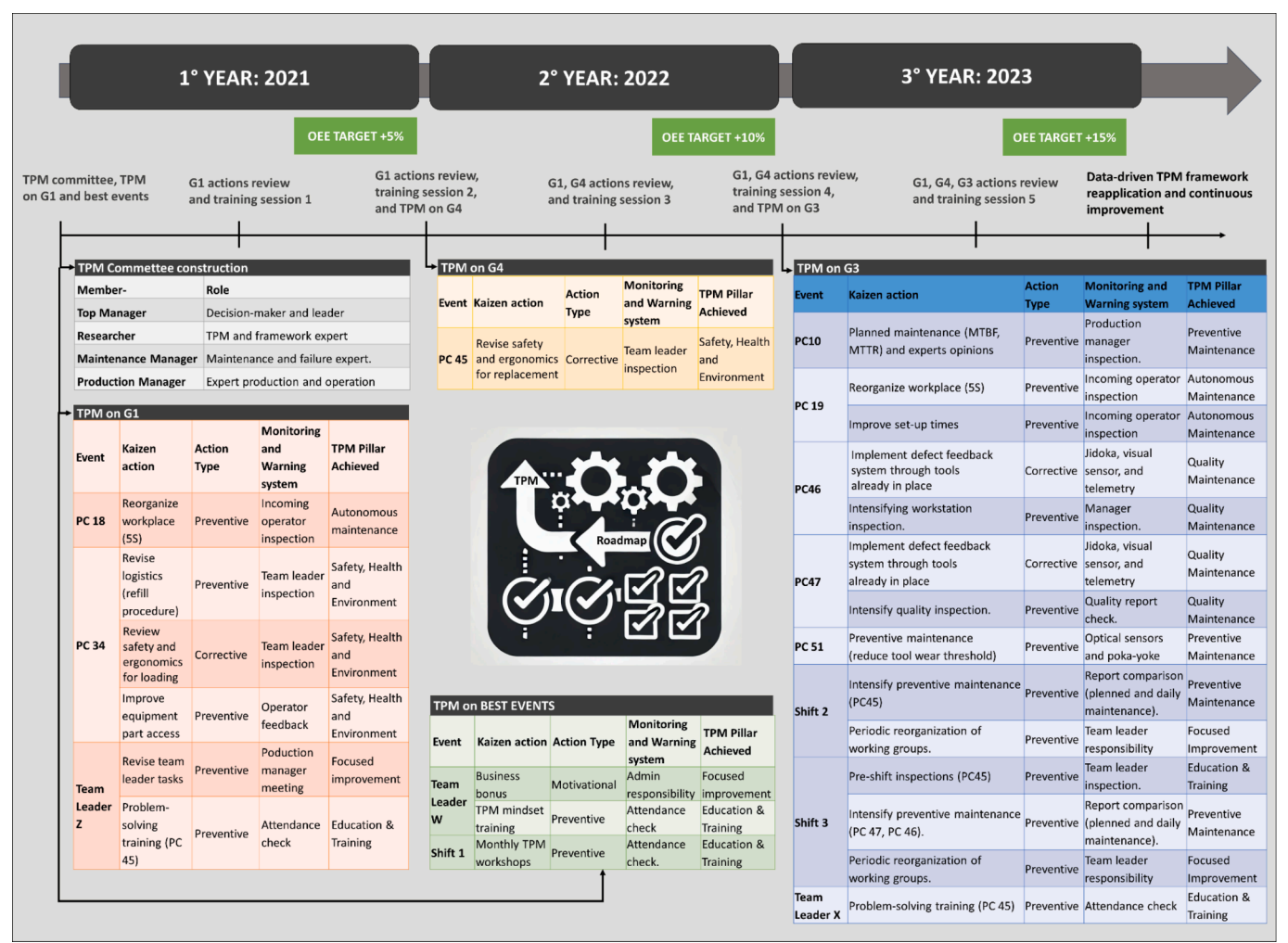


Fig. 8. Long-term TPM roadmap.

- reduced manual monitoring and shifted maintenance report reviews from daily to monthly.
- **Education & Training:** managers, team leaders, and operators gained theoretical knowledge about the data-driven TPM framework, as well as TPM pillars and objectives, and practical skills for managing minor breakdowns, fostering a sense of ownership over equipment. This allowed team leaders and operators to ask for maintenance interventions just in case of critical issues. For instance, Team Leaders Z and X were trained in proactive problem-solving for PC45, while a reward system incentivized continuous improvement for Team Leader W.
 - **Safety, Health and Environment:** addressed ergonomic and safety concerns related to PC34 and PC45 by optimizing the pump body loading process at the first workstation and the replacement procedures along the entire assembly line. Specifically, unsafe conditions, such as obstacles, were eliminated, and ergonomic aspects were enhanced through the implementation of a suitable structure to facilitate the task. These improvements resulted in a 7 % reduction in loading time, as measured during the second actions review.
 - **Focused Improvements (Kaizen):** initiated from day one, this pillar involved identifying and addressing critical production events, rewarding best performers (e.g., the business bonus for Team Leader W), and conducting biannual action reviews to measure OEE and refine maintenance strategies, as well as ongoing training of the Kaizen team.

Examining the causal relationships between variables can provide

valuable insights for industry managers and future research. This analysis allowed managers to focus their efforts on better managing and controlling variables that positively or negatively influenced others.

Although this approach was validated in a specific case study, its application was not limited to this industrial context. Since TPM and OEE are widely adopted across both manufacturing and service environments (e.g., Barriga et al., 2024; Muchiri & Pintelon, 2008), the framework was adaptable to various industries. The main prerequisite for implementation was the company's ability or willingness to collect machinery or process status data, including operating time, losses, maintenance reports, and product quality. This framework also proved suitable for SMEs, where using open-source platforms mitigated significant monetary investments, often a critical limitation for smaller companies. However, the lack of expertise among personnel during the initial implementation could be considered as a barrier. Overcoming this barrier required selecting relevant data sources and creating a robust dataset. Conversely, the framework's ease of interpretation, supported by association rules, network analysis, and graphical classification of production events, simplified its implementation even for non-experts, streamlining the creation of a TPM roadmap.

To conclude, a further aspect that should be considered to contextualize the proposed framework is data availability and quality. As a data-driven methodology, its effectiveness heavily depended on the consistency and reliability of the data analyzed, regardless of whether it was collected manually or automatically. Inaccurate or redundant data undermined the framework's ability to improve overall system performance (Crespo Márquez et al., 2019).

6.2. Theoretical contribution

The data-driven TPM framework proposed in this study addresses the growing need for innovative big data analytics solutions to ensure a continuous flow in implementing Lean 4.0 principles and TPM pillars (Lucantoni et al., 2022; Valamede & Akkari, 2020). By leveraging LA, the proposed framework overcomes the limitations of both traditional tools, such as Pareto Charts for loss assessment and Ishikawa diagrams for cause identification, and blended approaches that combine advanced DM techniques with traditional methods. Instead, the proposed framework integrates ARM to uncover hidden relationships between failure events and OEE values, and NA to create relationship patterns that support operators in decision-making processes. The framework incorporates a data-driven decision support system aligned with Jidoka Toyota’s philosophy, which views downtimes as opportunities to stop production and resolve issues, ensuring uninterrupted manufacturing (Kitazuka & Moretti, 2012). The theoretical contributions of the proposed framework can be summarized as follows:

- Development of advanced data analytics tools to support TPM teams in decision-making, optimizing maintenance schedules, identifying root causes, and implementing targeted improvement strategies.
- Introduction of a novel decision-making methodology for OEE improvement, accounting for all performance, quality, and availability factors.
- Conducting detailed resource efficiency analyses to evaluate operational variables influencing both the best and worst OEE factors.
- Providing a roadmap for TPM procedures and recommendations, enabling prioritization of operational variables for long-term OEE improvement.
- Facilitating the progressive achievement of TPM pillars through continuous monitoring and an integrated warning system.
- Preliminary approach towards Industry 5.0 where and human expertise are seamlessly integrated (Jan et al., 2023).

One of the challenges addressed by this framework was the integration of complex datasets from diverse sources with varying speeds and quality. To support the manufacturing company’s decision-making process, an N-dimensional analysis of performance data was required, utilizing datasets and streams gathered from wired and wireless sensors and meters. While the proposed dataset included data archival, business managers require help shifting through the vast volumes of data to find the necessary information. By introducing a framework capable of integrating and analyzing complex datasets through a big data analytics approach, this work addresses this gap. The massive, varied, and regularly generated data collections posed significant management challenges due to their complexity and volatility, as well as increasing volume, which makes processing and analysis highly demanding. Conventional analytical tools were unsuitable for such data because:

- The high number of potential predictors complicated typical statistical analysis.
- The independence hypothesis is commonly adopted in a parametric analysis, but the relationship between the independent variables may present problems in this study.
- The distribution of the data showed non-homogeneity and non-linearity, further complicating traditional approaches.

This study highlighted ARs as an effective alternative to conventional parametric techniques often applied to TPM problems. ARs proved particularly adept at defining “all” potential links among factors and causes affecting OEE factors. Furthermore, NA offered complementary capabilities by:

- a. Mapping comprehensive connections among factors and causes affecting the OEE. This approach identifies nodes or vertices joined in pairs by lines or edges.
- b. Identifying local and global patterns and locating influential entities.

Eventually, to evaluate the proposed methodology’s effectiveness (ARM + NA) as a foundation for TPM implementation, a comparative analysis was conducted (Table 9) against the only existing data-driven TPM methodology that employs ARM in the literature. This alternative methodology uses ARM and FMEA to analyze concurrent failure events occurring concurrently (Antomarioni et al., 2023).

The proposed methodology introduces significant innovations compared to the existing approach:

- Scope of analysis: unlike the FMEA-based methodology that focuses on failure modes, the proposed approach expands the scope to include pattern recognition across all operational variables, emphasizing their impact on critical OEE factors.
- Methodology: the proposed methodology goes beyond the existing approach by integrating a dual application of ARM and NA, enabling the analysis of all OEE factors (availability, performance, and quality) rather than just availability. This comprehensive approach provides deeper insights into both positive and negative influences on system performance. Advanced graph theory metrics, such as Out-degree, In-degree, and Centrality Degree, further enhance the traditional support and confidence analysis.
- Improvement strategy and actions: The new approach extends beyond the existing methodology, which focuses exclusively on “urgent and important” events categorized using the Eisenhower matrix. Instead, the proposed methodology introduces a structured TPM roadmap addressing all categories. The roadmap incorporates strategic action plans, including corrective, preventive, and motivational Kaizen actions, supported by continuous monitoring. This holistic strategy drives operational improvements and fosters cultural transformation across the organization.
- Effectiveness and efficiency: The proposed methodology achieves 7 out of 8 TPM pillars, compared to 4 in the existing approach, while balancing operational improvements and with organizational culture change to ensure long-term impact. With a three-year timeframe for strategic planning and periodic follow-ups, the methodology is more time-efficient and sustainable, reducing the need for frequent updates. Moreover, it engages the entire organization, from top management to operators, promoting collaboration and knowledge sharing—unlike the existing approach, which primarily involves maintenance staff.

The comparative analysis demonstrates that the ARM + NA methodology offers a more effective and efficient TPM program through a single, integrated application that also facilitates cultural transformation. This is further strengthened by the Kaizen policy, which combines actionable plans with a philosophy, where the whole company

Table 8
Yearly overview of OEE metrics and improvement targets.

Year	Availability	Performance	Quality	OEE	ΔOEE	Cumulative ΔOEE ACTUAL	Cumulative ΔOEE TARGET
0	84.0 %	93.5 %	88.0 %	69.1 %	—	—	—
1	88.2 %	95.4 %	89.7 %	75.5 %	6.4 %	6 %	5 %
2	90.1 %	96.7 %	90.0 %	78.4 %	2.9 %	9 %	10 %
3	92.6 %	96.3 %	92.8 %	82.7 %	4.3 %	14 %	15 %

Table 9

Comparative analysis with the only data-driven TPM approach presented in the literature using ARM: Comparison of existing literature methodology vs proposed methodology.

Features	Existing Literature Methodology	Proposed Methodology
Scope of the analysis	Identification of hidden relationships across failure modes	Identification of hidden relationships + Patterns recognition across operational variables
Methodology Variables analyzable	FMEA + ARM Failure modes	Dual application of ARM + NA All operational variables across the production line
OEE factors considered	Availability	Comprehensive OEE factors (availability, performance, quality)
Impact analysis on OEE	Identification of variables negatively influencing OEE	Analysis of both positive and negative variables influencing OEE and its factors
Association Rules Analysis	Support and Confidence factors analysis	Support and confidence analysis enhanced with advanced graph theory metrics (Out-degree, In-Degree, Centrality Degree)
Variables' classification	Risk Priority Number + Eisenhower matrix	Metric-based classification (Out-degree, In-degree, Centrality Degree), subject to review and approval by the TPM committee
Improvement strategy	Prioritization of "Urgent and Important" variables	TPM roadmap + strategic action plans for each critical category based on data insights
Improvement actions	Corrective actions	Kaizen action plan incorporating corrective, preventive, and motivational actions, integrated with a monitoring and feedback system
Effectiveness	4/8 TPM pillars addressed	7/8 TPM pillars addressed
Impact on organizational changes	Operational improvements	Transformation of operational processes and organizational culture
Efficiency	Time-consuming requiring frequent updates	Time-efficient with long-term strategic planning
Reimplementation frequency	Iterative methodology with frequent updates	Long-term strategic planning with periodic structured follow-ups
Planning timeframe	Not explicitly defined	3 years' timeframe for sustained improvements
Involvement	Maintenance personnel and operators	Full organizational involvement in continuous improvement, fostering collaboration and knowledge transfer

is actively involved in suggesting and implementing improvements, complemented by a reward system.

7. Conclusions and future developments

The proposed data-driven framework represents a significant advancement in the progressive implementation of TPM by leveraging advanced DM techniques. The primary goal was to strengthen the traditional TPM approach with more robust analytical foundations, enabling the identification of key production events that can most benefit from TPM strategies.

This methodology advances beyond existing applications that combine ARM with conventional methods by exclusively leveraging DM techniques, including a dual application of ARM e NA. Previous approaches, which primarily focused on failure events and availability, were limited in scope. In contrast, the proposed methodology enables a comprehensive exploration of all operational variables and their impact

on OEE and its factors—availability, performance, and quality—while expanding upon previous fully DM methodologies that focused solely on specific TPM pillars or availability losses. In addition, by considering only fully DM applications in the TPM domain—an area that remains underexplored—the proposed methodology provides an innovative contribution by not only analyzing individual OEE factors but also discovering hidden relationships between both high-impact positive and negative operational variables. This allows for a deeper understanding of system performance, surpassing traditional approaches that again focus predominantly on negative availability values. Furthermore, the new framework incorporates a metrics-based assessment, utilizing Out-degree, In-degree, and Centrality Degree to refine the traditional classification approach, advancing beyond the existing application based on support and confidence factors analysis and classification via the Eisenhower matrix. The proposed methodology offers a holistic action plan targeting all priority event categories—not just the “urgent and important” ones. A fundamental strength of the proposed approach is also its capacity to engage the entire organization, from top management to operators, fostering a culture of continuous improvement and enhancing collaboration across all levels. This inclusive approach lays a strong foundation for the transition towards Industry 5.0, where data-driven decision-making plays a pivotal role in achieving operational excellence.

Lastly, the effectiveness of the framework was demonstrated through a longitudinal industrial case study, which is noteworthy given the current relevance of the Lean 4.0 concept.

These advancements provide a holistic view of the production system, offering more precise and actionable insights into decision-making and continuous improvement strategies, resulting in fact in a 14 % improvement in OEE and successfully implementing 7 out of 8 TPM pillars quite in line with company expectations.

However, the work is subject to the following limitations:

1. Technical expertise and resources: Implementing such a data-driven approach requires substantial technical expertise and infrastructures, particularly in advanced data analytics and the integration of disparate data systems. This can pose challenges, especially during the early phases of implementation. For companies lacking systematic data collection processes, this could act as a barrier, but it can be mitigated through collaboration with academic or consulting experts.
2. Generalizability: The framework has been validated with a single case study, limiting the generalizability of the findings.
3. Exclusion of administrative TPM: While the framework addresses all TPM pillars except for administrative TPM, this pillar is rarely discussed in the literature as part of the complete TPM implementation process, though it is often explored independently.

In conclusion, the proposed framework demonstrates significant potential in improving the OEE and advancing TPM implementation. However, future research should focus on expanding its application across various industries to validate its robustness. Additionally, integrating the Administrative TPM pillar would complete the framework, addressing important areas such as material management and order processing, directly impacting production efficiency.

CRedit authorship contribution statement

Laura Lucantoni: Conceptualization, Methodology, Data curation, Writing – original draft. **Sara Antomarioni:** Data curation, Software, Validation, Writing – original draft, Methodology. **Filippo Emanuele Ciarapica:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Maurizio Bevilacqua:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

- Ale, F., Daniyan, I., Aderoba, O., & Adediran, A. A. (2024). An Artificial Intelligence-based Integrated Framework for Lean and Smart Manufacturing: A Case Study of the Rail Industry. In *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)* (pp. 1–10). <https://doi.org/10.1109/SEB4SDG60871.2024.10629752>
- Alenizi, F. A., Abbasi, S., Hussein Mohammed, A., & Masoud Rahmani, A. (2023). The artificial intelligence technologies in Industry 4.0: A taxonomy, approaches, and future directions. *Computers & Industrial Engineering*, 185, Article 109662. <https://doi.org/10.1016/j.cie.2023.109662>
- Amjad, M. S., Rafique, M. Z., & Khan, M. A. (2021). Leveraging Optimized and Cleaner Production through Industry 4.0. *Sustainable Production and Consumption*, 26, 859–871. <https://doi.org/10.1016/j.spc.2021.01.001>
- Antomarioni, S., Lucantoni, L., Ciarapica, F. E., & Bevilacqua, M. (2023). A Preliminary Implementation of Data-Driven TPM: A Real Case Study. In In: Crespo Márquez, A., Gómez Fernández, J.F., González-Prida Díaz, V., Amadi-Echendu, J. (eds) 16th WCEAM Proceedings. WCEAM 2022. Lecture Notes in Mechanical Engineering. Springer, Cham. (pp. 14–22). DOI: 10.1007/978-3-031-25448-2_2.
- Baghipur, M., Ganie, H. A., Ghorbani, M., & Andrade, E. (2022). On the spectral radius of the generalized adjacency matrix of a digraph. *Linear Algebra and Its Applications*, 653, 287–313. <https://doi.org/10.1016/j.laa.2022.08.017>
- Barriaga, A. L., Gonzales, M. G., & Ruiz-Ruiz, M. F. (2024). Lean Six Sigma and TPM for the Improvement of Equipment Maintenance Process in a Service Sector Company: A Case Study. *Communications in Computer and Information Science*, 2070 CCIS, 155–169. DOI: 10.1007/978-3-031-56373-7_13.
- Braglia, M., Castellano, D., Frosolini, M., Gabbriellini, R., Marrazzini, L., & Padellini, L. (2020). An ensemble-learning model for failure rate prediction. *Procedia Manufacturing*, 42, 41–48. <https://doi.org/10.1016/j.promfg.2020.02.022>
- Buddhakulsomsiri, J., Siradeghy, Y., Zakarian, A., & Li, X. (2006). Association rule-generation algorithm for mining automotive warranty data. *International Journal of Production Research*, 44(14), 2749–2770.
- Candra, N. E., Susilawati, A., & Heriswanto, & Setiady, W.. (2017). Implementation of Total Productive Maintenance (TPM) to Improve Sheeter Machine Performance. *MATEC Web of Conferences*, 135, 00028. <https://doi.org/10.1051/mateconf/201713500028>
- Chan, F. T. S., Lau, H. C. W., Ip, R. W. L., Chan, H. K., & Kong, S. (2005). Implementation of total productive maintenance: A case study. *International Journal of Production Economics*. <https://doi.org/10.1016/j.ijpe.2003.10.021>
- Correia Pinto, G. F., José Gomes da Silva, F., Octávio Garcia Fernandes, N., Carla Barros Casais, R., Baptista da Silva, A., & Jorge Vale Carvalh, C. (2020). Implementing a maintenance strategic plan using TPM methodology. *International Journal of Industrial Engineering and Management*, 11(3), 192–204. DOI: 10.24867/IJIE-M-2020-3-264.
- Crespo Márquez, A., de la Fuente Carmona, A., & Antomarioni, S. (2019). A Process to Implement an Artificial Neural Network and Association Rules Techniques to Improve Asset Performance and Energy Efficiency. *Energies*, 12(18), 3454. <https://doi.org/10.3390/en12183454>
- Djatna, T., & Alitu, I. M. (2015). An application of association rule mining in total productive maintenance strategy: An analysis and modelling in wooden door manufacturing industry. *Procedia Manufacturing*, 4, 336–343. <https://doi.org/10.1016/j.promfg.2015.11.049>
- Dogana, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166.
- Han, J., Cheng, H., & Xin, D. (2007). Frequent pattern mining : Current status and future directions. *Data Mining and Knowledge Discovery*, 15(1), 55–86. <https://doi.org/10.1007/s10618-006-0059-1>
- Jan, Z., Ahamed, F., Mayer, W., Patel, N., Grossmann, G., Stumptner, M., & Kuusk, A. (2023). Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities. *Expert Systems with Applications*, 216, Article 119456. <https://doi.org/10.1016/j.eswa.2022.119456>
- Kitazuka, R. E., & Moretti, C. (2012). Jidoka. In *Toyota by Toyota* (1st Edition, Vol. 1, p. 12).
- Kufre, E. U., Onokwai, A. O., & Onifade, M. K. (2024). Integrating Maintenance Practices with Industry 4.0 in Manufacturing System. In *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)* (pp. 1–5). <https://doi.org/10.1109/SEB4SDG60871.2024.10630010>
- Lopes, I., Senra, P., Vilarinho, S., Sá, V., Teixeira, C., Lopes, J., Alves, A., Oliveira, J. A., & Figueiredo, M. (2016). Requirements specification of a computerized maintenance management system - A case study. *Procedia CIRP*, 52, 268–273. <https://doi.org/10.1016/J.PROCIR.2016.07.047>
- Lucantoni, L., Antomarioni, S., Bevilacqua, M., & Ciarapica, F. E. (2022). Implementation of industry 4.0 techniques in lean production technology: A literature review. *Management and Production Engineering Review*, 13(3), 83–93.
- Lucantoni, L., Antomarioni, S., Ciarapica, F. E., & Bevilacqua, M. (2023). A rule-based machine learning methodology for the proactive improvement of OEE: A real case study. *International Journal of Quality & Reliability Management*. <https://doi.org/10.1108/IJQRM-01-2023-0012>
- Mendes, D., Gaspar, P. D., Charrua-Santos, F., & Navas, H. (2023a). Enhanced real-time maintenance management model—A step toward industry 4.0 through lean: conveyor belt operation case study. *Electronics*, 12(18), 3872. <https://doi.org/10.3390/electronics12183872>
- Mendes, D., Gaspar, P. D., Charrua-Santos, F., & Navas, H. (2023b). Integrating TPM and industry 4.0 to increase the availability of industrial assets: a case study on a conveyor belt. *Processes*, 11(7), 1956. <https://doi.org/10.3390/pr11071956>
- Mouhib, Z., Naciri, L., Gallab, M., Merzouk, S., Souli, A., Bhiri, B. E. L., & Dinardo, M. (2024). TPM and TQM: What Connections and How They Are Changing Through Industry 4.0 Technologies? (pp. 125–134). DOI: 10.1007/978-3-031-46849-0_14.
- Muchiri, P., & Pintelon, L. (2008). Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion. *International Journal of Production Research*, 46(13), 3517–3535. <https://doi.org/10.1080/00207540601142645>
- Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6), 441–453. <https://doi.org/10.1177/016555150202800601>
- Paiva, R. G. N., Melo, Y. R., A.V. Cavalcante, C., Tenório, V. A. S., & Do, P. (2024). Developing data-driven O&M policy through sequential pattern mining: A case study. *Computers & Industrial Engineering*, 193, 110318. DOI: 10.1016/j.cie.2024.110318.
- Pozzi, R., Cannas, V. G., & Rossi, T. (2024). Data science supporting lean production: evidence from manufacturing companies. *Systems*, 12(3), 100. <https://doi.org/10.3390/systems12030100>
- dos Reis, M. D. O., Godina, R., Pimentel, C., Silva, F. J. G., & Matias, J. C. O. (2019). A TPM strategy implementation in an automotive production line through loss reduction. *Procedia Manufacturing*, 38, 908–915. <https://doi.org/10.1016/j.promfg.2020.01.173>
- Sajid, S., Haleem, A., Bahl, S., Javaid, M., Goyal, T., & Mittal, M. (2021). Data science applications for predictive maintenance and materials science in context to Industry 4.0. *Materials Today: Proceedings*, 45, 4898–4905. <https://doi.org/10.1016/J.MATPR.2021.01.357>
- Samadhiya, A., Agrawal, R., & Garza-Reyes, J. A. (2024). Integrating Industry 4.0 and Total Productive Maintenance for global sustainability. *The TQM Journal*, 36(1), 24–50. <https://doi.org/10.1108/TQM-05-2022-0164>
- Tian Xiang, Z., & Jeng Feng, C. (2021). Implementing total productive maintenance in a manufacturing small or medium-sized enterprise. *Journal of Industrial Engineering and Management*, 14(2), 152. <https://doi.org/10.3926/jiem.3286>
- Tortorella, G. L., Fogliatto, F. S., Cauchick-Miguel, P. A., Kurnia, S., & Jurburg, D. (2021). Integration of Industry 4.0 technologies into Total Productive Maintenance practices. *International Journal of Production Economics*, 240, Article 108224. <https://doi.org/10.1016/J.IJPE.2021.108224>
- Tortorella, G. L., Saurin, T. A., Filho, M. G., Samson, D., & Kumar, M. (2021). Bundles of Lean Automation practices and principles and their impact on operational performance. *International Journal of Production Economics*, 235, Article 108106. <https://doi.org/10.1016/j.ijpe.2021.108106>
- Valamede, L. S., & Akkari, A. C. S. (2020). Lean 4.0: A new holistic approach for the integration of lean manufacturing tools and digital technologies. *International Journal of Mathematical, Engineering and Management Sciences*, 5(5), 854–868. <https://doi.org/10.33889/IJMEMS.2020.5.5.066>
- Wolska, M., Gorewoda, T., Roszak, M., & Gajda, L. (2023). Implementation and improvement of the total productive maintenance concept in an organization. *Encyclopedia*, 3(4), 1537–1564. <https://doi.org/10.3390/encyclopedia3040110>