



A Preliminary Implementation of Data-Driven TPM: A Real Case Study

Sara Antomarioni, Laura Lucantoni, Filippo Emanuele Ciarapica,
and Maurizio Bevilacqua^(✉)

Università Politecnica delle Marche, Ancona, Italy

{s.antomarioni,l.lucantoni,f.ciarapica,m.bevilacqua}@univpm.it

Abstract. Total Productive Maintenance (TPM) is one of the methodologies widely used to increase the availability of existing facilities reducing downtimes, stops, and defects, by improving manufacturing methods, usage, and maintenance equipment. Considering the large amount of data currently available thanks to the Industry 4.0 (I4.0) digitization processes, extending the analysis performed in TPM with the support of new techniques and tools is interesting. More in detail, in this work, Association Rule Mining (ARM) is used to identify the hidden relationships between different failure events, allowing their monitoring and prediction and improving the continuity and resilience of the production flow. The final aim pursued by the proposed approach is the development of a maintenance strategy in order to improve the Overall Equipment Effectiveness (OEE) of the selected process. An example case based on real data from an automotive company is used to present the approach and extract proper knowledge from its implementation.

1 Introduction and Background

Maintenance is nowadays considered one of the main strategic business activities for company performance improvement within Lean Production (LP) (Lucantoni et al. 2019): it is particularly useful to ensure the continuity of the production flow from a resilient perspective. Among LP practices, TPM is one of the most widely applied methodologies to increase the availability of existing facilities. Specifically, TPM has a relevant role in reducing stoppages, wastes, and defects and promoting employee participation in operation and maintenance (Au-Yong et al. 2022). TPM is usually combined with OEE assessment to find the cause of low values and provide suggestions for improvements (Sukma et al. 2022), paving the way towards perfect production. The current maintenance management systems, however, need a certain degree of personalization since their main features do not meet the requirements of each company when dealing with a wide amount of data (Lopes et al. 2016). Within TPM, Planned Maintenance is widely regarded in the literature as the main pillar (Morales Méndez and Rodriguez 2017): its main weakness can be recognized in the fact that it relies on the historical failure rate of the equipment but does not include any probability measure (Adesta et al. 2018). However, Predictive Maintenance is nowadays extensively used for failure prediction, equipment cost reduction, and performance improvement (Sahal et al. 2020). Novel data-driven techniques are required due to the large amount of data

available for knowledge extraction (Antomarioni et al. 2021). In parallel, some authors highlighted how Lean Automation can be applied for the concomitant implementation of I4.0 technologies into LP practices, even though the complexity of IT infrastructure necessary to fully integrate I4.0 into TPM could make such adoption less desirable (Tortorella et al. 2021). In line with this perception, despite I4.0 being one of the primary paradigms of the current industrial context (Marcucci et al. 2021), the current literature appears poor on how I4.0 techniques can really support LP principles and practices (Ferreira et al. 2022) showing that more research is needed in this area. One of the few examples in the existing literature presents that data mining techniques, such as ARM, can be integrated with traditional Pareto Chart and Ishikawa diagrams or network analysis in order to assess the magnitude of the production losses and identify the related causes within TPM (Djatna and Alitu 2015; Antomarioni et al. 2022).

Considering the existing research gap and the opportunities related to the importance of this research field, the focus of the proposed application is based on relating a metric derived from the well-known Failure Modes Effects Analysis – namely, the Risk Priority Number (RPN) - and ARM: from a practical point of view, they will be used to prioritize failure events; from a theoretical point of view, the aim of the proposed research approach is bridging the existing lack of research in this area through a novel data-driven approach. More in detail, RPN is used to identify the risk associated with each failure mode, considering the current best practices implemented in the company object of the study. Through ARM, instead, the hidden relationship existing between the occurrence of different failure events will be investigated. The last goal is to propose improvement actions that benefit the TPM strategy, improving the OEE and the continuous process flow. A case study from the automotive industry has been used as a pilot project to explain the proposed research approach.

In the rest of the paper, a general explanation of the proposed approach is provided in Sect. 2, while Sect. 3 contains its application to the case study. Conclusions and future research directions are drawn in Sect. 4.

2 Data-Driven TPM Approach

In order to introduce an effective data-driven TPM strategy in manufacturing, the proposed methodology can be summarized as in Fig. 1. Three main steps can be identified in carrying out such an application, as explained in the following sub-sections.

2.1 Data Collection and Pre-processing

Data collection and pre-processing: data represent the basis for an effective maintenance strategy; thus, this module is the foundation of the developed approach. It is fundamental to be able to access data from different sources and integrate them into a unique and reliable dataset. Indeed, the quality of the whole process relies on the quality of data, and the correctness of the decision that will be made is strictly related to them.

2.2 Data Analytics

Data analytics: The analytics phase is carried out on the integrated dataset produced in the previous step. It mainly consists of two further sub-steps: RPN calculation and Association Rule Mining. Firstly, indeed, Failure Mode and Effect Analysis (FMEA) is carried out to identify any possible failure modes in the production processes and the related RPN values. At this point, ARM is implemented to identify the failure events often occurring concurrently. The analysis can be limited to those failure modes having a high value of RPN, i.e., the ones that are considered more critical by the company or could be extended to the whole set of the identified failure modes. The objective will be, at this point, determining which are the failure modes frequently occurring concurrently: indeed, ARM aims to identify the relations among attributes and values stored in large datasets that frequently co-occur (Buddhakulsomsiri et al. 2006).

ASSOCIATION RULE MINING

In the following, a formal definition of the ARM process is provided: given a set of items (i.e., Boolean data) $I = \{\iota_1, \iota_2, \dots, \iota_n\}$ and given a set of transactions $T = \{\tau_1, \tau_2, \dots, \tau_m\}$ each of whom is composed by an itemset included in I . An Association Rule (AR) $\alpha \rightarrow \beta$ can be defined as an implication between itemsets - α and β - belonging to I ($\alpha, \beta \subseteq I$) and having no elements in common ($\alpha \cap \beta = \emptyset$). ARs' quality is determined through the calculation of different metrics. The Support (Supp) (1) and the Confidence (Conf) (2). Basing on them, ARM aims to identify relationships between failure events and select the ones requiring urgent and essential interventions. The association rules reported in Table 1 have been extracted through the ARM application. The co-occurrence of the failure events is obtained and, through them, decisions benefiting the continuous flow of the production can be made, prioritizing the rules having the highest Supp and Conf.

$$Supp(\alpha \rightarrow \beta) = \frac{\#(\alpha, \beta)}{\#T} \quad (1)$$

$$Conf(\alpha \rightarrow \beta) = \frac{Supp(\alpha \rightarrow \beta)}{Supp(\alpha)} \quad (2)$$

2.3 Decision Making

As a third step of the proposed approach, the main criticalities of the production process can be identified through the information provided by the ARM. Thus, an Eisenhower matrix is filled to classify the main criticalities and, most importantly, prioritize them. Such a matrix is built considering the RPN of the failure modes and the relationships identified through the ARM; its aim is to allow to classify the critical failure modes and prioritize components for maintenance intervention, defining appropriate preventive strategies. In addition, when a failure event occurs, the occurrence of the related failure modes should be inspected as a priority, in order to be able to intervene promptly.

3 Data-Driven TPM Implementation

The production system is an assembly line of an automotive company, composed of twelve fully automated stations. In standard conditions, the line operates during the whole



Fig. 1. Research approach

day, considering three shifts of eight hours each. It daily produces about 3,400 pieces (140 units per hour). The preventive maintenance system currently in place has optical sensors controlling parts position, manual operator checks, and planned maintenance for early equipment replacement every 1,000 parts produced. However, this strategy is currently not effective since unwanted failure events and stoppages of the production flow often occur, requiring immediate corrective interventions.

In the proposed case study, two data frames regarding daily production data and failure events are used to build the dataset. In all, it contains 1,122 integrated records referring to a time interval of six months. An extraction of the data frames with only the main columns is reported below.

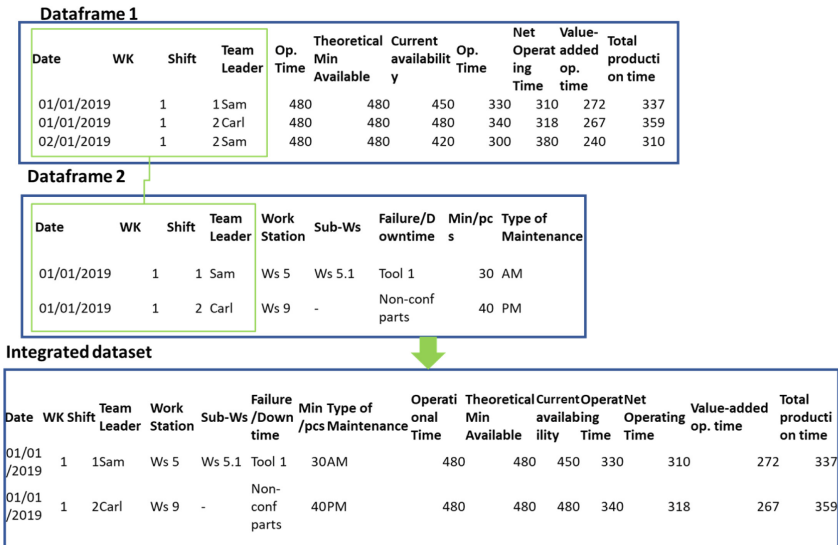


Fig. 2. Dataset for the analytics process

In order to build a reliable dataset for analysis, cleaning and standardization processes are carried out, removing all the inconsistencies and missing values. Specifically, empty columns (e.g., lack of operators, blackout, strike, etc.) were directly removed, while attributes with empty or negative values were well-analyzed and corrected through

brainstorming with line experts. In addition, the difference between the *Current availability* and the *Operating time* values of DataFrame1 was compared with the sum of all values in the *Min/pcs* column of DataFrame2 having the same merging key, also considering any further intermediate values in the first dataframe. Finally, both the *Failure/downtime* and *Description* columns of Dataframe2 were analyzed in order to ensure consistent nomenclature.

Lastly, 919 failure or downtime events are recorded during the monitored time interval. The resulting time dedicated to maintenance interventions corresponds to 28,417 min. In order to verify whether the implementation of the approach can be considered successful, the Overall Equipment Effectiveness (OEE) will be monitored. The *as-is* OEE is calculated daily with an average of 67%.

The analytics phase is carried out on the integrated dataset (see Fig. 2 for an excerpt). It mainly consists of two further steps: RPN calculation and Association Rule Mining. Firstly, Failure Mode and Effect Analysis (FMEA) is carried out to identify any possible failure modes in the production processes and RPN values associated with them. Taking into consideration the production process and the company's best practices, numerical ranges have been defined to classify the criticality of the RPN: *excellent-good* from 1 to 10, *good-sufficient* from 10 to 100, and *sufficient-poor* from 100 to 1000. According to this classification, 53% of the fault events showed an RPN higher than 10, which is the threshold value and required further investigation.

At this point, ARM is implemented to identify the failure events often occurring concurrently when the RPN is higher than the identified threshold. It should be noticed that the selected threshold resulted in being valid for the proposed application, while other case studies or different processes could increase it, decrease it or extend the analysis to the whole dataset.

The ARs are then mined using the integrated dataset as a starting point, but, as mentioned before, excluding the events concerning failure modes having an RPN under the threshold. Table 1 shows an excerpt of the results obtained: it can, for example, be noticed that failures 43 and 49 occur together in 3% of cases since the support of both the rules $43 \rightarrow 49$ and $49 \rightarrow 43$ is 3%. Conversely, when failure 43 occurs, 49 verifies in 24% of cases ($\text{Conf}(43 \rightarrow 49) = 24\%$); the opposite rule, instead, indicated that, when failure 49 occurs, the probability of occurrence of failure 43 is 29%.

Once the ARs have been mined, the Eisenhower matrix can be filled: Table 2 displays how the association rules are represented in it, so that prioritization can be made. First of all, the failure modes whose RPN is under the defined threshold are inserted into the non-important and non-urgent quadrant, since they are not object of the current study: forty failure modes are then excluded from the rest of the analysis. For the classification of the remaining ones, the ARs are used to fill the matrix: if a failure mode has an RPN above the defined threshold and it appears in the ARs mined both on the left- and right-hand side, then it is classified as urgent and important; if it appears in the left- or right-hand side of the ARs, then it is considered important but not urgent; if it does not appear in any rule, then it is not considered important, even though it can be urgent. The non-importance of the failure modes is related to the fact that they do not appear to be triggered by the occurrence of other failure events and they do not trigger others

Table 1. Some of the association rules among failure events with RPN above the threshold value

X		Y	Supp	Conf
26	→	46	5%	32%
43	→	47	4%	35%
43	→	49	3%	24%
46	→	26	5%	29%
46	→	47	4%	25%
49	→	43	3%	29%
49	→	47	4%	36%
59	→	26	5%	24%
59	→	46	4%	21%
59	→	47	9%	41%
91	→	46	5%	41%
91	→	47	4%	35%

either. The most critical failure modes in the upper red area are identified, namely, those requiring urgent action. Then, the extraction of the graphical results of the methodology is shown in Fig. 3.

Table 2. Criteria for the failure modes prioritization

RPN	ARs	Eisenhower matrix
Above threshold	“X” AND “Y”	Urgent and Important
Above threshold	“X” OR “Y”	Important and Non-Urgent
Above threshold	No	Urgent and Non-Important
Under threshold	Not analyzed in this work	Non-Urgent and Non-Important

In the upper right area, urgent and important failure modes are prioritized, namely, those with an RPN index above the threshold revealed by the FMEA analysis and closely related to each other as revealed by the ARs.

In this way, a data-driven strategy is defined to select which improvement strategies should be prioritized. TPM pillars are first applied to the failure modes considered urgent and important in order to mitigate their occurrence and anticipate their causes. Table 3 shows the improvement actions taken. From a continuous improvement perspective, the research approach is expected to be iteratively implemented, in order to gradually improve the quality of the overall process. After the implementation of such measures and an observation period of two months, an average improvement of the OEE by 2% has been achieved and a significant reduction of the failure events (about 70%). They can both be related to the actions taken and also to the monitoring of the process through

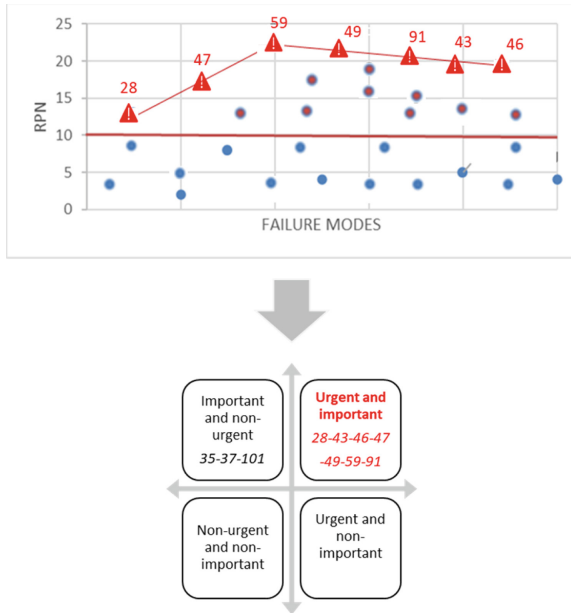


Fig. 3. Eisenhower matrix: failure modes classification and prioritization

the ARs results: indeed, when a failure event happens, the occurrence of the related failure modes should be inspected as a priority, in order to be able to intervene promptly. For instance, in the event of the occurrence of failure mode 59, from Table 1 we can see that three further failure modes could happen, i.e., 26, 46, and 47. Considering the confidence values of the association rules, it is evident that failure mode 47 is the most likely one ($Conf(59 \rightarrow 47) = 41\%$), followed by number 26 and 46 ($Conf(59 \rightarrow 26) = 24\%$; $Conf(59 \rightarrow 46) = 21\%$). In this way, preventive replacement of components can be performed when the probability is high (e.g., in case of number 47), while an inspection could be enough for the remaining ones.

Going into more detail of the results obtained, the 70% reduction in failure events is due to addressing those in the urgent and important quadrant (28–43–46–47–49–59–91). Our interest was to reduce the failure events of these modes being classified as the most urgent so 70% can be considered as satisfactory. OEE is, however, also affected by all other events, as well as other parameters such as, the item “cycle time different from standard” which is part of Performance losses. Longer monitoring would have allowed a broader view of the results obtained. In general, based on these considerations, while the reduction of mode-specific failure events is visible right away, the improvement of OEE definitely requires longer time and one with the continuous implementation of the methodology, as well as for continuous learning by operators.

That said, the implementation of the proposed methodology was also matched by a change of suppliers related to the current difficulties in the supply environment, which caused a deterioration in the quality of delivered materials and, consequently, a further

OEE penalty. Excluding this additional OEE penalizing factor, the overall improvement achieved as a result of the implemented approach was 10%.

Table 3. Improvement actions to mitigate the Important & Urgent failure modes

Failure mode code	Solution
FM 28	Production line refilling improvement
FM 43	Substitution of robot gripper sensors
FM 46	Production line refilling improvement
FM 47	Production line refilling improvement
FM 49	Programmed maintenance interval reduction
FM 59	Programmed maintenance interval reduction
FM 91	Training for workplace organization

4 Conclusions

The final goal of the proposed methodology is the identification of improvement actions for failure events prioritization in the field of lean automation. The added value deal with the application of a new data-driven approach as the I4.0 technique in a real case study supporting the TPM implementation and OEE improvement allowing a continuous production flow. After implementing the proposed data-driven TPM methodology, the results have been monitored for two months and four of the eight TPM pillars have been achieved in the short term: *Planned Maintenance*, avoiding recurrent failure events of FM49 and FM 59 keeping the equipment more operational, *Quality Management* reducing defects caused by FM 43 keeping the system more performant, *Education and Training* empowering maintenance operators about the new data management system and the workplace organization to avoid FM91, *Autonomous Maintenance* actively involving operators in minor maintenance tasks for the regular management of equipment as to cope with the FM 28, 46 and 47. In addition, it should be emphasized that *Continuous Improvement* may also be achieved through the periodic application of the proposed methodology. In conclusion, an average OEE improvement and reduction in the occurrence of failure modes important and urgent have been obtained due to the actions taken. When considering the improvement actions, the focus should surely be on the technical arrangements of the production systems, as well as on improving the policies currently in place. However, the training of the maintainers is fundamental in carrying out this approach since it ensures an improvement in operations quality.

Future research direction will focus on iteratively extending the proposed case study, also focusing on the remaining TPM pillars and on the failure modes not addressed in this study. With a view to exploiting association rules for monitoring the production system, a time dimension will be added in future developments to provide a more precise indication of the time at which it is necessary to be ready to perform preventive replacements.

Acknowledgment. This work was supported in part by the European Commission under the Project Reference No. 2021-1-IT01-KA220-VET-000033287. The authors would like to acknowledge the support of the researchers participating in the project “Business Continuity Managers Training Platform” (CONTINUITY) (business-continuity-project.eu).

References

- Adesta, E.Y.T., Prabowo, H.A., Agusman, D.: Evaluating 8 pillars of Total Productive Maintenance (TPM) implementation and their contribution to manufacturing performance. In: IOP Conference on Series: Materials Science and Engineering (2018)
- Antomarioni, S., et al.: Data-driven decision support system for managing item allocation in an ASRS: a framework development and a case study. *Expert Syst. Appl. Pergamon* **185**, 115622 (2021). <https://doi.org/10.1016/J.ESWA.2021.115622>
- Antomarioni, S., Ciarapica, F.E., Bevilacqua, M.: Association rules and social network analysis for supporting failure mode effects and criticality analysis: framework development and insights from an onshore platform. *Saf. Sci.* (2022)
- Au-Yong, C.P., Azmi, N.F., Myeda, N.E.: Promoting employee participation in operation and maintenance of green office building by adopting the total productive maintenance (TPM) concept. *J. Clean. Prod.* (2022)
- Buddhakulsomsiri, J., Siradeghyan, Y., Zakarian, A., Li, X.: Association rule-generation algorithm for mining automotive warranty data. *Int. J. Prod. Res.* **44**(14), 2749–2770 (2006)
- Djatna, T., Alitu, I.M.: An application of association rule mining in total productive maintenance strategy: an analysis and modelling in wooden door manufacturing industry. *Procedia Manuf.* (2015). <https://doi.org/10.1016/j.promfg.2015.11.049>
- Ferreira, W.D.P., et al.: Extending the lean value stream mapping to the context of Industry 4.0: an agent-based technology approach. *J. Manuf. Syst.* (2022)
- Lopes, I., et al.: Requirements specification of a computerized maintenance management system - a case study. In: *Procedia CIRP. Elsevier B.V.*, vol. 52, pp. 268–273 (2016). <https://doi.org/10.1016/J.PROCIR.2016.07.047>
- Lucantoni, L., Bevilacqua, M., Ciarapica, F.E.: An enhanced approach for implementing risk-based maintenance in a total productive maintenance perspective. In: XXIV Summer School “Francesco Turco” – Industrial Systems Engineering (2019)
- Marcucci, G., et al.: The impact of operations and IT-related Industry 4.0 key technologies on organizational resilience. *Prod. Plann. Control* (2021). <https://doi.org/10.1080/09537287.2021.1874702>
- Morales Méndez, J.D., Rodriguez, R.S.: Total productive maintenance (TPM) as a tool for improving productivity: a case study of application in the bottleneck of an auto-parts machining line. *Int. J. Adv. Manuf. Technol.* **92**(1–4), 1013–1026 (2017). <https://doi.org/10.1007/s00170-017-0052-4>
- Sahal, R., Breslin, J.G., Ali, M.I.: Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. *J. Manuf. Syst. Elsevier* **54**, 138–151 (2020). <https://doi.org/10.1016/J.JMSY.2019.11.004>
- Sukma, D.I., Prabowo, H.A., Setiawan, I., Kurnia, H., Fahturizal, I.M.: Implementation of total productive maintenance to improve overall equipment effectiveness of linear accelerator synergy platform cancer therapy. *Int. J. Eng. Trans. A: Basics* (2022)
- Tortorella, G., et al.: Towards the proposition of a lean automation framework: integrating industry 4.0 into lean production. *J. Manuf. Technol. Manag.* (2021)