Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 217 (2023) 1038–1047

Procedia Computer Science

www.elsevier.com/locate/procedia

4th International Conference on Industry 4.0 and Smart Manufacturing

Predictive maintenance on injection molds by generalized fault trees and anomaly detection

Pedro Nunes[a,b,*], Eugénio Rocha[c,d], José Santos[a,b], Ricardo Antunes[e]

[a]*Department of Mechanical Engineering, University of Aveiro, 3810-193, Portugal*
[b]*Centre for Mechanical Technology and Automation, Aveiro, 3810-193, Portugal*
[c]*Department of Mathematics, University of Aveiro, 3810-193, Portugal*
[d]*Center for Research and Development in Mathematics and Applications (CIDMA), Aveiro, 3810-193, Portugal*
[e]*OLI - Sistemas Sanitários, Aveiro, 3800-314, Portugal*

Abstract

Predictive maintenance (PdM) plays a key role in the Industry since it allows optimization of the schedule for proactive interventions and to take the maximum advantage of the useful lifetime of industrial assets. The reliability-centered maintenance (RCM) is based on equipment's reliability and allows the use of different maintenance strategies to optimize maintenance costs. With a recently proposed data-driven methodology entitled generalized fault trees (GFT), it is possible to assess the reliability of industrial equipment in real-time, based on their actual condition. In this paper, we exploit the GFT methodology in a completely different industrial scenario. A new training algorithm that intends to minimize operational costs, together with an anomaly detection technique (isolation forest) is presented to perform the predictive maintenance of injection molds at OLI, an enterprise specialized in producing plastic parts by the injection process. The results show that the proposed methodology may allow savings of 27.05% compared with preventive maintenance (PM) in optimized constant periods, and 63.43% compared to corrective maintenance (CM).

Keywords: Predictive Maintenance; Generalized Fault Trees; Reliability-Centered Maintenance; Industry 4.0

1. Introduction

Maintenance is one of the major concerns of enterprises nowadays, since its associated costs represent one of the biggest slices in the operational expenses [16]. There are different maintenance strategies, such as corrective maintenance (CM), where actions are taken only when there is a critical failure and the industrial asset is unable to work without intervention. On the other hand, preventive maintenance (PM) intends to proactively schedule interventions

* Corresponding author. Tel.: +351-234-370-359 ; fax: +351-234-382-014.
E-mail address: pnunes@ua.pt

1877-0509
10.1016/j.procs.2022.12.302

in equally spaced time intervals or working cycles. However, these strategies have clear limitations. CM allows the occurrence of failures that can damage several parts of the equipment and whose repair is often expensive and time-consuming, causing the unavailability of the asset. The PM typically under-exploits the asset's lifetime, since actions are taken when it could perform a considerably higher number of working cycles. In more severe scenarios, the failure occurs before the assigned period, and corrective actions need to be performed. According to [20], more than 30% of maintenance costs derived from poor maintenance scheduling. For this reason, predictive maintenance (PdM) and condition-based maintenance (CBM) strategies have received increasing attention, since they take into account the actual degradation state of the assets, usually determined from the monitoring and analysis of relevant working parameters.

Reliability-centered maintenance (RCM) is considered a type of PM [3]. It is not a specific maintenance strategy, but a decision-making process based on reliability analysis to ensure an efficient, cost-effective, and reliable maintenance [13]. Sometimes, when the failure is very simple to solve, spending time and resources on proactive actions is unnecessary, the RCM not only uses proactive actions but combines several maintenance policies to optimize its process. Reliability analysis is concerned with the failure probability of a given system or part, and relies on statistical data from relevant parameters of a given process. These distribution probabilities are usually approximated to known distributions, such as Weibull, Gamma, Log-Normal or Exponential [6]. One of the most well-known and used of these techniques is the Fault Tree Analysis (FTA) [19]. A fault tree (FT) is a directed acyclic graph that represents the failure path or chain of events that originated a problem. It represents graphically the root causes of a given problem through a tree-shaped structure, where the root represents the failure or top event (TE), leaves represent the root causes and are also called basic events (BEs). The relation between the BEs is modeled by gates, which are operators that calculate a new probability based on the probability of input BEs. For example, an OR gate calculates the probability that at least one of the BEs occurs, while the AND calculates the probability that both BEs have occurred. The cumulative density function (CDF) of the TE is calculated by computing all the gates.

Previously in [18], some limitations of the FTA were discussed, namely the fact that existing approaches assume that BEs follow an exponential (or Weibull) distribution, which in general can be a rough approximation of the real distribution. Moreover, the FTA and traditional reliability analysis methods rely on expert knowledge to determine the relevant parameters and components to take into account in the reliability assessment. To overcome these issues, a data-driven generalization of the FTA, the so-called Generalized Fault Tree (GFT) analysis, was presented by the authors, and proved to be effective in realistic scenarios, namely, it was employed to assess the reliability of a stamping press, and to automatically select the most suitable parameters to describe its failure. The fact that the BEs employed in the GFT approach are data-driven, allows the probability of TE to be updated according to the occurrence of each BE, and then the reliability assessment is performed based on the actual condition of components. In this work, we exploit the GFT methodology together with an anomaly detection technique (isolation forest), to assess the reliability of a mold (an important and expensive component of injection molding machines) in real-time, and adopt an RCM decision-making, by combining proactive and corrective maintenance actions to optimize its costs and increase the mold's availability.

Injection molding machines (IMMs) are industrial assets with many components subjected to failures, and whose maintenance costs are considerable. Furthermore, this equipment is widely used, due to the high productivity of the process and the flexibility in producing parts with different shapes. For this reason, much research has been done to improve the maintenance of this equipment. Some of them employ reliability analysis methods and RCM to reduce costs with maintenance interventions. For example, [6] applied a fuzzy technique to reduce the annual cost of the maintenance of an IMM, while [5] provided a copulas model for the same purpose. The authors in [17] presented an interesting work, which is still in an embryonic stage and where outlier detection is employed together with a statistical method based on kernel density estimation, and on the supervision of the distribution of variables over the time to study the possibility of applying PdM on the hydraulic pump of an IMM. The process parameters of injection were analyzed by [15], who adopted Nelson Rules [14], to detect abnormal patterns in the injection process. Other authors exploited machine learning (ML) techniques to handle IMMs maintenance. Namely, the authors in [2] employed several data mining techniques to forecast machine-related disruptions on IMMs, while [8] employed reinforcement learning for real-time scheduling, and [16] exploited ML classification algorithms to distinguish between optimal functioning and borderline functioning of an IMM. Other researches focused on creating the necessary apparatus to monitor and collect the necessary data for further analysis [7, 12].

Machine learning models are good to find patterns and non-linear relations in the data, but they are less efficient in globally capturing the failure behavior of industrial equipment in the presence of high imbalanced data sets [10], where failures represent a very small percentage of the data, which is quite common in industrial scenarios. Specially, in scenarios where there is a tool degradation process, it is interesting to consider all the data from the beginning of the component's life cycle, or the last maintenance action, to accurately determine its reliability or health condition, which usually is not accounted by ML techniques. Furthermore, these techniques are not directly interpretable by humans, since most of the time they are "black boxes". RCM demonstrated to be an effective decision-making process to reduce costs with the maintenance of IMMs. The works developed by [6] and [5] demonstrated very significant results concerning the savings with maintenance in real use cases. However, both approaches have common issues with traditional reliability techniques, namely, they rely on expert knowledge and approximate the probabilities distributions to known distributions. The GFT approach can mitigate these issues, by automatically selecting the relevant parameters for reliability assessment and providing a numerical approach that allows BEs to follow an arbitrary compact support distribution. Furthermore, the approach is completely data-driven, allowing the failure probability of the TE to be calculated based on the actual state of the asset, in real-time, which provides better results compared to an optimized PM strategy, as demonstrated by the results. Besides the application of the GFT methodology in a completely different use case, an anomaly detection technique, isolation forest, is applied to generate BEs that are used in the GFT's leaves, and a new training method, based on operational costs is proposed. Since in the present use case, proactive actions are not always the most suitable strategy, the new training method optimizes the GFT structure to propose proactive actions for the most serious failures, and let corrective actions play a role when a proactive action is not justified.

The remaining of this document is organized as follows; Section 2 presents the IMM use case at OLI. Section 3 and its subsections describe the GFT approach. Section 4 highlights the main results, and Section 5 presents the conclusions and final remarks.

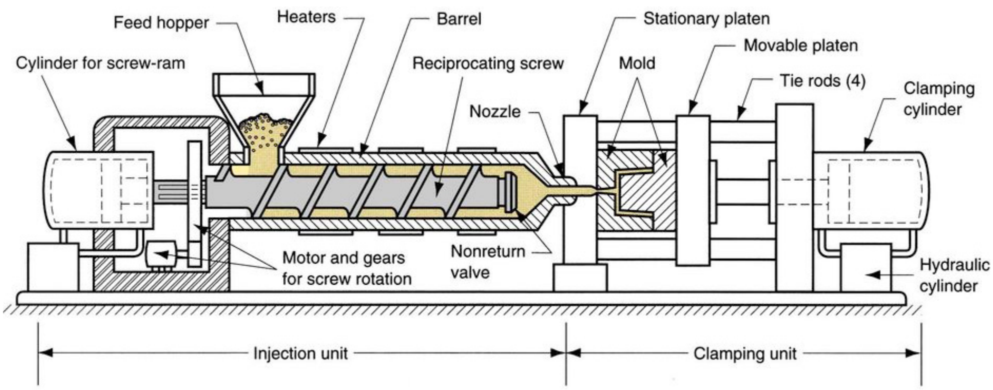

Fig. 1. Schematic of the IMM components [9].

2. The Injection Molds Problem

The injection molding process starts with plastic pellets in a feed hopper. These pellets are posteriorly melted and injected with pressure into a mold cavity, which is filled with plastic material. After a solidification phase, the material is ejected from the mold cavity, producing a 3-dimensional part as depicted by Fig. 1. This process causes degradation of the IMM and its components. In the concrete case of this work, the component that accounts for the highest number of malfunctions are the injection molds, and their most common failure is the mechanical one. When these problems occur, the mold must be repaired, since there are no spare molds of the same type, due to the fact that they are very expensive. It is important to note that proactive actions have a cost, namely with technicians that need to inspect the IMM and decide the actions to take. The mechanical failures of molds differ in severity, i.e., there are failures that are easily resolved with corrective actions, while some of them imply long unavailability times and expensive repairs. Having this in consideration, and taking into account that the CM is the only strategy applied currently by the enterprise, the objective of this work is to reduce the unavailability of the injection molds due to mechanical failures and consequently reduce the cost of maintenance, which is achieved not only with proactive actions but by selecting properly the correct maintenance strategy in each case, i.e. failures that are easily repaired do not justify proactive actions, while more severe ones can be mitigated with proactive actions.

The IMM is equipped with built-in sensors to monitor the injection process, following the standards of EUROMAP 63 [1]. The data collection is based on shared folders' architecture, and the collected data are stored in a file within the assigned shared folder, and then it is sent to a REST API, as demonstrated by Fig. 2. Together with the sensors' data, contextual data from the STAIN manufacturing execution system (MES) is also sent to the API. The REST API, when receiving a request, generates a Kafka message [21] that is forwarded to the Kafka Cluster, where processing will take place. The data from sensors and the contextual data from the MES are employed to train an optimized GFT model to reduce costs with molds' maintenance, where the BEs are anomalies determined by an isolation forest. Then, proactive actions are suggested to the maintenance team, when the failure probability given by the GFT model is higher than a threshold determined during the training phase.

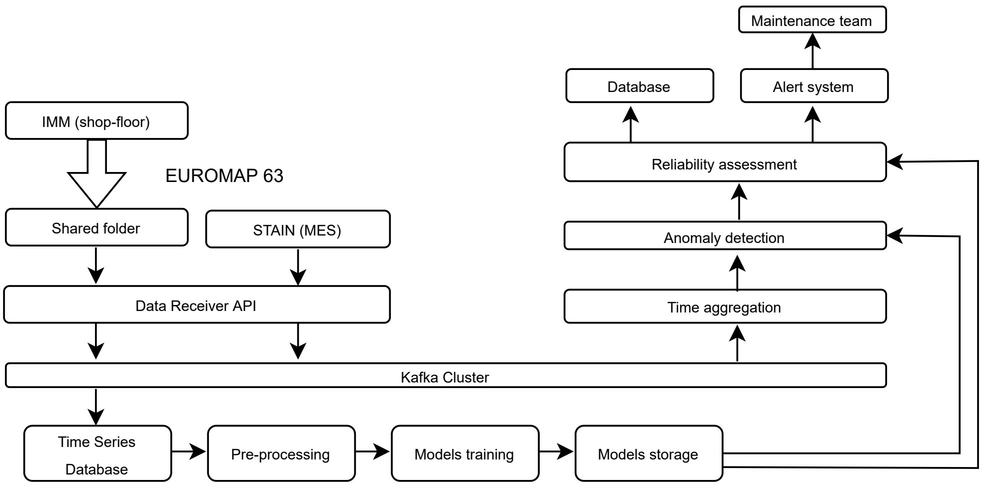

Fig. 2. Scheme of the data pipeline deployed as a micro-service architecture.

3. The GFT approach

The GFT approach encompasses several steps and could be divided into two main parts: the training phase, and the reliability assessment in real-time, as depicted by Fig. 3. The training phase consists in generating the tree-shaped structure that minimizes maintenance costs with a given mold. It starts with a pre-processing, common to both phases, which includes the following steps:

- The data from different molds are separated and labeled, by combining the information of real-time sensors and the contextual data from STAIN;
- The periods where the IMM is stopped are removed from the dataset to not bias the results;
- The data are aggregated in 5-minute time windows, and the aggregation variables, namely, mean, standard deviation (std), maximum (max), and minimum (min) are calculated;
- The dataset is enriched with new variables, resulting from the difference between consecutive time windows of the same variable.

Following the pre-processing, an anomaly detection technique, the isolation forest is trained, and the obtained model is stored to be used in real-time. The BEs that are used to train the GFT structure are obtained from the anomaly detection output, i.e., the anomalies in each variable are considered BEs. The distribution of BEs, as well as the obtained GFT model, are employed in real-time to assess the reliability of molds and suggest proactive actions. The next subsections describe in more detail the mentioned steps.

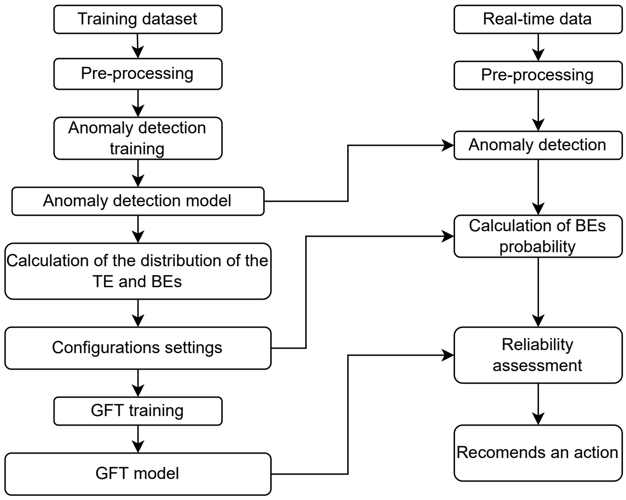

Fig. 3. General overview of steps in the GFT approach.

3.1. Basic events distributions

The data obtained from the IMM sensors are typically continuous, however, the GFT approach works with events and their associated probability distribution, thus, BEs are generated from a well-known anomaly detection technique, the isolation forest. This technique is based on a tree structure to isolate all the instances. Since anomalies are by definition instances that differ significantly from the majority of data, anomalies are isolated closer to the tree's root [11]. This method was employed, since it isolates anomalies, instead of profiling the normal behavior of data. Considering that in the IMM use case there are different configuration settings (temperatures, pressures, cycle time ...), due to the fact that different parts are produced, and the working parameters may change even for the same mold, there is no unique normal behavior model, since the which is normal for a given setup may be anomalous in another, thus isolating anomalous points is a better approach to detect anomalies in this case.

The anomaly detection is performed individually on each variable. The points where anomalies are detected are the ones where it is considered to exist a BE to build the GFT structure. The distribution of these BEs is calculated by considering the elapsed time since the last mechanical failure for that mold until the BE occurred, as illustrated by Fig. 4. The elapsed time only accounts for working time, since the inactive periods (when the IMM is not producing) were previously removed. The CDF of each variable can be obtained with a discretization step, h, of 5 minutes, which is the time window used to aggregate the data. From the CDF of each BE, the probability of BEs can be calculated and updated according to its occurrence, i.e., at a given time after the last repair, the probability of a given BE is the value of the CDF curve for this elapsed time, however, if the event already occurred, its probability is updated to the maximum value, i.e., to "1".

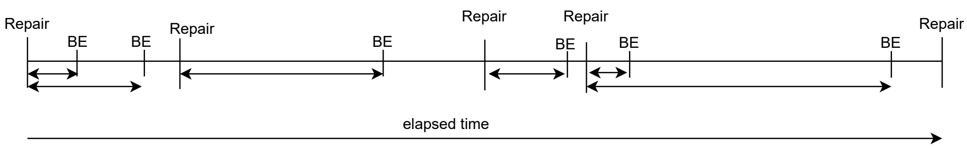

Fig. 4. Calculation of BEs distribution.

3.2. Cost-based GFT training

The training consists in finding the GFT structure and the probability threshold, p, that minimizes the costs of the maintenance of injection molds. Since the number of possible trees explodes with the number of BEs and the size of the set of operators, O, a pruning technique is employed, as presented by algorithm 1. To understand the training algorithm described by the pseudo-code, it is important to note that here we only use binary operators that take the probability of each BE in each time window and output a new cumulative probability, according to equations 1 and 2. The function *calculateProb* computes a new probability according to the input BEs and gate. At the beginning of the process, a dictionary of dictionaries (Python data structure) exists to save the probability in each time window, according to each tree $Tree$. Then, all the combinations with 2 BEs are performed, and the corresponding trees are calculated. The metric used to evaluate the trees and chose the most suitable one, consists in calculating the minimum cost of mold maintenance that could be achieved with a given tree. These costs are proportional to the time spent in the reparation of the mold, thus the percentage of the mold's unavailable time can be used as a cost metric, which is defined by equation 3, where N_i is the number of interventions performed according to the failure probability computed by each tree, ut_i is the unavailable time caused by each intervention and wt_i is the working time of the machine since mold's maintenance intervention $i-1$, until intervention i. Note that for the considered data set, only corrective interventions were performed, thus one has the unavailable time of mold for each failure, by the contextual data of the enterprise's MES. However, these data lack information concerning the different mechanical failures that may occur, thus it is hard to estimate in each case what would be the reduction of unavailable time by changing a corrective action to a proactive one. Since the time spent with corrective interventions has a wide range of possible values (see next section), it is reasonable to assume that the benefit of performing a proactive action, instead of a corrective one when the last one takes a short time period to be performed is irrelevant, or it can even increase unavailability of the mold since proactive revisions and maintenance actions can take longer than a simple and short corrective measure. Considering these aspects, we considered that ut_i for a proactive action assumes a constant value, which is the median of ut for all the corrective actions.

A proactive action is suggested each time the failure probability computed by the GFT is higher than a threshold value. Thus, for each computed $Tree$, an iterative process, varying p from 0.50 to 1.00 in intervals of 0.01 is performed to find the optimal p. The best GFT structure is the one whose optimal p represents the lowest cost, defined by equation 3. To reduce computational time, the calculation of more complex trees is made from the less complex ones. For example, a tree with 4 BEs can be calculated with a gate that has a tree with 3 BEs and a single BE as inputs, or 2 trees with 2 BEs each. These trees are saved in the data structure $Tree$ to speed up the calculation process. Note that the process is not exhaustive, since, for a tree with K BEs, only the trees that contain all the BEs of the best tree with $K-1$ BEs are considered. This avoids the explosion of possible trees and reduces the computational time of the process, compared with an exhaustive search.

$$AND(F_X, F_Y) = F_X F_Y, \tag{1}$$

$$OR(F_X, F_Y) = F_X + F_Y - F_X F_Y, \tag{2}$$

$$Cost = \sum_{i=0}^{N_i} \frac{ut_i}{wt_i + ut_i}. \tag{3}$$

4. Results and discussion

The methodology described in the previous section was applied to the injection mold which presented the higher number of malfunctions during the period of one month at OLI's facilities. The initial data set contains more than 100 variables of the process, described by the EUROMAP 63 protocol [1] and sampled with a 1Hz rate. After removing the variables that do not change during the process and receiving the feedback of engineers in the enterprise, a total of 10 variables that have an influence on the mold's operation were selected. Note that the GFT approach is not dependent on expert knowledge, since all the initial variables could be considered, however, it would result in more computational complexity. Since engineers from OLI have the insight about some parameters that are not related to

Algorithm 1 Cost-based GFT training with a pruning technique.

INPUT: $\mathfrak{D}(BE_i), i \in \{1, 2, 3, ..., n\}$ the updated CDFs of BEs in each timestamp, n is the number of BEs, C an array of costs, O, the set of operators, and N the maximum number of BEs to use in the leaves.

OUTPUT: GFT structure, GFT, and the threshold probability, p

```
Tree[1][BE_i] = 𝔇(BE_i)
Compute all the combinations, combs of 2 from all the BEs
for ∀ operator ∈ O  do
    for ∀ comb ∈ combs  do
        cdf = calculateProb(comb, operator)
        Tree[2][comb + operator] = cdf
    end for
end for
ps = [0.50, 0.51, 0.52, ..., 1.00]
for i ∈ {3, 4, ..., N}  do
    Cost(𝔇(TE),C,ps), ∀ T_t ∈ Tree[i − 1].items()
    Find the tree with lower cost and create a list BestBEs of their BEs
    for j ∈ {1, 2, ..., floor(i/2)} do
        for key1, value1 ∈ Tree[i − j].items() do
            for key2, value2 ∈ Tree[j].items() do
                if BestBEs ∈ BEs(key1)|BEs(key2) then
                    cdf = calculateProb(value1, value2, operator)        ∀ operator ∈ O}
                    Tree[i][(key1, key2) + operator] = cdf
                end if
            end for
        end for
    end for
end for
Output the tree with lower cost from Tree, and the respective p
```

the mold's failure, their feedback was exploited to speed up the training process. After the aggregation process, and the enrichment with new variables, the total number of variables is 80.

The GFT structure obtained from the training process is represented in Fig. 5 and the description of BEs is presented in Table 1. One of the main features of GFT analysis is related with the fact that a qualitative analysis of the process may be performed by the analyzing the tree-shaped structure, namely, minimal cut sets (MCS) can be determined. These are the sets with the minimum number of BEs whose occurrence ensures the occurrence of the TE [4]. From the presented skeleton, the following four MCS are obtained:

- $C_1 = \{E\_1, E\_3, E\_4\}$,
- $C_2 = \{E\_1, E\_3, E\_5, E\_6\}$,
- $C_3 = \{E\_2, E\_3, E\_4\}$,
- $C_4 = \{E\_2, E\_3, E\_5, E\_6\}$.

From this, it can be concluded that parameters such as the position of a cushion (for each injection unit), the cycle time, and the injection start position are the most relevant process parameters to infer the mechanical failure of molds since the chosen BEs during the training process are anomalies related to these variables. Moreover, the event E_3 is present in all the MCS, thus among others, it is the one that impacts more in the calculation of failure probability.

Table 1. Description of BEs.

BE	Description
E_1	The difference of the standard deviation of the cycle time between consecutive time windows is anomalous.
E_2	The difference of the mean stroke position at cushion for each injection unit between consecutive time windows is anomalous.
E_3	The mean stroke position at cushion for each injection unit is anomalous.
E_4	The difference of the maximum injection start position between consecutive time windows is anomalous.
E_5	The difference between the mean cycle time between consecutive time windows is anomalous.
E_6	The difference of the max stroke position at cushion for each injection unit between consecutive time windows is anomalous.

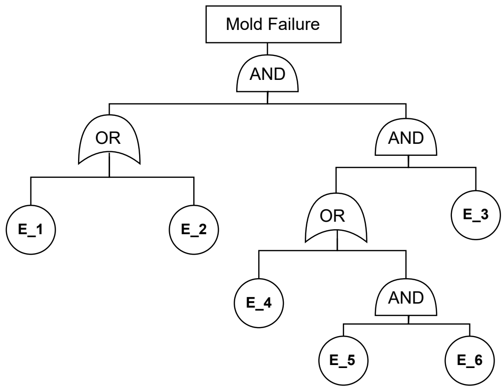

Fig. 5. Skeleton of the obtained GFT model.

Table 2 presents an overview of the mechanical failures of the selected mold in the reported period. As can be seen, the time that the mold works until the failure, ranges from 55 to 3030 minutes, and the unavailable time due to mold repairs has similar behavior, since it ranges from 5 to 3915 minutes. Two things may be concluded from here: it is difficult to estimate equally spaced intervals of time for proactive actions since mechanical failures of molds do not follow temporal patterns; proactive actions may not be the best approach for all types of failure, since, for example, failures number 2, 9, 10, and especially 5, took little time to be solved. Since these malfunctions are easily repaired, proactive actions are unnecessary.

Due to a lack of information about the failure mode and actions that were performed to overcome a given failure (this information is not maintained in the enterprise's MES), the cost of a proactive action was fixed as the median of the unavailable time with the current CM policy (247.5 minutes), which, when discussed with experts, was considered a conservative approach, since this value is an upper bound of the time that would be needed. The fourth column presents what would happen with the PM strategy in optimized equally spaced intervals of time. Since it only accounts for the working time, the majority of mechanical failures of the mold would not be avoided. However, this policy avoids the failure that causes the most impact in the unavailability of the mold, which may represent savings of 49.86% compared to the CM policy. When using the obtained GFT to optimize maintenance actions, we can see that the four major failures are avoided (penultimate column of Table 2). In fact, the only failures that are not avoided, are the ones whose unavailable time is lower than the estimated time needed for a proactive action (247.5 minutes). The fact that the GFT account for the actual state of molds, given by the BEs' probability, allows a reduction of the costs by 27.05% compared with the optimized PM strategy, and may allow a reduction of 63.43% compared to the CM.

An interesting observation that can be obtained from the results, is the fact the mold failures may be distinguished according to the unavailability caused by them, since, by setting the cost of a proactive action higher than some malfunctions, the GFT suggested an optimal combination of PdM and CM strategies, by suggesting proactive action near to severe failures, and allowing CM actions being taken in failures where the CM cost is lower compared to proactive actions. It is interesting, since it automatically balances proactive and corrective actions based on the cost

Table 2. Different maintenance strategies.

Failure number	wt_i CM (min)	Unavailable time (CM) (min)	wt_i PM (min)	Unavailable time (PM) (min)	wt_i GFT (min)	Unavailable time (GFT) (min)
1	2485	315	-	315	**2155**	**247.5**
2	140	35	-	35	-	35
3	300	180	-	180	-	180
4	55	585	-	585	**5**	**247.5**
5	2895	5	-	5	-	5
6	1890	390	-	390	-	390
7	850	725	-	725	**680**	**247.5**
8	3030	3915	**3025**	**247.5**	**2945**	**247.5**
9	2955	35	-	35	-	35
10	170	95	-	95	-	95

of the different failures and the model describes and distinguishes the failures with no previous knowledge more than the unavailability caused by each one.

5. Conclusions

In this work, an improvement of the GFT methodology (introduced by the authors in [18]) was applied to minimize the unavailability and cost of an injection mold at OLI's facilities. The training process was modified to optimize the maintenance costs, and a pruning strategy was adopted to speed up the process. The results show that estimated savings of 27.05% may be achieved, compared to a PM policy, and 63.43%, when compared with the CM policy.

The fact that the unavailability caused by each failure is very different, makes that not always a proactive action is the best action to take, since it can take longer time periods when compared to a simple corrective action that takes a short time. When using the cost-based training process, the GFT structure optimizes interventions according to their cost and can distinguish the failures according to their severity. Its relevance resides in an optimal combination of predictive and corrective actions, which is automatically generated to optimize maintenance costs.

The GFT approach brings important features compared with other techniques such as ML, since the generated model is easily understandable by technicians and engineers, due to its graphical simplicity, moreover, it accounts for all the events since the last repair. In comparison with statistical methods used in RCM, it uses the real distributions of data, instead of an approximation to known distributions. Although the computation of the failure probability by a GFT model is computationally light, the computational complexity of the training process is an aspect that may be improved, since it is dependent on the number of variables. Some aspects that need further improvement in this use case, encompass the creation of universal syntax to detail the causes of molds' malfunctions, which could result in a diagnostics model, that would give more insights and useful information to the maintenance team.

Acknowledgements. The present study was partially developed in the scope of the Project Augmented Humanity (PAH) [POCI-01-0247- FEDER-046103], financed by Portugal 2020, under the Competitiveness and Internationalization Operational Program, the Lisbon Regional Operational Program, and by the European Regional Development Fund. The first author has a PhD grant supported FCT – Fundação para a Ciência e a Tecnologia, I.P. for the PhD grants ref. 2020.06926.BD. The second author was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA), through the Portuguese Foundation for Science and Technology, reference UIDB/04106/2020. The first and third authors would like to acknowledge the University of Aveiro, FCT/MCTES for the financial support of TEMA research unit (FCT Ref. UIDB/00481/2020 & UIDP/00481/2020) and CENTRO01-0145-FEDER-022083 - Regional Operational Program of the Center (Centro2020), within the scope of the Portugal 2020 Partnership Agreement, through the European Regional Development Fund. The fourth author was supported by OLI - Sistemas Sanitários.

Data availability statement. The data sets used in this work are confidential information of OLI company manufacturing system, so they are not publicly available.

References

[1] Ayer, L., Ayer, P.H.L., 2000. Euromap 63 - Data Exchange Interface. Technical Report.

[2] Bazargan-Lari, M.R., Taghipour, S., 2021. A Data Mining Approach for Forecasting Machine Related Disruptions, in: 2021 Annual Reliability and Maintainability Symposium (RAMS), IEEE. pp. 1–7. URL: https://ieeexplore.ieee.org/document/9605729/, doi:10.1109/RAMS48097.2021.9605729.

[3] Deshpande, V., Mahant, P., 2013. Application of Reliability Centred Maintenance Methodology to Develop Maintenance Program for a Heavy Duty Hydraulic Stretching Machine. Australian Journal of Multi-Disciplinary Engineering 9, 177–184. URL: https://www.tandfonline.com/doi/full/10.7158/14488388.2013.11464858, doi:10.7158/14488388.2013.11464858.

[4] Durga Rao, K., Gopika, V., Sanyasi Rao, V.V., Kushwaha, H.S., Verma, A.K., Srividya, A., 2009. Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. Reliability Engineering and System Safety 94, 872–883. doi:10.1016/j.ress.2008.09.007.

[5] Fuentes-Huerta, M.A., González-González, D.S., Cantú-Sifuentes, M., Praga-Alejo, R.J., 2018. RCM implementation on plastic injection molding machine considering correlated failure modes and small size sample. The International Journal of Advanced Manufacturing Technology 95, 3465–3473. URL: http://link.springer.com/10.1007/s00170-017-1402-y, doi:10.1007/s00170-017-1402-y.

[6] Fuentes-Huerta, M.A., González-González, D.S., Cantú-Sifuentes, M., Praga-Alejo, R.J., 2021. Fuzzy reliability centered maintenance considering personnel experience and only censored data. Computers & Industrial Engineering 158, 107440. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360835221003442, doi:10.1016/j.cie.2021.107440.

[7] García, Á., Bregon, A., Martínez-Prieto, M.A., 2022. A non-intrusive Industry 4.0 retrofitting approach for collaborative maintenance in traditional manufacturing. Computers & Industrial Engineering 164, 107896. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360835221008007, doi:10.1016/j.cie.2021.107896.

[8] Ghaleb, M., Namoura, H.A., Taghipour, S., 2021. Reinforcement Learning-based Real-time Scheduling under Random Machine Breakdowns and Other Disturbances: A Case Study. Proceedings - Annual Reliability and Maintainability Symposium 2021-May. doi:10.1109/RAMS48097.2021.9605791.

[9] Groover, M.P., 2019. Fundamentals of Modern Manufacturing: Materials, Processes, and Systems, Enhanced eText, 7th Edition — Wiley. 4th ed. URL: https://www.wiley.com/en-se/Fundamentals+of+Modern+Manufacturing%3A+Materials%2C+Processes%2C+and+Systems%2C+7th+Edition-p-9781119475217.

[10] Leevy, J.L., Khoshgoftaar, T.M., Bauder, R.A., Seliya, N., 2020. Investigating the relationship between time and predictive model maintenance. Journal of Big Data 7, 36. URL: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00312-x, doi:10.1186/s40537-020-00312-x.

[11] Liu, F.T., Ting, K.M., Zhou, Z.H., 2008. Isolation forest, in: Proceedings - IEEE International Conference on Data Mining, ICDM, IEEE. pp. 413–422. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4781136http://ieeexplore.ieee.org/document/4781136/, doi:10.1109/ICDM.2008.17.

[12] Moreira, E.E., Alves, F.S., Martins, M., Ribeiro, G., Pina, A., Aguiam, D.E., Sotgiu, E., Fernandes, E.P., Gaspar, J., 2020. Industry 4.0: Real-time monitoring of an injection molding tool for smart predictive maintenance, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE. pp. 1209–1212. URL: https://ieeexplore.ieee.org/document/9212167/, doi:10.1109/ETFA46521.2020.9212167.

[13] Moubray, J., 1997. Reliability-centered maintenance, Butterworth Heinemann, p. 426.

[14] Nelson, L.S., 1984. The Shewhart Control Chart—Tests for Special Causes. Journal of Quality Technology 16, 237–239. URL: https://www.tandfonline.com/doi/abs/10.1080/00224065.1984.11978921https://www.tandfonline.com/doi/full/10.1080/00224065.1984.11978921, doi:10.1080/00224065.1984.11978921.

[15] Park, C., Moon, D., Do, N., Bae, S.M., 2016. A predictive maintenance approach based on real-time internal parameter monitoring. The International Journal of Advanced Manufacturing Technology 85, 623–632. URL: http://dx.doi.org/10.1007/s00170-015-7981-6http://link.springer.com/10.1007/s00170-015-7981-6, doi:10.1007/s00170-015-7981-6.

[16] Pierleoni, P., Palma, L., Belli, A., Sabbatini, L., 2020. Using Plastic Injection Moulding Machine Process Parameters for Predictive Maintenance Purposes, in: 2020 International Conference on Intelligent Engineering and Management (ICIEM), IEEE. pp. 115–120. URL: https://ieeexplore.ieee.org/document/9160120/, doi:10.1109/ICIEM48762.2020.9160120.

[17] Rivera, D.L., Scholz, M.R., Bühl, C., Krauss, M., Schilling, K., 2019. Is Big Data About to Retire Expert Knowledge? A Predictive Maintenance Study. IFAC-PapersOnLine 52, 1–6. URL: https://linkinghub.elsevier.com/retrieve/pii/S2405896319322645, doi:10.1016/j.ifacol.2019.12.364.

[18] Rocha, E.M., Nunes, P., Santos, J., 2022. Reliability Analysis of Sensorized Stamping Presses by Generalized Fault Trees, in: Proceedings of the International Conference on Industrial Engineering and Operations Management Istanbul, Turkey, March 7-10, 2022.

[19] Ruijters, E., Stoelinga, M., 2015. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. doi:10.1016/j.cosrev.2015.03.001.

[20] Salonen, A., Deleryd, M., 2011. Cost of poor maintenance: A concept for maintenance performance improvement. Journal of Quality in Maintenance Engineering 17, 63–73. doi:10.1108/13552511111116259/FULL/PDF.

[21] Zelenin, A., Kropp, A., 2021. Apache Kafka, in: Apache Kafka. Carl Hanser Verlag GmbH & Co. KG, München, pp. I–XVII. URL: https://kafka.apache.org/documentation/https://www.hanser-elibrary.com/doi/10.3139/9783446470460.fm, doi:10.3139/9783446470460.fm.