Maintenance Planning for Prognostic Health Management of Multi-Component Systems: A Case Study of Multi-Objective Optimisation for Railway Vehicles

Bernd G. Wagner[1,2], Sofoklis Kitharidis[1], Furong Ye[1], Thomas H.W. Bäck[1] and Niki van Stein[1]

[1] *Leiden University, Einsteinweg 55, 2333 CC Leiden, the Netherlands*
s.kitharidis@liacs.leidenuniv.nl
f.ye@liacs.leidenuniv.nl
t.h.w.baeck@liacs.leidenuniv.nl
n.van.stein@liacs.leidenuniv.nl

[2] *WDW-Rail Nederland BV, Kristiansandstraat 18, 2548 JT Den Haag, the Netherlands*
b.g.wagner@liacs.leidenuniv.nl

Abstract

Condition-based maintenance (CBM) offers an opportunity to improve classical preventive maintenance strategies by transitioning to a dedicated prognostics & health management (PHM) strategy for individual components. This means that time-based fixed maintenance intervals are being replaced by condition-dependent and component-specific interventions. For multi-component, complex systems, such a component-condition-based maintenance strategy yields multiple remaining useful life (RUL) prognostics. Maintainers must combine these into optimal maintenance plans that account for Reliability, availability, maintainability, safety and costs (RAMS-C). Combining multiple RUL prognostics with the system's real-world dynamics, including architecture, anticipated future operating conditions, economic interdependence, stochastic interdependence and structural interdependence, yields complex maintenance planning decisions. To address this problem, several studies propose solutions based on multi-objective optimisation algorithms. However, these are often based on a priori knowledge and do not account for the day-to-day dynamics of operational variations. This paper presents a methodology for managing the complexity of maintenance scheduling under these conditions. The model is applied to the maintenance history of a Voith Maxima 30CC cargo locomotive. A priori maintenance models are compared with real-life implementation and optimised models using evolutionary multi-objective optimisation (EMO) to assess the rigour of the a priori plans under varying operating conditions. The study shows that applying such algorithms can significantly reduce costs while increasing overall system reliability and availability.

1. Introduction

Traditionally, maintenance management strategies can be positioned between run-to-failure strategies, in which an asset is used until it fails, or preventive maintenance strategies, which aim to improve on run-to-failure models by converting general component run-to-failure statistics into pre-emptive maintenance plans. In these preventive maintenance plans, a component is scheduled for maintenance before failure to meet system reliability. CBM aims to further optimise these generic preventive maintenance plans by tracking the behaviour of individual components and forecasting their failure points and RUL. Thus, instead of relying on general failure statistics, the component- and condition-specific RULs serve as triggers for maintenance activities.

RUL is a widely used concept that has predominantly been applied to simple, single-component systems such as roller bearings (Heng and Nor 1998). Increasingly, RUL methodologies have been studied on more complex systems like battery lifetime calculations (Lipu et al. 2018), wind turbines (Rezamand et al. 2020), mechanical tools (Zhou et al. 2022) and industrial machinery (Ji et al. 2021). An important role in this development is played by the publicly available NASA dataset of aircraft engines (Saxena et al. 2008), which has led to extensive research on RUL concepts on these engines (Boujamza and Elhaq 2024). With the application of RUL to more complex systems, when compared to bearings, new challenges arise. Challenges such as inter-component damage propagation, multiple degradation models arising from multiple fault causes and degradation imbalance across varying operating conditions. These challenges do not arise in single-component systems (Lei et al. 2022).

Single-component systems, such as roller bearings or gun barrels (Kumar, Kalra, and Jha 2022), as well as seemingly more complex aircraft engine studies, typically consist of a single component comprising one or more parts. Based on historical behavioural run to failure (RTF) data, an RUL is calculated. In contrast, complex systems, such as road, air, marine, and rail vehicles, comprise multiple components that interact dynamically and under varying operating conditions. CBM of complex, multi-component systems requires tracking the RUL of multiple different components. These components may be structurally related, and degradation may depend on different physical quantities or utilisation measurements (Olde Keizer, Flapper, and Teunter 2017). As a result, asynchronous component degradation and multiple failure causes become typical challenges of multi-component systems. Maintainers must therefore aggregate multiple components RUL into a system view. This must then be translated into a meaningful maintenance schedule that minimises the number of interventions by grouping similar maintenance events.

The optimal grouping of maintenance activities in multi-component systems is an established research domain. However, existing approaches predominantly formulate the problem as a single-objective optimisation problem, typically minimising cost while satisfying reliability constraints (Wu et al. 2020). Alternatively, treat maintenance as a binary event, not distinguishing between various maintenance activities. However, in real-world applications, maintenance involves a more diverse set of conditions that influence planning. The decision for maintenance depends on costs, time, relevance, and the scope of maintenance. They are dependent on component-specific factors and are dynamic rather than constant. Maintainers consider these factors in their day-to-day maintenance planning to determine an optimal maintenance schedule. However, combining multiple components, asynchronous degradation, and diverse maintenance requirements complicate maintenance scheduling. As such, maintenance scheduling optimisation is challenging when accounting for the dynamic behaviour of multiple components' RULs.

Real-world optimised maintenance schedules try to balance maintenance and business requirements. Maintenance requirements can be derived from RUL statistics, while RAMS-C serves as a widely adopted framework for evaluating maintenance performance at the business level throughout an asset's lifecycle (Chengshan et al. 2026). In many industrial applications, safety and maintainability are treated as baseline requirements ensured through design standards and regulatory frameworks, while operational optimisation primarily focuses on reliability, availability, and lifecycle cost. Reliability characterises the probability of failure-free operation over time, whereas availability reflects the proportion of time a system remains capable of delivering its intended function.

As RUL and RAMS-C serve different business and engineering objectives, a maintenance model must account for both to determine an optimal maintenance schedule. Optimisation of schedules across multiple components is typically addressed in multi-objective optimisation (Deb, Sindhya, and Hakanen 2016). In multi-objective optimisation, Evolutionary computation (EC) (Bäck et al. 2023) is widely recognised as the most common approach, as it does not require prior knowledge of specific problem representations. Its population-based scheme naturally aligns with the objective of multi-objective optimisation, which aims to obtain a set of Pareto solutions. Several evolutionary multi-objective algorithms, such as NSGA-II (Deb, Pratap, et al. 2002) and MOEA/D (Zhang and Li 2007), have been extensively applied across a range of domains.

Based on the above, this paper aims to provide three contributions:

1. We formulate the maintenance planning problem for multi-component systems with multiple RULs as a multi-objective optimisation problem that considers reliability, availability, safety, and cost.
2. We analyse the discrepancy between real plans, theoretical models, and optimised schedules using a case study of the Voith Maxima 30CC locomotive.
3. We evaluate how operational variability affects the robustness of maintenance plans and discuss implications for PHM-driven maintenance optimisation.

This paper is organised as follows: Section 2 reviews the related work. Section 3 presents a generic mathematical foundation and model. Section 4 applies the model to a real-world case study of a Voith Maxima locomotive. Section 5 evaluates the model's outcomes. Finally, Section 6 provides limitations and recommendations for future research.

2. RELATED WORK

Despite that, both RUL and MOO have been studied extensively; the application of Multi-objective optimisation (MOO) in the context of PHM has received limited attention in the literature. Existing studies include the minimisation of RUL at engine replacement while maximising inspection intervals for turbofan propulsion engines (Chen and Liu 2025); the optimisation of maintenance costs, spare parts availability, and RUL for offshore wind turbine bearings (Du et al. 2025); and the maximisation of RUL while minimising maintenance costs for wind turbine transmissions (K. Wang, Deng, and Ding 2020). Other work considers the simultaneous minimisation of maintenance cost, makespan, total weighted job completion time, and total weighted tardiness, while maximising machine availability (Jin, Jiang, and Hou 2008).

These existing studies primarily optimise maintenance schedules under the assumption of a single RUL variable. Additionally, further research is required to evaluate trade-offs between competing objectives, such as reliability and maintenance costs (L. Wang, Zhu, and Zhao 2024). In complex systems composed of multiple components, multiple RUL estimates must be considered simultaneously. Each component has its own degradation trajectory and failure threshold, resulting in different predicted maintenance requirements. This necessitates consolidating individual maintenance events into coordinated interventions to avoid multiple events occurring within a short time frame.

Balancing maintenance costs, system downtime, and system reliability, therefore, constitutes a multi-objective optimisation problem. Maintenance grouping strategies for multi-component systems that explicitly account for system availability, reliability, and cost remain underexplored in the literature (Z. Wang et al. 2025). Moreover, many studies assume a single type of maintenance activity, whereas real-world maintenance planning is often more complex and depends on varying operating conditions and component wear profiles.

This paper addresses these challenges by optimising maintenance schedules to balance system reliability, availability, and cost. Based on multiple RUL predictions under specific operating conditions, long-term maintenance plans are generated. These plans are subsequently optimised using an EMO algorithm with respect to maintenance cost, reliability, and availability, resulting in maintenance strategies that best match the given operational conditions and the dynamics of multiple degradation patterns.

3. PROBLEM DESCRIPTION

This section presents the maintenance optimisation problem for multi-component systems by combining an RUL-driven maintenance schedule with an evolutionary multi-objective optimisation algorithm. The problem is addressed in two main steps. First, a generic maintenance event schedule (MES) is generated based on actual operating conditions. Second, this schedule is optimised into a maintenance plan using an evolutionary multi-objective optimisation (EMO). The EMO optimises system availability, reliability, and cost. The complete algorithm is referred to as MESEMO.

Component RULs can be expressed in various utilisation variables, such as kilometres, engine hours, or calendar time. However, a maintenance schedule must be presented in terms of time. Therefore, all other individual utilisation parameters ultimately need to be aligned with time. Further, each component might be subject to multiple failure modes. Consequently, the maintenance plan must account for one or more RUL predictions per component.

3.1. Construction of Events Schedule

Let

$$S = \{S_1, S_2, \dots, S_M\},$$

denote a maintenance schedule consisting of $M \in \mathbb{N}$ maintenance steps. Each maintenance step S_m, $m \in \{1, \dots, M\}$, is defined as a tuple:

$$S_m = (W_m, D_m, U_m, R_m, C_m)\,, \tag{1}$$

where:

- $W_m = (k_m, \lambda_m)$ contains the Weibull shape ($k_m > 0$) and scale ($\lambda_m > 0$) parameters, describing the failure probability of the maintenance step as a function of the specific utilisation U.
- $D_m \in \mathbb{R}_{>0}$ denotes the downtime required for the maintenance step.
- $U_m \in \mathcal{U}$ is the utilisation variable, where $\mathcal{U}$ represents admissible domains such as time, cycles, or mileage.
- $R_m \subseteq \{1, \dots, M\}$ represents dependencies with other maintenance steps.
- $C_m \in \mathbb{R}_{\geq 0}$ is the cost of performing the maintenance step m.

The schedule can be formally represented as:

$$S = \{(W_m, D_m, U_m, R_m, C_m) \mid m = 1, \dots, M\}\,. \tag{2}$$

The failure probability of any specific component is derived from the failure probability of each maintenance step S_m by applying the shape and scale parameters of W_m to the Weibull cumulative distribution function (CDF) such that:

$$F_m(u; k_m, \lambda_m) = 1 - \exp\left[-\left(\frac{u}{\lambda_m}\right)^{k_m}\right], \quad u \geq 0. \tag{3}$$

where:

- $F_m =$ is the failure probability of step m,
- $k_m > 0$ is the shape parameter,
- $\lambda_m > 0$ is the scale parameter,
- u is the utilisation variable, which is derived from U_m

Now, assuming a specific utilisation horizon T consists of a constant of all U_m reflecting the specific operating conditions considered for a specific case, the MES algorithm now projects utilisation-specific operating parameters T (e.g., project-specific kilometres, engine hours, cycles, time intervals, etc.) onto the standard maintenance schedule S_m, which results in a maintenance plan that considers specific operating conditions:

$$S_m^t = \Pi_{S_m}(t), \tag{4}$$

where the resulting un-optimised maintenance schedule S_m^t serves as the input to the multi-objective optimisation.

3.2. Evolutionary optimisation objectives

Existing RUL models focus solely on the singular probable failure time of an individual component without considering potential operational imbalances in multiple component environments. Balancing multiple RUL models requires consideration of multiple optimisation objectives: availability, reliability, and cost. This is addressed in the following three objectives:

1. Minimise total maintenance downtime; in multicomponent environments, this can be achieved by reducing the number of events and by effectively grouping them into joint interventions, maximising work that can be done in parallel.
2. Minimise aggregate reliability risk through the performance of maintenance as early as needed, given a specific reliability threshold given by the nature of the business.
3. Minimise total maintenance cost, including intervention setup cost, direct event costs, and planned downtime costs.

From an optimisation perspective, not only is the sum of all events relevant, but also how they can be efficiently bundled into a meaningful intervention. This reduces overall downtime by combining activities that can be done simultaneously and reduces general set-up costs associated with planning and performing a maintenance step.

As such, further to the denotion of $S = \{S_1, S_2, \ldots, S_M\}$ as the set of maintenance events, $I = \{I_1, I_2, \ldots, I_n\}$ denotes an intervention plan $n \in \mathbb{N}$ and each I_n consisting of 1 or more maintenance steps S_m^t.

The objectives are formulated as follows.

3.2.1. Objective 1: Minimising Total Downtime

System availability is inversely related to total downtime. In the implemented optimisation model, the quantity that is directly minimised is the total downtime, consisting of planned intervention downtime and expected unplanned downtime due to failure risk:

$$\min_S D_{\text{total}}(S) = D_{\text{planned}}(S) + D_{\text{unplanned}}(S). \tag{5}$$

The planned downtime is computed per intervention as one baseline delivery day plus the number of workshop days required by the longest effective maintenance task scheduled in that intervention:

$$D_{\text{planned}}(S) = \sum_{v \in \mathcal{V}(S)} \left(1 + \left\lceil \frac{\max_{m \in \mathcal{E}(v)} D_m}{H_{\text{day}}} \right\rceil\right), \tag{6}$$

where $\mathcal{V}(S)$ denotes the set of maintenance visits induced by schedule S, $\mathcal{E}(v)$ denotes the set of effective maintenance tasks performed during visit v, and H_{day} is the number of working hours per day.

The expected unplanned downtime is computed as:

$$D_{\text{unplanned}}(S) = \sum_{m \in \mathcal{E}(S)} F_m(t)\, D_m^{\text{fail}}, \tag{7}$$

where D_m^{fail} denotes the expected downtime associated with failure of maintenance task m.

3.2.2. Objective 2: Minimising Failure Risk

Reliability in the implemented model is represented through the failure probabilities of the effective maintenance events in the schedule. For each maintenance event m, a Weibull-based failure probability $F_m(t)$ is computed at the scheduled intervention time. Rather than controlling reliability at the level of individual maintenance events, the model evaluates the overall reliability of a schedule through the average failure probability across all effective events:

$$\bar{R}_{\text{risk}}(S) = \frac{1}{|\mathcal{E}(S)|} \sum_{m \in \mathcal{E}(S)} F_m(t), \tag{8}$$

where $\mathcal{E}(S)$ denotes the set of effective maintenance events in schedule S.

A maximum acceptable average risk threshold τ is prescribed. Schedules with average risk above this threshold are considered infeasible. For feasible schedules, the optimisation favours solutions whose average reliability risk lies as close as possible to the threshold from below, since this allows maintenance to be delayed where possible without violating the reliability requirement. This is expressed as:

$$\min_{S} R_{\text{slack}}(S) = \max(0, \tau - \bar{R}_{\text{risk}}(S)). \tag{9}$$

This formulation reflects the idea that reliability should be controlled at the level of the maintenance schedule as a whole. This formulation ensures that reliability is controlled at the schedule level through a prescribed upper bound on average failure risk.

3.2.3. Objective 3: Minimising Maintenance Costs

In the implemented optimisation model, the total maintenance cost comprises four components: a setup cost per maintenance intervention, the direct execution cost of the effective maintenance tasks, the cost of planned downtime, and the expected cost associated with failure risk. The resulting cost objective is defined as:

$$\begin{aligned}\min_{S} C_{\text{maint}}(S) = \sum_{v\in\mathcal{V}(S)} C_{\text{SETUP}} + \sum_{m\in\mathcal{E}(S)} C_m + \\ C_{\text{DAY}}\, D_{\text{planned}}(S) + \sum_{m\in\mathcal{E}(S)} F_m(t)\, C_m^{\text{fail}},\end{aligned} \tag{10}$$

where $\mathcal{V}(S)$ denotes the set of maintenance visits induced by schedule S, $\mathcal{E}(S)$ denotes the set of effective maintenance tasks, excluding tasks that are already covered by a higher-level maintenance action. C_{SETUP} is the fixed setup cost per visit, C_m is the direct cost of maintenance task m, C_{DAY} is the cost per downtime day, and C_m^{fail} is the failure cost associated with maintenance task m.

3.3. Multi-Objective Optimisation Formulation

In the implemented optimisation model, the search is guided by total maintenance cost, total downtime, and a schedule-level risk measure based on the average task-level failure probability. Let

$$\bar{F}(S) = \frac{1}{|\mathcal{E}(S)|} \sum_{m\in\mathcal{E}(S)} F_m(t) \tag{11}$$

denote the average effective task-level failure probability over schedule S, and let τ denote the prescribed risk threshold. Two auxiliary quantities are defined:

$$R_{\text{viol}}(S) = \max(0, \bar{F}(S) - \tau), \tag{12}$$

$$R_{\text{slack}}(S) = \max(0, \tau - \bar{F}(S)). \tag{13}$$

The implemented optimisation problem is therefore:

$$\min_{S} \mathbf{F}(S) = \Big(C_{\text{maint}}(S), D_{\text{total}}(S), R_{\text{slack}}(S)\Big), \tag{14}$$

subject to

$$R_{\text{viol}}(S) = 0, \tag{15}$$

and the admissibility constraints induced by feasible maintenance windows, legal obligations, grouping restrictions, and dependency relations.

where $\mathcal{S}$ is the set of admissible maintenance actions. The Pareto-optimal front is obtained using the specific EMO:

$$\widehat{\mathcal{P}}^{*} = \mathcal{A}_{\text{EMO}}\Big(\min_{S} \mathbf{F}(S) \mid S_m \in \mathcal{S}\Big). \tag{16}$$

Decision variables $\mathbf{s} = (s_1, \ldots, s_d)$ represent the unoptimised maintenance schedule derived from RUL predictions. Objective functions and constraints are defined as:

$$\mathbf{F}(\mathbf{s}) = (f_1(\mathbf{s}), \ldots, f_K(\mathbf{s})), \quad g_j(\mathbf{s}) \le 0,\ h_\ell(\mathbf{s}) = 0, \tag{17}$$

where g_j and h_ℓ represent admissibility and scheduling constraints, such as feasible maintenance windows, legal obligations, and grouping restrictions.

4. CASE STUDY AND MODEL CONSTRUCTION

A historical maintenance dataset of a Voith Maxima (30CC) locomotive was used in this study. The Maxima is a six-axle Co-Co diesel locomotive equipped with a Voith Turbo hydraulic drive and an ABC marine diesel engine adapted for railway service (Figure 1). Built in Germany between 2006 and 2010 for heavy freight transport, it remains in active service. Compared to typical railway vehicles, the Maxima features a complex maintenance regime spanning five major components: vehicle body, traction engine, running gear, brake system, and side panels. Each component has a single utilisation parameter (e.g., time, distance, or engine hours) and multiple failure causes that drive maintenance events. These are based on average RULs of the components' parts. Accordingly, this system can be considered as a multi-component system with multiple failure categories per component. Appendix A: Component Maintenance Plan, provides the overview of maintenance criteria for each component.

Figure 1. Voith Maxima 30CC locomotive

The dataset covers over 15 years, including 871,248 km, 30,352 engine hours, and more than 300 inspection reports. Partially handwritten reports were manually transcribed and decoded, yielding 200 maintenance events. The resulting historical maintenance table is validated by field experts and forms the basis for the comparative analysis.

For modelling purposes, each component's utilisation parameter serves as the independent variable in the RUL prediction. Higher-level maintenance events include all lower-level events. As such, a form of economic dependence among events emerges, while structural dependence is determined by the design of individual maintenance events. The system architecture is such that failure of any single component results in overall system failure (full system-level functional dependencies). Components themselves function independently. Maintenance activities from different components can be performed in parallel; otherwise, interventions are executed sequentially. Individual maintenance steps can be grouped into maintenance interventions. A single maintenance intervention may combine multiple component-level maintenance events, and costs and downtime are accounted for on a per-intervention basis. Maintenance time and cost assumptions follow industry benchmarks, with each intervention requiring 1 delivery day, EUR 2,000 for delivery and return, and EUR 1,000/day for workshop costs. Infinite maintenance resources are assumed. System-level *availability* is calculated as the vehicle's operational time minus the days spent on all interventions. *Cost* is the sum of all maintenance costs and intervention costs. To derive *reliability* statistics of each maintenance event, a failure probability is projected using the Weibull cumulative distribution function:

$$F(u;\beta,\eta) = 1 - \exp\left[-\left(\frac{u}{\eta}\right)^{\beta}\right], \qquad (18)$$

The RUL model has been derived by fitting a Weibull distribution to each individual maintenance event. In line with industry-wide reliability assumptions about general failure and their conversion into maintenance intervals default $\beta = 3$, and $P_{fail} = 0.01$ are suggested for modelling each singular failure probability function. The resulting CDF for each maintenance step is considered to represent the RUL model for each failure cause.

Failure to complete safety-critical and legally mandated inspections is treated differently in the implementation. For selected capped maintenance categories, the effective utilisation in the failure model is not allowed to exceed the due interval, such that:

$$u^{\star} = \min(u, u_{\mathrm{due}}). \qquad (19)$$

This means that, for these categories, the Weibull failure probability is capped at the due-point value rather than increasing further beyond the prescribed inspection interval. This reflects the legal obligation to only use vehicles that comply with legal and safety regulations.

The individually suggested RULs of all components are grouped into optimal interventions using an NSGA-II algorithm with a 30-day minimum time window between interventions to prevent repeated short-term workshop visits. Due to NSGA-II's proven effectiveness and computational efficiency in EMO, it is a solid reference model. Furthermore, it remains one of the most widely used and best-proven algorithms for models with up to 3 objectives. (Rahimi et al. 2022)

Simulations were conducted under three maximum failure probability thresholds to explore often-used reliability numbers:

$P_{fail} > F_{threshold}, F_{threshold} \in \{0.01, 0.05, 0.10\}$.

The NSGA-II algorithm is parametrised with the following settings: Window size: 30 days Minimum reliability: 0.01, 0.05, and 0.1. Crossover Probability: $P_c = 0.9$ Mutation probability: $P_m = 0.1$ Generations: 100

The 3 resulting Pareto fronts provide interesting insights into the dynamics of accepting higher failure risk and are provided in Appendix B: Pseudo-Code.

5. RESULTS, DISCUSSION AND CONCLUSION

The maintenance scheduling of the Voith Maxima poses a real-world asset management challenge due to the complexity arising from multiple RULs, their interdependencies, and the daily variation in utilisation. The locomotive and its maintenance concept were designed with the potential implementation of CBM in mind, including the installation of an onboard diagnostic system to collect sensor data. Despite the technical capability, the integration of condition monitoring into maintenance planning has not been realised in practice. One of the main barriers appears to be the complexity of managing the re-

Table 1. Comparison of real, manufacturer-based, and optimised maintenance schedules. For each optimised setting, we report the selected representative Pareto solution.

Parameter	Real	Theoretical	Opt. (1%); Sol. 1	Opt. (5%); Sol. 5	Opt. (10%); Sol. 10
Events	186	146	142	142	142
Interventions	81	127	73	71	67
Cost (€)	2.900.086	1.373.500	1.119.000	1.107.000	1.084.500
Reliability	98.59%	99.13%	99.00%	95.12%	90.09%
Availability	83.68%	94.35%	96.16%	96.04%	95.94%
Downtime (D)	918	326	215	220	228

sulting maintenance schedules, which places significant demands on planners and maintenance personnel.

Combining the manufacturer's maintenance plan with the real-world utilisation provides a theoretical event schedule. Using the suggested MESEMO framework, the Pareto-optimal solutions for this theoretical schedule were identified for 3 reliability thresholds $P_fail = 0.01, 0.05, 0.1$, enabling a systematic comparison of cost, reliability and availability. The best-performing solutions are summarised in the results table 1, while the complete set of Pareto fronts is presented in Appendix C: Pareto Views.

The results of the various scenarios highlight interesting differences between theoretical planning and real-world practice. The first notable finding is the substantial cost difference between the RUL based plans and the real-world situation. In reality, maintainers appear to combine events into interventions economically (only 81 interventions, compared to the theoretical models' 127), but availability and reliability fall below the baseline model. They do so at nearly double the cost. One reason is the substantial discrepancy between actual maintenance events and theoretical recommendations. A detailed event-based comparison between the executed events and the model shows that 78 of the 146 required maintenance events were skipped or added. This explains the overshoot of 186 real events versus the required 146. This is also reflected in the downtime: 326 vs 918, a significant difference. However, the difference in downtime can also be attributed to other factors, such as a lack of business or inaccuracies in the downtime benchmark figures. If we apply the same availability calculation to the real model as to the theoretical model, the actual downtime would be 232 days (95.87%), which is lower than in the theoretical model. In conclusion, although the real world can efficiently bundle events into interventions, it comes at a price. Something that maintainers should consider when budgeting life cycle models.

Across the evaluated optimised solutions, all models outperform the real and theoretical cases. Compared to the theoretical model, cost savings are around 20%. Which is substantial. The number of interventions is 45% lower than in the manufacturer's model, and availability is over 100 days lower compared to the reference case. All in all, significant improvements. Trying to reflect the benefits of the optimised models to the real case is slightly more subjective due to the ambiguous nature of the root data. However, it is reasonable to expect similar pro-rate improvements across all fronts. Improving a manufacturer's model will likely improve its real-world derivative as well. Furthermore, presenting the optimised results to maintenance organisations might provide insights and awareness into the trade-off among intervention reduction, costs, and failure risks. By understanding the dynamics, maintenance organisations might learn from such an experience and adopt their behaviour. Given the limited one-time effort required to set up and apply MESEMO and similar algorithms, there is a clear case for maintenance organisations to consider integrating optimisation algorithms into their planning process.

Comparing the Pareto fronts across all three failure thresholds reveals only marginal differences in cost, reliability, and availability. Surprisingly, the reliability threshold of $P_fail(0.1)$ outperforms $P_fail(0.01, 0.05)$ on downtime and reliability while being only marginally more expensive (3%). Even though the number of interventions dropped to 67 at the 0.1 failure threshold. Over the 15-year time horizon, these differences are negligible for rail vehicle maintenance. These results are specific to the vehicle's parameters, the utilisation, and the model inputs. As such, generalising the results is not possible. Also, other business cases might assign different weights to the optimal model, for example, by emphasising reliability in safety-critical environments or by limiting the number of events that continue processes.

In conclusion, this case study demonstrates that increasing the number of RULs leads to more complex scheduling challenges. Maintenance organisations struggle with this and consequently experience lower equipment availability, higher failure risk, and greater costs. Applying dynamic optimisation algorithms, such as the presented MESEMO, can significantly support maintenance decision-making and provide significant returns. It further provides a benchmark of current maintenance planning performance and raises operational awareness of the dynamics of RAMS-C targets.

6. LIMITATIONS AND RECOMMENDATIONS

This study is subject to some limitations. First, it focuses exclusively on preventive maintenance, whereas practical maintenance strategies involve both preventive and corrective actions. Consequently, the effects of corrective failures on system costs, availability, and reliability are not captured. This might unfairly favour models with higher failure risks. Second, the comparison with historical maintenance records, while informative, is also subject to uncertainty. Inspection reports contain inconsistencies and occasional inaccuracies, and their transcription requires interpretative judgment. Moreover, the dataset spans approximately fifteen years, during which personnel, documentation practices, and interpretations of maintenance procedures evolved, introducing additional variability. Third, the case study accounts for economic interdependence but only addresses structural dependencies to a limited extent. It excludes stochastic interdependencies and variations in operating conditions, and is based on a simplified functional dependency structure. NSGA-II may become less suitable for more complex models incorporating these effects, as they introduce non-linear interactions between decision variables and objectives, potentially leading to non-convex search spaces, constraint propagation, and reduced convergence efficiency. Alternative approaches, such as MOEA/D, NSGA-III, and surrogate-assisted evolutionary optimisation methods, may offer improved convergence behaviour for highly coupled systems. Finally, the theoretical and optimised models assume linear utilisation across all parameters, whereas real-world operations are affected by seasonality and random noise. Also, the modelling construction is based on a posteriori knowledge, where the full lifecycle utilisation is already known. In practice, utilisation evolves, and maintainers must act on uncertain anticipated future utilisation.

Future research should extend the modelling framework to incorporate corrective maintenance and its influence on performance metrics. Methodologies that account for non-linear utilisation patterns could enhance the realism and robustness of optimisation results. Furthermore, this study considers a single set of utilisation parameters. Evaluating maintenance schedules across a landscape of utilisation scenarios could provide deeper insights into the robustness of optimised schedules and the potential benefits of (CBM).

REFERENCES

Bäck, Thomas HW et al. (2023). “Evolutionary algorithms for parameter optimization—thirty years later”. In: *Evolutionary Computation* 31.2, pp. 81–122.

Boujamza, Abdeltif and Saâd Lissane Elhaq (2024). “Optimizing Remaining Useful Life Predictions for Aircraft Engines: A Dilated Recurrent Neural Network Approach”. In: *IFAC-PapersOnLine* 58.13, pp. 811–816. DOI: 10.1016/j.ifacol.2024.07.582.

Chen, Yan and Cheng Liu (Sept. 2025). “Sequential Multi-Objective Multi-Agent Reinforcement Learning Approach for System Predictive Maintenance of Turbofan Engine”. In: *Advanced Engineering Informatics* 67, p. 103553. DOI: 10.1016/j.aei.2025.103553.

Chengshan, Yi et al. (2026). “Rams-based modeling and decision optimization for condition-based maintenance strategies of EMU components”. In: *Measurement* 268, p. 120523. DOI: 10.1016/j.measurement.2026.120523.

Deb, Kalyanmoy, Amrit Pratap, et al. (2002). “A fast and elitist multiobjective genetic algorithm: NSGA-II”. In: *IEEE Transactions on Evolutionary Computation* 6.2, pp. 182–197.

Deb, Kalyanmoy, Karthik Sindhya, and Jussi Hakanen (2016). “Multi-objective optimization”. In: *Decision sciences*. CRC Press, pp. 161–200.

Du, Yuanhao et al. (Dec. 2025). “Dynamic preventive maintenance model for offshore wind turbine bearings based on remaining useful life prediction”. In: *Sustainable Energy Technologies and Assessments* 84. DOI: 10.1016/j.seta.2025.104750.

Heng, R.B.W. and M.J.M. Nor (Jan. 1998). “Statistical Analysis of Sound and Vibration Signals for Monitoring Rolling Element Bearing Condition”. In: *Applied Acoustics* 53.1-3, pp. 211–226. DOI: 10.1016/S0003-682X(97)00018-2.

Ji, Daan et al. (Jan. 2021). “A Review: Data Driven-Based Fault Diagnosis and RUL Prediction of Petroleum Machinery and Equipment”. In: *Systems Science & Control Engineering* 9.1, pp. 724–747. DOI: 10.1080/21642583.2021.1992684.

Jin, Yulan, Zuhua Jiang, and Wenrui Hou (Nov. 2008). “Multi-objective integrated optimization research on preventive maintenance planning and production scheduling for a single machine”. In: *International Journal of Advanced Manufacturing Technology* 39.9-10, pp. 954–964. DOI: 10.1007/s00170-007-1268-5.

Kumar, Deepak, Sahil Kalra, and Mayank Shekhar Jha (Dec. 2022). “A Concise Review on Degradation of Gun Barrels and Its Health Monitoring Techniques”. In: *Engineering Failure Analysis* 142, p. 106791. DOI: 10.1016/j.engfailanal.2022.106791.

Lei, Jianxin et al. (Dec. 2022). “A Review: Prediction Method for the Remaining Useful Life of the Mechanical System”. In: *Journal of Failure Analysis and Prevention* 22.6, pp. 2119–2137. DOI: 10.1007/s11668-022-01532-4.

Lipu, M.S. Hossain et al. (Dec. 2018). "A Review of State of Health and Remaining Useful Life Estimation Methods for Lithium-Ion Battery in Electric Vehicles: Challenges and Recommendations". In: *Journal of Cleaner Production* 205, pp. 115–133. DOI: 10.1016/j.jclepro.2018.09.065.

Olde Keizer, Minou C. A., Simme Douwe P. Flapper, and Ruud H. Teunter (Sept. 2017). "Condition-Based Maintenance Policies for Systems with Multiple Dependent Components: A Review". In: *European Journal of Operational Research* 261.2, pp. 405–420. DOI: 10.1016/j.ejor.2017.02.044.

Rahimi, Iman et al. (2022). "Scheduling by NSGA-II: Review and Bibliometric Analysis". In: *Processes* 10.1, p. 98. DOI: 10.3390/pr10010098.

Rezamand, Milad et al. (Dec. 2020). "Critical Wind Turbine Components Prognostics: A Comprehensive Review". In: *IEEE Transactions on Instrumentation and Measurement* 69.12, pp. 9306–9328. DOI: 10.1109/TIM.2020.3030165.

Saxena, Abhinav et al. (Oct. 2008). "Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation". In: *2008 International Conference on Prognostics and Health Management*. Denver, CO, USA: IEEE, pp. 1–9. DOI: 10.1109/PHM.2008.4711414.

Wang, Kui, Chao Deng, and Lili Ding (Sept. 2020). "Optimal Condition-Based Maintenance Strategy for Multi-Component Systems under Degradation Failures". In: *Energies* 13.17. DOI: 10.3390/en13174346.

Wang, Lubing, Zhengbo Zhu, and Xufeng Zhao (2024). "Dynamic predictive maintenance strategy for system remaining useful life prediction via deep learning ensemble method". In: *Reliability Engineering & System Safety* 245.

Wang, Zongyao et al. (Oct. 2025). "A Multi-Objective Optimization Model for Predictive Opportunistic Maintenance of Lithium-Ion Batteries". In: *Journal of Energy Storage* 132, p. 117902. DOI: 10.1016/j.est.2025.117902.

Wu, Tianyi et al. (2020). "Dynamic maintenance strategy with iteratively updated group information". In: *Reliability Engineering System Safety* 197, p. 106820. ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2020.106820. URL: https://www.sciencedirect.com/science/article/pii/S0951832019309354.

Zhang, Qingfu and Hui Li (2007). "MOEA/D: A multiobjective evolutionary algorithm based on decomposition". In: *IEEE Transactions on Evolutionary Computation* 11.6, pp. 712–731.

Zhou, Yang et al. (Aug. 2022). "Tool Wear Mechanism, Monitoring and Remaining Useful Life (RUL) Technology Based on Big Data: A Review". In: *SN Applied Sciences* 4.8, p. 232. DOI: 10.1007/s42452-022-05114-9.

APPENDIX A: COMPONENT MAINTENANCE PLAN

Interval Name	Engine hours	Kilometres	Time	Maintenance time (h)
MH00	750			2
MH01	1.500			4
MH02	3.000			8
MH03	4.500			16
MH04	6.000			16
MH05	9.000			24
MH06	18.000			40
MH07	36.000			640
MH08	72.000			720
MM01		30.000		2
MM02		60.000		4
MM03		360.000		8
MM04		720.000		240
MM05		1.440.000		280
MT00			0,25 y	2
MT01			0,5 y	4
MT02			1 y	8
MT03			5 y	8
MT04			6 y	16
MT05			8 y	480
MT06			10 y	40
MTSW			4 y	16
BR 1			1 y	8
BR 2			4 y	8
BR 3			8 y	160

Table 2. Component standard RUL times as defined by the manufacturer and showing the name of each interval as well as its intended inspection frequency in engine hours, kilometres or time. Maintenance is due when the first threshold is reached. MH is engine-hour-based maintenance on the engine; MM is mileage-based maintenance on the running gear; MT is time-based maintenance on the vehicle body. MTSW and BR1, BR2 and BR3 are legally obligated maintenance intervals of the brake system and on the bonding connection of the side panels.

APPENDIX B: PSEUDO-CODE

Algorithm 1 NSGA-II for Component-Based Maintenance Scheduling

Require: dataset path/worksheet; config cfg (horizon, usage profile, costs, GA parameters); optional includes map
Ensure: Pareto set $\mathcal{P}^\star$ of trade-off schedules

$df \leftarrow$ LOADDATASET(cfg)
$(km_{\mathrm{cum}}, eh_{\mathrm{cum}}) \leftarrow$ BUILDUSAGEPROFILES(cfg)
$\mathcal{T} \leftarrow$ EXPANDOCCURRENCES(df, cfg, $km_{\mathrm{cum}}, eh_{\mathrm{cum}}$)
includes_map $\leftarrow$ APPLYINCLUDES($\mathcal{T}$, df, cfg)
if $|\mathcal{T}| = 0$ **then**
    return $\emptyset$
Initialize population $\mathcal{P}_0$ of size P by sampling $d_i \sim \mathcal{U}[L_i, U_i]$ for flexible tasks and fixing legal tasks at d_i^{due}
for each $\mathbf{d} \in \mathcal{P}_0$ **do**
    $\mathbf{d} \leftarrow$ REPAIRSCHEDULE($\mathbf{d}$, $\mathcal{T}$, cfg, includes_map)
    evaluate $\mathbf{f}(\mathbf{d}) \leftarrow$ COMPUTEOBJECTIVES($\mathbf{d}$, $\mathcal{T}$, cfg, includes_map, $km_{\mathrm{cum}}, eh_{\mathrm{cum}}$)
$\mathcal{P} \leftarrow \mathcal{P}_0$
for $g = 1$ to G **do**
    $\mathcal{O} \leftarrow \emptyset$
    while $|\mathcal{O}| < P$ **do**
        Select parents $\mathbf{d}^A, \mathbf{d}^B$ by tournament selection using non-dominated rank and crowding distance
        $(\mathbf{c}^1, \mathbf{c}^2) \leftarrow$ SINGLEPOINTCROSSOVER($\mathbf{d}^A, \mathbf{d}^B; p_{\mathrm{cx}}$)
        $\mathbf{c}^1 \leftarrow$ MUTATE($\mathbf{c}^1; p_{\mathrm{mut}}$)
        $\mathbf{c}^2 \leftarrow$ MUTATE($\mathbf{c}^2; p_{\mathrm{mut}}$)
        $\mathbf{c}^1 \leftarrow$ REPAIRSCHEDULE($\mathbf{c}^1$, $\mathcal{T}$, cfg, includes_map)
        $\mathbf{c}^2 \leftarrow$ REPAIRSCHEDULE($\mathbf{c}^2$, $\mathcal{T}$, cfg, includes_map)
        append $\mathbf{c}^1$ (and $\mathbf{c}^2$ if room) to $\mathcal{O}$
    $\mathcal{U} \leftarrow \mathcal{P} \cup \mathcal{O}$
    for each $\mathbf{d} \in \mathcal{U}$ **do**
        evaluate $\mathbf{f}(\mathbf{d})$
    Partition $\mathcal{U}$ into fronts $F_1, F_2, \ldots$ by fast non-dominated sorting
    $\mathcal{P} \leftarrow \emptyset$
    for $k = 1, 2, \ldots$ **do**
        if $|\mathcal{P}| + |F_k| \le P$ **then**
            $\mathcal{P} \leftarrow \mathcal{P} \cup F_k$
        else
            rank F_k by crowding distance
            fill $\mathcal{P}$ to size P
            break
$\mathcal{P}^\star \leftarrow F_1$ of the final population
Export selected Pareto schedules and summary tables
return $\mathcal{P}^\star$

APPENDIX C: PARETO VIEWS

Table 3. Full Pareto front obtained with $p_{\text{fail}} = 0.01$.

Sol.	Interv	Events	Cost (€)	Downtime	Availability	Reliability
1	73	142	1,119,000.00	214.84	0.9616	0.9900
2	80	142	1,154,000.00	228.84	0.9588	0.9900
3	74	142	1,121,500.00	214.81	0.9616	0.9901
4	81	142	1,156,500.00	228.84	0.9590	0.9900
5	77	142	1,136,000.00	222.84	0.9600	0.9900

Table 4. Full Pareto front obtained with $p_{\text{fail}} = 0.05$.

Sol.	Interv	Events	Cost (€)	Downtime	Availability	Reliability
1	78	142	1,144,500.00	237.20	0.9572	0.9500
2	75	142	1,125,000.00	228.20	0.9589	0.9500
3	75	142	1,124,500.00	227.18	0.9591	0.9501
4	76	142	1,135,500.00	232.20	0.9585	0.9500
5	71	142	1,107,000.00	220.86	0.9604	0.9512
6	74	142	1,119,000.00	227.13	0.9593	0.9503
7	73	142	1,118,000.00	227.18	0.9592	0.9501
8	74	142	1,120,500.00	227.18	0.9592	0.9501
9	72	142	1,107,000.00	222.12	0.9601	0.9503
10	74	142	1,123,500.00	227.15	0.9593	0.9502

Table 5. Full Pareto front obtained with $p_{\text{fail}} = 0.10$.

Sol.	Interv	Events	Cost (€)	Downtime	Availability	Reliability
1	83	142	1,169,000.00	260.40	0.9535	0.9000
2	74	142	1,120,500.00	241.40	0.9566	0.9000
3	69	142	1,096,000.00	232.39	0.9582	0.9000
4	69	142	1,093,000.00	232.37	0.9581	0.9001
5	74	142	1,122,500.00	242.40	0.9566	0.9000
6	66	142	1,077,500.00	224.92	0.9594	0.9017
7	70	142	1,102,500.00	234.40	0.9583	0.9000
8	74	142	1,121,000.00	242.40	0.9563	0.9000
9	69	142	1,095,500.00	231.37	0.9588	0.9001
10	67	142	1,084,500.00	228.14	0.9594	0.9009