Data and Metadata. 2025; 4:1102
doi: 10.56294/dm20251102

ORIGINAL

Reliable prediction of industrial components Remaining Useful Life using Cox and Weibull models: A Comparative Study

Predicción fiable de la vida útil remanente de componentes industriales utilizando los modelos de Cox y Weibull: un estudio comparativo

Hamdi Alaoui Abdelhafid[1], Anwar Meddaoui[1], Ahmed En-nhaili[1]

[1]ENSAM, Hassan II university, industrial engineering department, Casablanca, Morocco.

Cite as: Alaoui Abdelhafid H, Meddaoui A, En-nhaili A. Reliable prediction of industrial components Remaining Useful Life using Cox and Weibull models: A Comparative Study. Data and Metadata. 2025; 4:1102. https://doi.org/10.56294/dm20251102

Submitted: 22-03-2025 **Revised:** 01-07-2025 **Accepted:** 09-10-2025 **Published:** 10-10-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding author: Hamdi Alaoui Abdelhafid

ABSTRACT

Predicting the Remaining Useful Life (RUL) of industrial equipment is a cornerstone of predictive maintenance strategies aimed at minimizing downtime and optimizing maintenance costs. This study conducts a comparative evaluation of two prominent survival analysis techniques the Cox Proportional Hazards (Cox PH) model and the Weibull model for RUL prediction in industrial machinery. We analyzed the AI4I 2020 Predictive Maintenance Dataset using a robust analytical framework incorporating Kaplan-Meier survival curves, log-rank tests, and multivariate survival modeling. Our methodology involved detailed data preprocessing, model validation through the concordance index (C-index) and Akaike Information Criterion (AIC), and the identification of significant predictors of failure. The results demonstrated that the Cox PH model outperforms the Weibull model in terms of flexibility, predictive accuracy, and the ability to handle multiple covariates. This work highlights the strengths and limitations of both models and emphasizes the superior applicability of the Cox PH model for complex industrial datasets. These findings offer actionable insights for developing more reliable, data-driven maintenance strategies in Industry 4.0 environments.

Keywords: Remaining Useful Life (RUL); Predictive Maintenance; Reliability; Cox Proportional Hazards Model; Weibull Model; Survival Analysis.

RESUMEN

Predecir la Vida Útil Restante (RUL) del equipamiento industrial es un pilar fundamental de las estrategias de mantenimiento predictivo orientadas a minimizar el tiempo de inactividad y optimizar los costos de mantenimiento. Este estudio realiza una evaluación comparativa de dos técnicas destacadas de análisis de supervivencia: el modelo de Riesgos Proporcionales de Cox (Cox PH) y el modelo de Weibull, para la predicción del RUL en maquinaria industrial. Analizamos el conjunto de datos AI4I 2020 de Mantenimiento Predictivo utilizando un marco analítico sólido que incorpora curvas de supervivencia de Kaplan-Meier, pruebas log-rank y modelado de supervivencia multivariante. Nuestra metodología incluyó un detallado preprocesamiento de datos, validación de modelos mediante el índice de concordancia (C-index) y el Criterio de Información de Akaike (AIC), así como la identificación de predictores significativos de fallo. Los resultados demostraron que el modelo Cox PH supera al modelo de Weibull en términos de flexibilidad, precisión predictiva y capacidad para manejar múltiples covariables. Este trabajo destaca las fortalezas y limitaciones de ambos modelos y enfatiza la superior aplicabilidad del modelo Cox PH para conjuntos de datos industriales complejos. Estos hallazgos ofrecen perspectivas prácticas para desarrollar estrategias de mantenimiento más fiables y basadas en datos en entornos de la Industria 4.0.

Palabras clave: Vida Útil Restante (RUL); Mantenimiento Predictivo; Fiabilidad; Modelo de Riesgos Proporcionales de Cox; Modelo de Weibull; Análisis de Supervivencia.

INTRODUCTION

In modern industrial environments, predictive maintenance has become essential to ensure equipment reliability, reduce unexpected downtimes, and optimize operational efficiency.(1) A key component of this strategy is the accurate estimation of the Remaining Useful Life (RUL) of machines. By anticipating when a failure is likely to occur, industries can plan timely interventions, thus avoiding unnecessary repairs and minimizing production losses.(2)

Survival analysis offers a powerful statistical framework for modeling time-to-event data, especially when equipment failures do not occur at fixed intervals and observations may be censored.(3) Among the most widely used survival models are the Cox Proportional Hazards (Cox PH) model and the Weibull model. These two approaches differ in assumptions and applicability: the Cox PH model is semi-parametric and capable of incorporating multiple covariates without assuming a specific baseline hazard,(4) while the Weibull model is fully parametric and provides a predefined structure for the failure rate over time.(5)

The prediction of Remaining Useful Life (RUL) employs various approaches, including physics-based, statistical, data-driven, and hybrid models. Physics-based models, though highly interpretable, require expert knowledge and are limited to specific failure mechanisms. Statistical methods like the Cox Proportional Hazards model and the Weibull distribution are widely used in reliability engineering for handling censored data and estimating failure probabilities under diverse conditions, as noted by Liu et al.(6) and Gao et al.(7) Meanwhile, data-driven techniques, increasingly favored due to abundant sensor data and computational power, encompass machine learning and deep learning models, offering high accuracy for complex, high-dimensional datasets despite lower interpretability.

This study aims to apply both models to the AI4I 2020 Predictive Maintenance Dataset, which contains real-world-inspired sensor data and machine failures. Our goal is to compare their effectiveness in predicting RUL, interpreting key predictors, and validating the results through established statistical metrics.

The main contributions of this paper are as follows:

We present a comparative study of the Cox PH and Weibull models for RUL prediction in an industrial context. We integrate Kaplan-Meier (KM) survival curves and the log-rank test to assess group-based survival behavior.(8) We include a detailed data preprocessing pipeline and a robust validation framework using the Concordance Index (C-index) and the Akaike Information Criterion (AIC).(9) We identify significant covariates affecting machine failure and provide insights for maintenance strategy planning.(10) The remainder of the paper is structured as follows: Section 2 provides a review of relevant literature; Section 3 describes the theoretical background of survival models; Section 4 details the methodology including data preprocessing, feature selection, and validation; Section 5 presents the results and model comparisons; Section 6 discusses the findings; and Section 7 concludes the study and outlines future directions.

LITERATURE REVIEW

Recent research by (6,11) has demonstrated the integration of deep learning with survival models for improved RUL estimation. Xu et al. proposed a federated learning architecture combining adaptive sampling and ensemble strategies for aircraft engine RUL prediction, offering both data privacy and generalization performance. Similarly, Zhang et al.(12) presented a Wiener process-based degradation model enhanced with external condition variables, showing that hybrid models can capture both observable and latent factors affecting degradation.

Hybrid approaches aim to combine the strengths of physics-based and data-driven methods. For instance, Zhang et al.(13) developed a hybrid framework combining Weibull-based degradation modeling with machine learning, balancing interpretability and predictive power. Such models are especially useful in industrial settings where partial knowledge of the failure mechanism exists alongside rich sensor data.

Reliability-centered maintenance (RCM) approaches also emphasize the strategic integration of reliability principles into predictive maintenance. As Smith et al.(14) suggested, embedding RCM in predictive frameworks leads to significant improvements in uptime and resource allocation.

Despite the increasing interest in AI-based methods, traditional statistical models like Cox PH remain highly relevant. They provide interpretable hazard ratios, allow the inclusion of multiple covariates,(15) and perform well in scenarios with censored data an aspect often overlooked by black-box machine learning models.

In summary, although deep learning models are advancing rapidly, survival analysis models remain essential tools thanks to their interpretability and proven reliability in industrial applications. Nevertheless, there is a notable lack of studies directly comparing the Cox Proportional Hazards (Cox PH) model and the Weibull model on real-world industrial data, particularly those that simultaneously address model validation, the influence of predictive features, and the interpretability of failure mechanisms. To address this gap, the present study aims to conduct a comprehensive and rigorous evaluation of both models using the AI4I 2020 Predictive Maintenance Dataset, with the general objective of determining their relative strengths and suitability for predicting the Remaining Useful Life (RUL) of industrial equipment.

THEORICAL BACKGROUND

Cox Proportional Hazards Model

The Cox Proportional Hazards model is a semi-parametric method that estimates how covariates affect the hazard of failure without specifying the baseline hazard function. Its hazard function is:

$$h(t|x) = h_0(t) * exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)$$

where h0(t) is the baseline hazard and exp(β) gives hazard ratios, offering interpretable measures of risk

Weibull Model

The Weibull model is a parametric survival model used in reliability engineering, assuming failure times follow a Weibull distribution. Its density function is:

$$f(t;\lambda,\beta) = \frac{\beta}{\lambda}\left(\frac{t}{\lambda}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\lambda}\right)^{\beta}\right]$$

The shape parameter β\betaβ indicates decreasing (<1), constant (=1), or increasing (>1) hazard rates.

Kaplan-Meier Estimator

The Kaplan-Meier estimator is a non-parametric method for estimating survival probabilities from time-to-event data, especially useful with censored observations. The survival function is calculated as:

$$S(t) = \prod_{ti \leq t} \left(1 - \frac{di}{n_i}\right)$$

where is the number of failures at time and the number at risk. It's widely used for visualizing survival curves.

Log-Rank Test

The log-rank test is a non-parametric method for comparing survival curves between groups, testing the null hypothesis of no difference in survival. The statistic is:

$$LOG - Rank\ Statistic = \frac{\sum_i (O_i - E_i)^2}{\sum_I V_i}$$

Where Oi and Ei are the observed and expected events in group i, and Vi is the variance. A significant result indicates differences in survival distributions

METHOD

Dataset Overview

This study is an applied, quantitative research project based on secondary data analysis. Applied research focuses on solving practical problems using scientific methods, while quantitative research relies on numerical data and statistical techniques. The study employs survival analysis, a statistical approach used to examine the time until the occurrence of a specific event, in this case, machine failure. In this context, the variable (Tool wear [min]) serves as the time variable, representing the duration until failure, and the binary variable "Machine failure" acts as the event indicator, identifying whether a failure has occurred (1) or not (0). The research utilizes the AI4I 2020 Predictive Maintenance Dataset, which contains 10,000 records and 14 variables describing operational conditions of industrial equipment, including air and process temperatures, rotational speed, torque, tool wear, machine failure events, and failure types. Additionally, machines are categorized into three types L, M, and H enabling subgroup analysis to explore potential differences in failure behavior across

https://doi.org/10.56294/dm20251102 ISSN: 2953-4917

machine types.

Key Parameters

The AI4I 2020 Predictive Maintenance Dataset includes several critical variables essential for modeling machine failures and predicting Remaining Useful Life (RUL). The most relevant concepts and parameters for this study are summarized below:

- UDI (Unique Identifier): A unique code assigned to each data record, used for accurate referencing and tracking within the dataset.
- Product ID and Type: Identify the specific machine type or product category being manufactured. These variables enable differentiation between machines and analysis of failure patterns across different product lines.
- Air Temperature [K]: The ambient temperature surrounding the machine, measured in Kelvin. Environmental conditions like temperature can significantly influence machine performance and the likelihood of failure.
- Process Temperature [K]: Reflects the internal temperature during machine operation. Elevated process temperatures may signal abnormal conditions or emerging faults.
- Rotational Speed [rpm]: Indicates how fast machine components rotate, measured in revolutions per minute. Higher speeds generally accelerate wear and increase the risk of mechanical failures.
- Torque [Nm]: Measures the twisting force applied to machine components in Newton-meters. Excessive torque can cause mechanical stress, leading to premature wear or failure.
- Tool Wear [min]: Represents the cumulative operational time of cutting or machining tools, measured in minutes. Tool wear is a crucial indicator of degradation and serves as the time variable in survival analysis for RUL prediction.
- Machine Failure: A binary variable where 1 indicates the occurrence of a failure event, and 0 indicates normal operation. This serves as the event indicator in survival analysis models and is the primary target variable for predictive maintenance applications.

Failure Types: Categorize the specific nature of failures observed in the machines:

- Heat Dissipation Failure (HDF): Caused by inadequate heat removal, leading to overheating and potential damage to components like motors or bearings.
- Power Failure (PF): Results from fluctuations or interruptions in the power supply, potentially causing system shutdowns or damage to electrical parts.
- Overstrain Failure (OF): Occurs when machines operate beyond their designed capacity, resulting in excessive stress and potential breakdowns.
- Random Failure (RF): Unpredictable failures with no clear root cause, possibly linked to material defects, environmental changes, or operational errors.
- Tool Wear Failure (TWF): Progressive deterioration of tools due to friction, heat, and repeated use, ultimately leading to functional failures.

These parameters are essential for understanding the operational conditions, degradation processes, and failure mechanisms of the industrial equipment.

Data Preprocessing

A robust data preprocessing pipeline was implemented to ensure data quality:

- Missing Values: The dataset had no missing entries. However, we verified consistency by scanning for null values and applying imputation techniques when necessary.
- Outlier Detection: We employed interquartile range (IQR)-based filtering and visual methods (boxplots and scatter plots) to identify and treat outliers, especially in sensor readings (e.g., extreme torque or speed values).
- Censoring: As is typical in industrial settings, some observations were right-censored—i.e., machines that had not yet failed at the end of the observation period. These were appropriately handled by both the Cox and Weibull models.
- Feature Encoding: The categorical variable 'Type' (L, M, H) was encoded using one-hot encoding for compatibility with survival models.
- Normalization: Continuous variables (temperatures, torque, speed) were scaled using standard normalization to ensure numerical stability during model fitting.

Feature Selection and Multicollinearity

To avoid overfitting and reduce redundancy, we performed feature selection using the following steps:

- Correlation Matrix: Pearson correlation analysis was applied to assess collinearity between numeric variables.
- Variance Inflation Factor (VIF): VIF scores were calculated, and features with VIF > 5 were evaluated for potential exclusion.
- Domain Knowledge: Variables such as air temperature, torque, and tool wear were retained based on their known influence on mechanical degradation.

Model Development

Two survival analysis models were developed and trained:

- Cox Proportional Hazards (Cox PH): A semi-parametric model that estimates the impact of covariates on the hazard rate. Model fitting was done using partial likelihood estimation.
- Weibull Model: A fully parametric model where scale (λ) and shape (β) parameters were estimated using maximum likelihood estimation (MLE).
- We also used Kaplan-Meier estimators to generate survival curves for each machine type and the log-rank test to assess statistical differences between group survival distributions.

A pair plot analysis confirmed expected physical behavior: torque and rotational speed are inversely related, and tool wear appears evenly distributed.

Figure 1 present the heatmap of the correlation matrix visualizes the relationships between dataset features, with correlations ranging from -1 to 1. Key observations include: Strong Positive Correlation: A 0,88 correlation between Air temperature and Process temperature suggests they increase together, indicating environmental impact on machinery.

Strong Negative Correlation: A -0,88 correlation between Rotational speed and Torque reflects the principle that higher speeds require lower torque, confirming expected mechanical behavior.

Moderate Positive Correlations with Machine Failure: Machine failure shows moderate correlations with Heat Dissipation Failure (0,58) and Power Failure (0,52), highlighting these as significant contributors to overall failure.

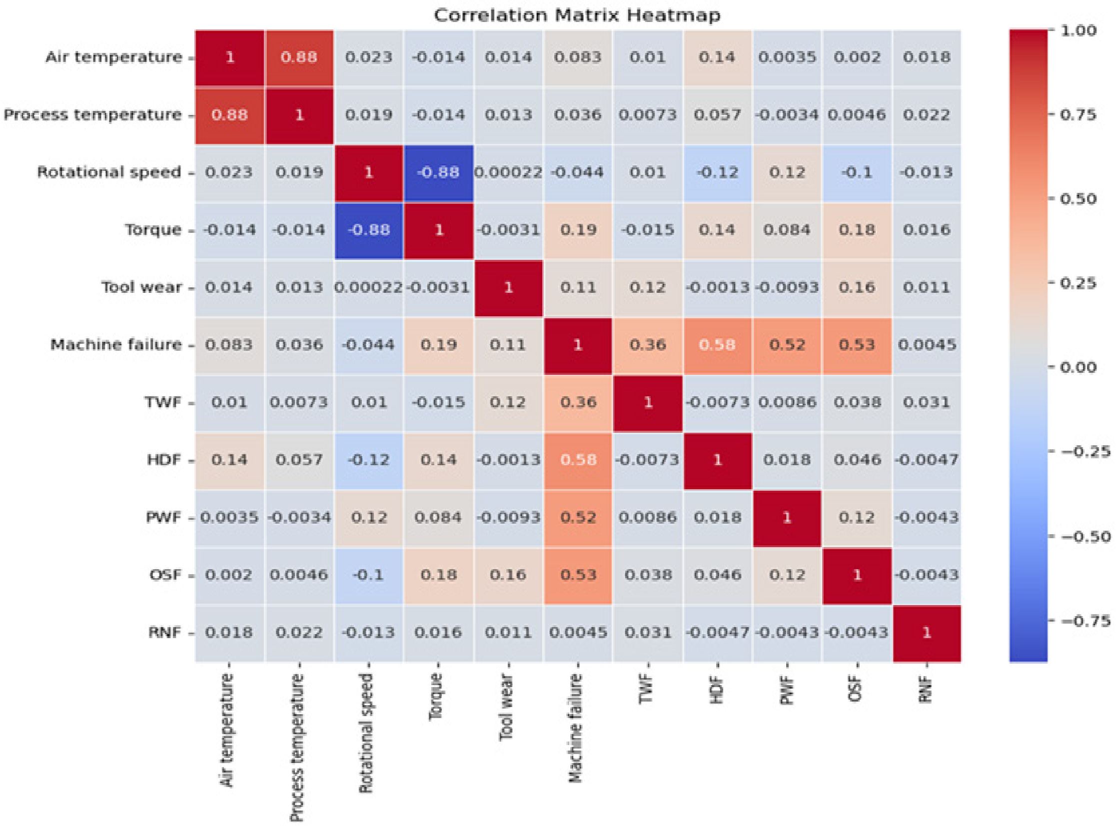

Figure 1. Correlation Matrix Heatmap

Workflow charter

The complete RUL prediction pipeline is shown in figure 2. It starts with data acquisition, followed by preprocessing, exploratory analysis, feature selection, and parallel modeling using Cox PH and Weibull.

https://doi.org/10.56294/dm20251102 ISSN: 2953-4917

Validation and comparison metrics guide model interpretation and deployment.

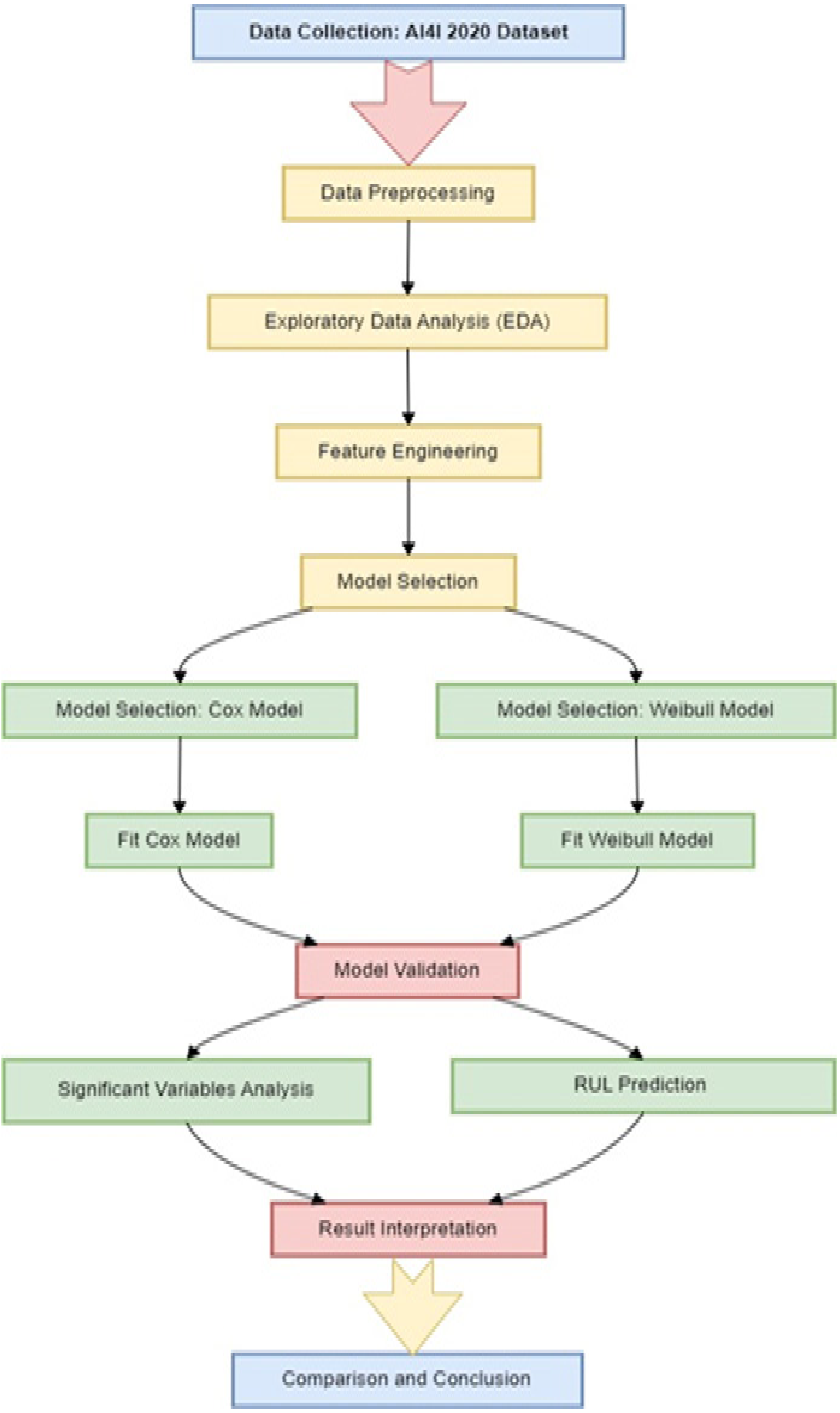

Figure 2. prediction methodology workflow chart

Model Validation Strategy

To evaluate model robustness and generalizability, we employed a stratified 5-fold cross-validation strategy on the training data. Performance was assessed using:

- Concordance Index (C-index): To measure the rank correlation between predicted and observed survival times.
- Akaike Information Criterion (AIC): To compare model fit quality.
- Residual Plots and Schoenfeld Tests: To assess the proportional hazards assumption for the Cox model.

The dataset was split into 70 % training and 30 % testing sets, maintaining the event distribution across sets. Validation metrics were computed on both training and testing partitions to assess overfitting.

Exploratory Data Analysis (EDA)

An initial correlation matrix heatmap revealed strong dependencies between variables:

- Air temperature and Process temperature: strong positive correlation (ρ = 0,88)
- Rotational speed and Torque: strong negative correlation (ρ = -0,88)
- Machine failure moderately correlated with Heat Dissipation Failure (ρ = 0,58) and Power Failure (ρ = 0,52)

RESULTS

The Kaplan-Meier estimator was used to generate survival curves for each machine type. Next, the log-rank test was applied to compare survival distributions between the machine types and to determine whether significant differences in failure rates actually exist [16,17] Both Cox PH and Weibull models were fitted to the data. Covariates in the Cox PH model included air and process temperature, rotational speed, torque, and tool wear, thus allowing estimation of how the effect of these variables influences the hazard for machine failure.[18]

For the Weibull model, shape and scale parameters are determined by the method of maximum likelihood. The process enables the model to account for different failure behaviors of the Machinery. The process may have some deviation from the normal procedures, which may be related to the scope of the research.

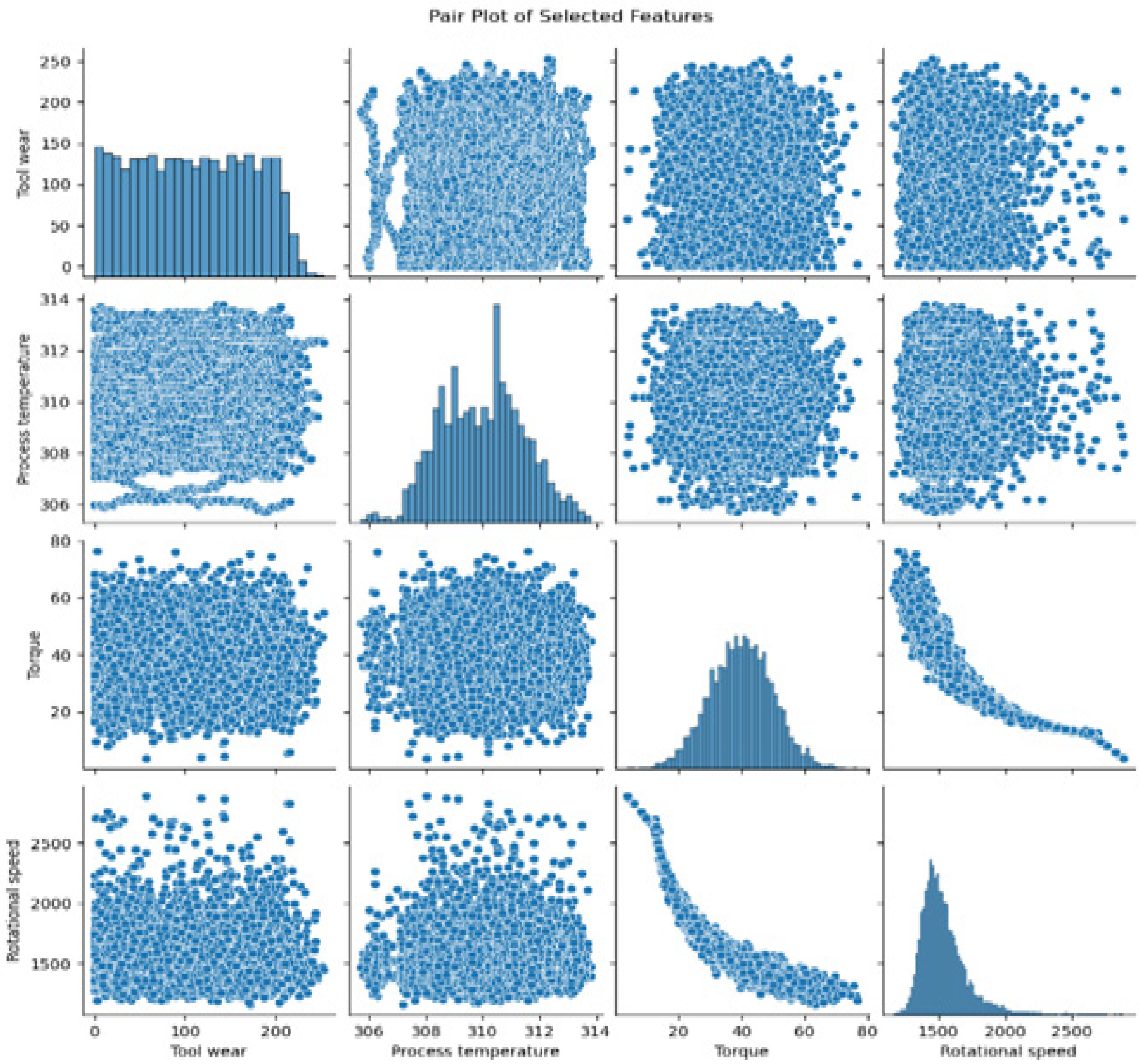

Figure 3. Pair plot of selected Features

Figure 3 present the pair plot reveals the distributions and relationships between Tool wear, Process temperature, Torque, and Rotational speed. Histograms show that Tool wear is uniformly distributed and

https://doi.org/10.56294/dm20251102 ISSN: 2953-4917

Process temperature has a normal distribution centered around 310°C, indicating stable operations.

A strong inverse relationship exists between Torque and Rotational speed, reflecting fundamental mechanical principles. However, Tool wear lacks strong linear relationships with other variables, suggesting it may be influenced by complex interactions. These insights highlight the need for advanced modeling techniques like Cox and Weibull models to capture non-linear effects and improve RUL predictions.

Torque and Rotational Speed: Torque displays a bell-shaped distribution, while Rotational speed shows a peak around 1500 RPM, indicating that most machines operate within specific, optimized ranges for these variables.

Kaplan-Meier Survival Curves

Kaplan-Meier curves were generated for the three machine types (L, M, H). Key findings:

- Type L showed the highest survival probability throughout the observed tool wear duration.
- Type M experienced the fastest degradation, with the steepest drop in survival probability.
- Type H had intermediate performance.

As Tool Wear exceeds 200 minutes, both the Kaplan-Meier survival curves and the Cumulative Distribution Function (CDF) curves figure 4 reveal critical differences in the failure behaviors of the different component types (M, L, and H).

The Kaplan-Meier curves show that after 200 minutes, the survival probabilities start to diverge, with Type L showing the highest likelihood of survival, followed by Type H, and then Type M, which declines more rapidly. Simultaneously, the CDF curves indicate that Type L components experience a steep rise in failure probability after 200 minutes, suggesting that these components are more prone to fail earlier compared to Types M and H, with Type H showing the slowest accumulation of failures. By 250 minutes, both sets of curves converge towards a survival probability near zero and a CDF probability of one, indicating that nearly all components have failed by this point regardless of type, marking the end-of-life for these components.

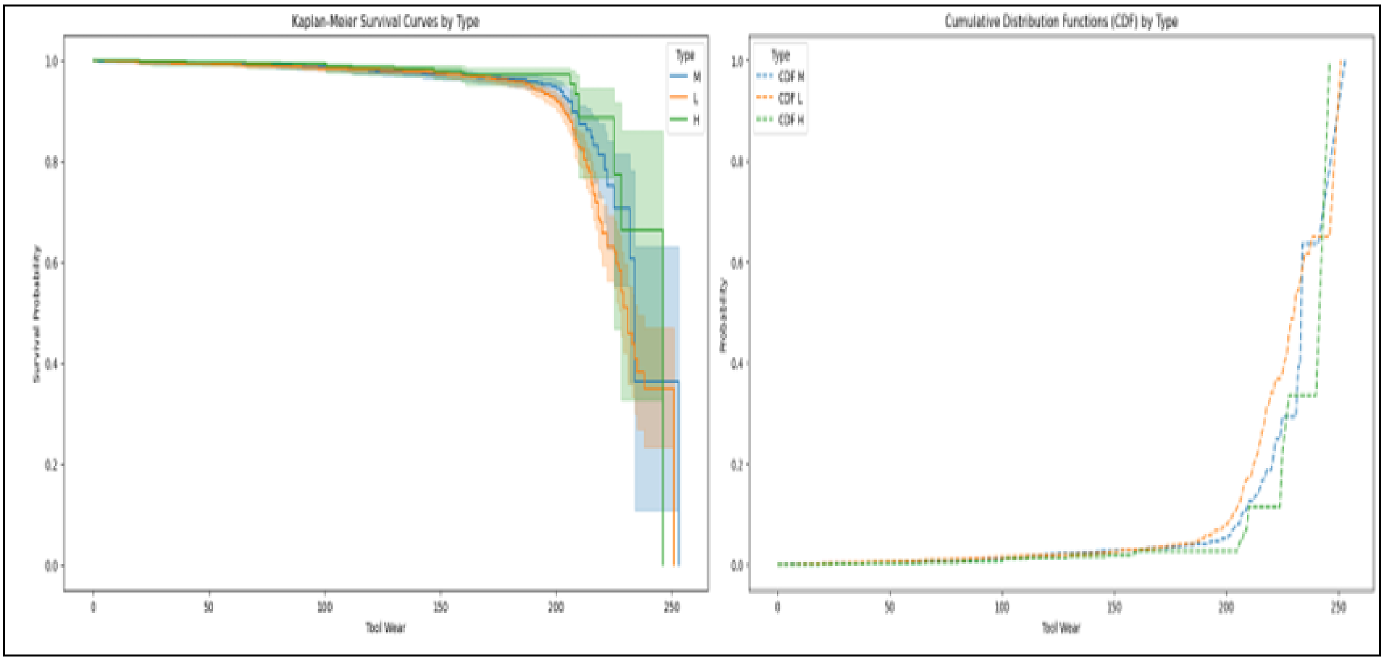

Figure 4. Kaplan Meier Survival and Cumulative Distribution Function curves by type

Log-Rank Test

The Log-Rank test was performed to compare the survival distributions of machine components under three different operational conditions, labeled as Types M, L, and H. The results revealed varying degrees of statistical significance between the groups.

When comparing Type M and Type L, the p-value was 0,0501, which is just above the conventional threshold of 0,05.

This suggests that there is no statistically significant difference in the survival distributions between these two types, although the result is borderline and may warrant further investigation.

The comparison between Type M and Type H yielded a p-value of 0,1046, indicating no significant difference in their survival times, implying that components under these two conditions exhibit similar failure behaviors.

However, the comparison between Type L and Type H showed a p-value of 0,0057, which is well below 0,05, indicating a statistically significant difference in their survival distributions.

This suggests that the operational conditions or inherent characteristics of Type L and Type H components lead to different failure rates, with Type L and Type H components experiencing significantly different patterns of reliability and failure.

Table 1. P-values for Pairwise Comparisons Between Machine Types Using the Log-Rank Test

Comparison	P-value	Interpretation
Type M vs. Type L	0,0501	No statistically significant difference (p-value ≈ 0,05)
Type M vs. Type H	0,1046	No statistically significant difference
Type L vs. Type H	0,0057	Statistically significant difference

Weibull Model Results

The Weibull model is fitted using the time-to-event data, which estimates the scale (λ) and shape (β) parameters of the Weibull distribution

The survival function plot using the Weibull model provides important insights into the likelihood that a machine component will continue to function over time as it accumulates tool wear

The survival curve estimated by the Weibull model, applied to the tool wear data (Tool Wear), reveals an initially high survival probability that decreases exponentially as wear increases. This behavior is typical of the Weibull model, where the scale parameter λ is estimated at 639,09 minutes, indicating a typical lifespan before a significant failure occurs. The confidence interval for λ, though not explicitly mentioned here, likely reflects robustness in the estimation, with very high statistical significance ($p < 0,001$).

The curve presented in the figure 5 illustrates the survival function estimated by the Weibull model, applied to the tool wear data (Tool Wear), expressed in minutes. The survival function shows an initial probability close to 1, which gradually decreases as tool wear increases.

This characteristic decline in the Weibull model reflects an increasing risk of failure as wear accumulates. The shape parameter β, estimated at 2,09 suggests a rising failure rate, indicating that as tools wear down, the risk of failure increases rapidly.

The confidence intervals, represented by the shaded areas around the curve, reflect the uncertainty of the estimates. In the early stages of wear, these intervals are relatively narrow, indicating greater certainty in survival estimation. However, as wear reaches higher levels, the confidence intervals widen, reflecting growing uncertainty, likely due to the scarcity of data for high wear values.

Although the Weibull model is known for its flexibility in modeling increasing, decreasing, or constant failure rates, the results obtained here suggest that, while it captures the general trend of component survival, it may under or overestimate the survival probability in certain specific wear ranges. The median survival time, estimated at around 536,29 minutes, provides a clear indication of when 50 % of the components are likely to fail. This observation highlights the need to consider the model's limitations, particularly for long-term forecasts, and potentially compare it to non-parametric methods like Kaplan-Meier to ensure the robustness of conclusions. These results offer a deep understanding of component survival behavior but also call for caution when interpreting survival predictions for high wear values.

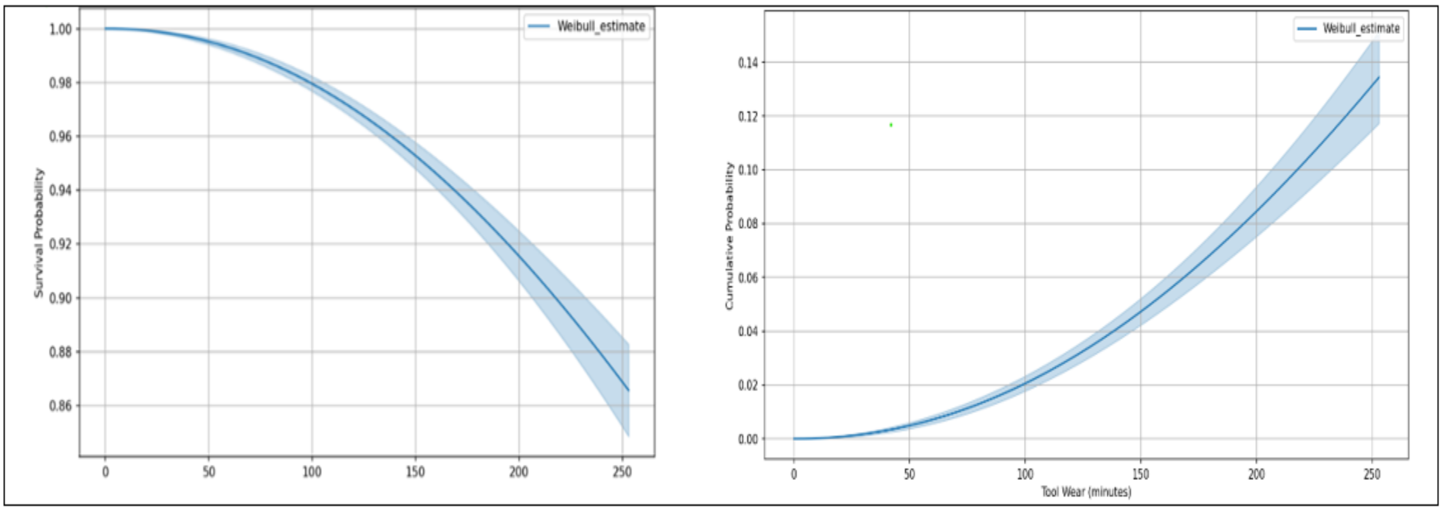

Figure 5. Survival Function and CDF using Weibull Model

https://doi.org/10.56294/dm20251102 ISSN: 2953-4917

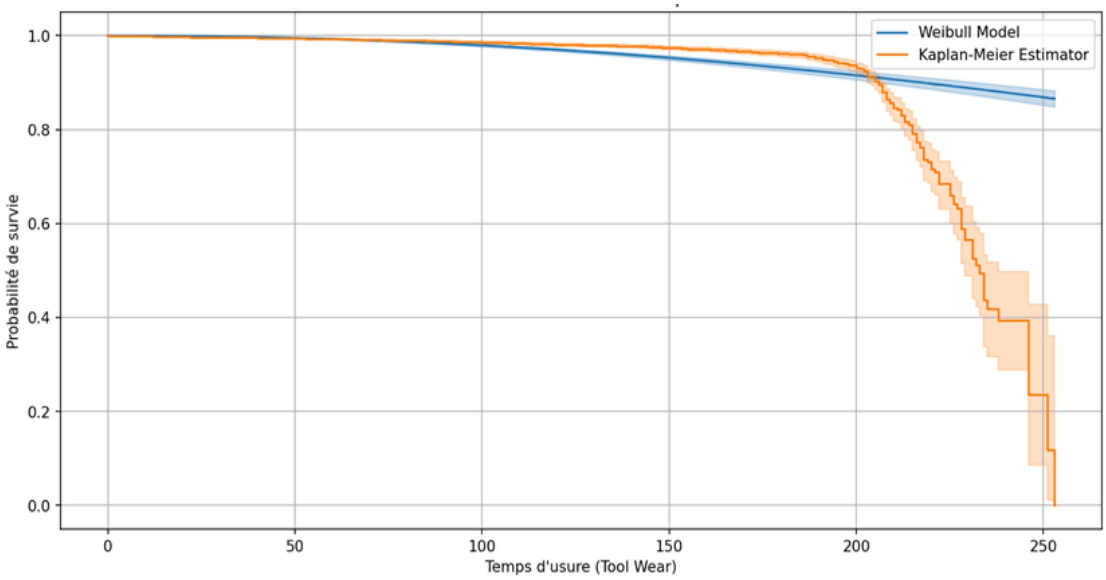

Figure 6. Weibull Survival Curve Vs Kaplan Meier Curve

The figure 6 compares two different approaches to estimating tool lifespan: the Weibull model and the Kaplan-Meier estimator. The blue curve represents the Weibull model, which shows a gradual decrease in survival probability as tool wear increases. In contrast, the orange Kaplan-Meier curve steps down at various points, reflecting actual tool failures observed in the data.

Initially, both curves show similar high survival probabilities, indicating tools rarely fail early in their life. Beyond 200 minutes, the Kaplan-Meier curve drops sharply, reflecting a sudden rise in failure risk, while the Weibull curve declines more gradually, suggesting a slower increase in risk. The Kaplan-Meier confidence intervals widen after 200 minutes due to fewer tools remaining, indicating greater uncertainty, whereas the Weibull intervals remain narrower, driven by model assumptions rather than data. Overall, Kaplan-Meier closely tracks actual data but loses reliability over time, while the Weibull model offers smoother predictions that may oversimplify reality. Combining both methods can provide a more comprehensive understanding of tool lifespan and reliability.

Cox PH Model Results

Table 2. Cox PH Estimates: Coefficients, Hazard Ratios, and Statistical Significance

Covariate	Coefficient (log(HR))	Hazard Ratio (HR)	p-value	95 % Confidence Interval (log(HR))
Air Temp	0,4882	1,629	1,79e-17	(0,376- 0,601)
Process Temp	-0,4938	0,610	1,78e-10	(-0,645- -0,342)
Rot Speed	0,0080	1,008	6,32e-123	(0,007-0,009)
Torque	0,2000	1,221	2,24e-168	(0,186-0,214)
Type (L)	0,5315	1,701	2,07e-02	(0,081- 0,982)
Type (M)	0,2896	1,336	0,239	(-0,192- 0,771)

The Cox proportional hazards model results provide valuable insights into the factors influencing machine failure. Air temperature has a significant positive effect on failure, with a hazard ratio of 1,63, indicating that as air temperature increases, the likelihood of machine failure rises significantly ($p < 0,001$).

Process temperature, on the other hand, has a protective effect, as it is negatively associated with failure ($HR = 0,61$, $p < 0,001$), meaning higher process temperatures reduce the risk of failure. Rotational speed and torque are also statistically significant, though their effects differ. Rotational speed has a minor positive effect on failure ($HR = 1,008$, $p < 0,001$) while torque strongly increases the hazard of failure ($HR = 1,22$, $p < 0,001$).

Regarding machine types, Type L machines have a 70 % higher risk of failure compared to the reference group ($HR = 1,70$ | $p = 0,02$), while Type M machines show no significant impact ($p = 0,239$), suggesting that machine type M does not significantly influence failure risk.

The significant covariates in this model highlight the need for careful management of air temperature, process temperature, torque, and rotational speed to reduce machine failure risks

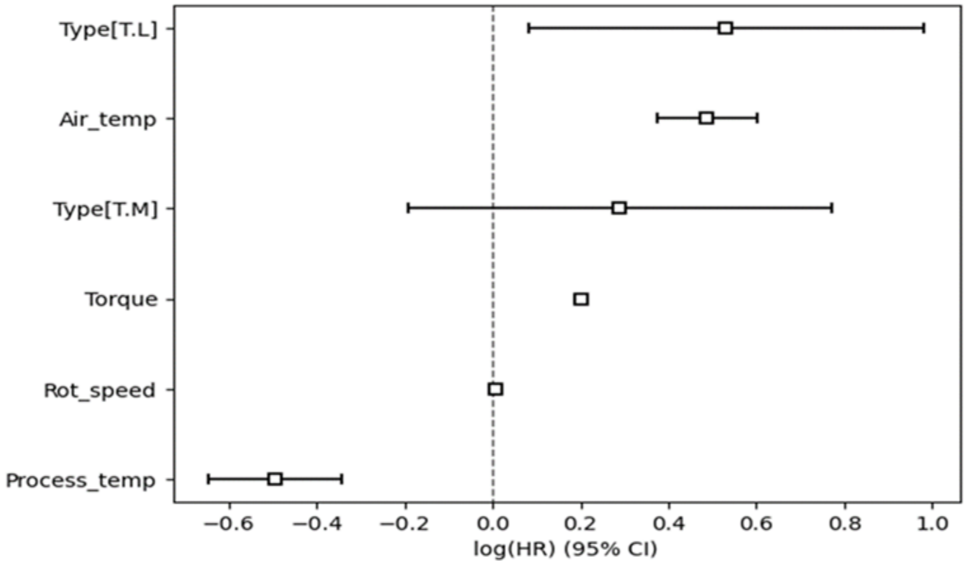

Figure 7. Coefficients du Modèle de Cox

The figure 7 illustrates the coefficients from the Cox Proportional Hazards Model, depicting the influence of various covariates on the likelihood of machine failure. Each point represents the log-hazard ratio (HR) for a given factor, with horizontal bars indicating 95 % confidence intervals.

Type L (Type[T.L]), it shows the largest positive effect on failure risk, with a significantly higher hazard compared to the baseline group (Type H), as indicated by its position far to the right of zero and a confidence interval that does not cross zero. This suggests that machines of Type L are much more likely to fail.

In contrast, Type M (Type [T.M]) shows no statistically significant effect on failure risk, with its confidence interval crossing zero, meaning it doesn't strongly affect machine longevity compared to Type H.

Air temperature (Air_temp) also has a notable positive impact on the risk of failure, with a narrow confidence interval demonstrating high confidence in the estimate. This indicates that as air temperature rises, so does the likelihood of machine failure. Torque shows a similar effect, with a positive coefficient and confidence interval that remains above zero, indicating that higher torque increases the risk of failure as well.

On the other hand, Process temperature (Process_temp) has a protective effect, as seen by its negative coefficient. The confidence interval doesn't cross zero, confirming that higher process temperatures significantly reduce the failure risk.

Finally, Rotational Speed (Rot_speed), although showing a very small positive effect on failure, remains statistically significant, suggesting that increased speed slightly increases the failure risk, though its influence is minor compared to the other variables.

The graph highlights that air temperature, torque, and process temperature are key drivers of machine failure, with air temperature and torque increasing the risk, while process temperature reduces it. Machine Type L is shown to be more prone to failure, whereas Type M appears to have no significant effect.

The Cox model identified the following covariates as statistically significant:

Table 3. Covariates statistically significant

Covariate	Hazard Ratio (HR)	p-value	Interpretation
Air Temperature	1,63	< 0,001	Higher air temp increases failure risk
Process Temp	0,61	< 0,001	Protective effect
Rot. Speed	1,008	< 0,001	Small positive effect
Torque	1,22	< 0,001	Strong positive effect
Type L	1,70	0,0207	Higher risk compared to Type H
Type M	1,34	0,239	Not significant

Model Validation and Performance Indicators

To compare the performance of the Cox Proportional Hazards model and the Weibull model for predicting the Remaining Useful Life (RUL), we implemented a rigorous evaluation approach based on three key indicators: the Concordance Index (C-index), the Akaike Information Criterion (AIC), and the median predicted RUL. These metrics assess both the accuracy, robustness, and goodness-of-fit of the models.

https://doi.org/10.56294/dm20251102 ISSN: 2953-4917

C-index Calculation:

The C-index measures the model's ability to correctly rank pairs of observations based on their predicted failure times. A C-index of 1 denotes perfect concordance, while a value of 0,5 corresponds to random ranking.

The C-index was computed using the concordance_index function from the Python lifelines library.

Right-censored pairs were appropriately accounted for in the evaluation.

C-index of 0,9148 for the Cox model (excellent predictive ability), compared to 0,639 for the Weibull model, indicating moderate performance.

AIC Calculation:

The Akaike Information Criterion (AIC) evaluates the trade-off between model fit and complexity, based on the maximized log-likelihood of the model. For the Cox model, the AIC was derived using the Breslow approximation to handle censored observations. For the Weibull model, AIC was calculated based on maximum likelihood estimates under the assumption of a fixed parametric form.

Median RUL Estimation:

The median RUL corresponds to the time at which the survival probability reaches 0,5.

For the Weibull model, the median RUL is constant and computed using the scale (λ) and shape (β) parameters (Median RUL= $\lambda.(\ln(2))^{(1/\beta)}$),this yields a fixed median RUL of 500,03 minutes. For the Cox model, due to its semi-parametric nature, the RUL estimation is individualized based on the covariates of each observation. Personalized survival functions were generated to predict the RUL for each unit.

Table 4. Performance Summary Table

Metric	Cox Model	Weibull Model
C-index	0,9148 (Excellent)	0,639 (Moderate)
AIC	4184,44	5915,87
Median RUL (minutes)	Variable (individualized)	500,03 (fixed)
Model Type	Semi-parametric	Parametric
Handles Covariates	Yes	Limited

DISCUSSION

Interpretability and Feature Significance

The comparison between the Cox Proportional Hazards (PH) model and the Weibull model using the AI4I 2020 dataset highlights their distinct strengths and limitations, consistent with prior studies in predictive maintenance contexts.[19,20] The Cox PH model excels at identifying significant variables by providing hazard ratios, p-values, and likelihood ratio tests, clearly quantifying the impact of covariates such as Air temperature, Process temperature, Torque, and Machine type on failure risks. Similar findings have been reported by Susto et al.[3], who emphasized the Cox model's strength in handling multiple covariates and censored data effectively. In contrast, the Weibull model primarily describes the overall failure distribution through its shape and scale parameters but struggles to isolate specific covariate effects clearly.[21] Consequently, while offering useful insights into general system behavior, the Weibull model lacks the flexibility and detailed diagnostics of the Cox model, particularly in scenarios with time-varying or numerous covariates.

Thus, the Cox model proves more effective for precisely identifying significant predictors of machine failure, whereas the Weibull model is more appropriate for broad analyses or when detailed covariate relationships are less critical.[22]

Predictive Accuracy and Robustness

The results clearly demonstrate significant performance differences between the Cox and Weibull models for RUL prediction using sensor data from industrial equipment. The Cox model exhibited superior predictive accuracy with a C-index of 0,9148, reflecting excellent discriminative capability. This aligns with previous studies indicating the strong predictive power of semi-parametric survival models when multiple dynamic covariates are present. In contrast, the Weibull model showed a moderate performance (C-index = 0,639), limiting its application when covariate effects are prominent or variable over time, a limitation also noted by recent literature.[23]

Moreover, the Akaike Information Criterion (AIC), which penalizes model complexity, further validated the Cox model's superior performance (4184,44 vs. 5915,87 for Weibull). This outcome aligns with previous research highlighting the Cox model's balance between goodness-of-fit and simplicity, effectively avoiding overfitting despite its semi-parametric structure.[19,23]

One major advantage of the Cox model identified here is its ability to provide personalized RUL predictions dynamically adapted to individual machine conditions, similar to findings by Parri et al.[21] Conversely, the Weibull model yields a single median RUL based on a global distribution, potentially overlooking variability inherent in complex industrial environments.

However, the Cox model relies heavily on the proportional hazards assumption, which, although validated in this study, requires careful assessment and potential adjustments in different scenarios.[24] Despite its underperformance in our dataset, the Weibull model remains valuable in simpler or predictable wear-out contexts due to its straightforward implementation and interpretability, as noted in recent industrial applications.[22,25]

Model Limitations

Cox Model *w*hile robust and effective, the Cox PH model assumes proportional hazards over time, which might not always hold true, and can be sensitive to multicollinearity. Therefore, careful feature selection is essential to avoid overfitting.[22]

Weibull Model *a*lthough straightforward to implement and widely utilized in reliability engineering, the Weibull model's assumption of a fixed distribution shape may lead to underfitting or misleading survival estimates, particularly in heterogeneous or noisy operational environments.[23,25]

CONCLUSION

Our findings show that the Cox Proportional Hazards (Cox PH) model outperforms the Weibull model for predicting Remaining Useful Life (RUL), achieving higher accuracy (C-index 0,91 vs. 0,64) and better model fit (AIC 4184 vs. 5916). The Cox model effectively handles multiple covariates, censored data, and provides individualized predictions, making it well-suited for complex industrial environments shaped by dynamic factors.

However, it relies on the proportional hazards assumption, requiring validation for each application. The Weibull model, though less accurate here, remains useful for systems with predictable wear-out patterns.

This comparative framework paves the way for future research combining survival models with machine learning, where hybrid approaches like neural networks integrated with Cox or Weibull models could improve prediction while maintaining interpretability

REFERENCES

1. Predictive Maintenance: A Novel Framework for a Data-Driven, Semi-Supervised, and Partially Online Prognostic Health Management Application in Industries. Disponible sur: https://www.mdpi.com/2076-3417/11/8/3380

2. Zhu T, Ran Y, Zhou X, Wen Y. A Survey of Predictive Maintenance: Systems, Purposes and Approaches. arXiv. Disponible sur: http://arxiv.org/abs/1912.07383

3. Si G, Xia T, Zhu Y, Du S, Xi L. Triple-level opportunistic maintenance policy for leasehold service network of multi-location production lines. Reliability Engineering & System Safety. 1 oct 2019;190:106519.

4. Zheng H, Kong X, Xu H, Yang J. Reliability analysis of products based on proportional hazard model with degradation trend and environmental factor. Reliability Engineering & System Safety. 1 déc 2021;216:107964.

5. Meeker WQ, Escobar LA, Pascual FG. Statistical Methods for Reliability Data. John Wiley & Sons; 2021. 708 p.

6. Liu J, Wang ,Bing Xing, Wang ,Shasha, and Zhou H. The prediction of RULs for Weibull or gamma k-out-of-n system with nonidentical components based on component data. Quality Technology & Quantitative Management. 0(0):1-26.

7. Zhang H, Gao Z, Du C, Bi S, Fang Y, Yun F, et al. Parameter estimation of three-parameter Weibull probability model based on outlier detection. RSC Advances. 2022;12(53):34154-64.

8. Feng J, Zhang H, Li F. Investigating the relevance of major signaling pathways in cancer survival using a biologically meaningful deep learning model. BMC bioinformatics. déc 2021;22(1):47.

9. Mathpati YC, More KS, Tripura T, Nayek R, Chakraborty S. MAntRA: A framework for model agnostic reliability analysis. Reliability Engineering & System Safety. 1 juill 2023;235:109233.

10. Zhou D, Sun CP, Du YM, Guan X. Degradation and reliability of multi-function systems using the hazard rate matrix and Markovian approximation. Reliability Engineering & System Safety. 1 févr 2022;218:108166.

11. Xu A, Wang R, Weng X, Wu Q, Zhuang L. Strategic integration of adaptive sampling and ensemble techniques in federated learning for aircraft engine remaining useful life prediction. Applied Soft Computing. 1 mai 2025;175:113067.

12. Zhang S, Zhai Q, Li Y. Degradation modeling and RUL prediction with Wiener process considering measurable and unobservable external impacts. Reliability Engineering & System Safety. 1 mars 2023;231:109021.

13. A Novel Remaining Useful Life Prediction Approach Combined eXtreme Gradient Boosting and Multi-Quantile Recurrent Neural Network | IEEE Journals & Magazine | IEEE Xplore. Disponible sur: https://ieeexplore.ieee.org/abstract/document/10478510

14. Smith AM, Hinchcliffe GR. RCM--Gateway to World Class Maintenance. Elsevier; 2003. 362 p.

15. Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. Sci Rep. 26 mars 2021;11(1):6968.

16. Collett D. Modelling Survival Data in Medical Research. 4e éd. New York: Chapman and Hall/CRC; 2023. 556 p.

17. Hasegawa T. Sample size determination for the weighted log-rank test with the Fleming-Harrington class of weights in cancer vaccine studies. Pharm Stat. 2014;13(2):128-35.

18. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. Statistics in Medicine. 15 juin 2005;24(11):1713-23.

19. Chen C, Liu Y, Wang S, Sun X, Di Cairano-Gilfedder C, Titmus S, et al. Predictive maintenance using cox proportional hazard deep learning. Advanced Engineering Informatics. 1 avr 2020;44:101054.

20. Yang Z, Kanniainen J, Krogerus T, Emmert-Streib F. Prognostic modeling of predictive maintenance with survival analysis for mobile work equipment. Sci Rep. 20 mai 2022;12(1):8529.

21. Parii D, Janssen E, Tang G, Kouzinopoulos C, Pietrasik M. Predicting the Lifespan of Industrial Printheads with Survival Analysis. arXiv; 2025. Disponible sur: http://arxiv.org/abs/2504.07638

22. Khuntia SR, Zghal F, Bhuyan R, Schenkel E, Duvivier P, Blancke O, et al. Use of survival analysis and simulation to improve maintenance planning of high voltage instrument transformers in the Dutch transmission system. arXiv; 2023. Disponible sur: http://arxiv.org/abs/2301.01239

23. Lillelund CM, Pannullo F, Jakobsen MO, Pedersen CF. Predicting Survival Time of Ball Bearings in the Presence of Censoring. arXiv; 2023. Disponible sur: http://arxiv.org/abs/2309.07188

24. Cavalcante T, Ospina R, Leiva V, Cabezas X, Martin-Barreiro C. Weibull Regression and Machine Learning Survival Models: Methodology, Comparison, and Application to Biomedical Data Related to Cardiac Surgery. Biology. mars 2023;12(3):442.

25. Petsinis P, Naskos A, Gounaris A. Analysis of key flavors of event-driven predictive maintenance using logs of phenomena described by Weibull distribution. arXiv; 2021. Disponible sur: http://arxiv.org/abs/2101.07033

26. Karmakar B, Pradhan B. Residual lifetime prediction for heterogeneous degradation data by Bayesian semi-parametric method. arXiv; 2025. Disponible sur: http://arxiv.org/abs/2504.15794

FINANCING

The researchers state that they did not receive any financial assistance, grants, or other support while preparing this manuscript.

CONFLICT OF INTEREST

The authors declare that they have no financial or non-financial conflicts of interest to disclose

AUTHORSHIP CONTRIBUTION

Conceptualization: Hamdi Alaoui Abdelhafid, Anwar Meddaoui, Professor
Data Curation: Hamdi Aloui Abdelhafid, Anwar Meddaoui.
Formal Analysis: Hamdi Alaoui Abdelhafid, Anwar Meddaoui.
Research: Hamdi Alaoui Abdelhafid, Anwar Meddaoui.
Methodology: Hamdi Alaoui Abdelhafid, Anwar Meddaoui.
Project Management: Hamdi Alaoui Abdelhafi, Anwar Meddaoui.
Resources: Hamdi Alaoui Abdelhafid, Anwar Meddaoui.
Software: Hamdi Alaoui Abdelhafid, Anwar Meddaoui.
Supervision: Anwar Meddaoui, Ahmed En-Nhaili.
Validation: Anwar Meddaoui, Ahmed En-Nhaili .
Display: Hamdi Alaoui Abdelhafid, Anwar Meddaoui, Ahmed En-Nhaili.
Drafting - Original Draft: Hamdi Alaoui Abdelhafid, Anwar Meddaoui, Ahmed En-Nhaili.
Writing - Proofreading and Editing: Hamdi Alaoui Abdelhafid, Anwar Meddaoui, Ahmed En-Nhaili.