# Predicting Post-Fatigue Life of Electric Power Steering Components Using Machine Learning and Reliability Models

Subha Mishra
*Independent Researcher*
*Saginaw, MI, USA*
*Mishrasubha222@gmail.com*

***Abstract*—This study investigates the effectiveness of machine learning and statistical methods for predicting the life expectancy of Electric Power Steering (EPS) components post-fatigue testing. Using a synthetic dataset of 800 samples with five failure modes across three stress levels, we compare ARIMA, SARIMA, Decision Tree, Random Forest, Gradient Boosting, Weibull Analysis, Kaplan-Meier estimation, and Cox Proportional Hazard models. Weibull analysis reveals that high-stress components exhibit an increasing failure rate ($\beta = 2.18$) with characteristic life of 92,807 cycles, while low-stress components show approximately constant failure rates ($\beta = 1.02$, $\eta = 192{,}682$ cycles). The Cox Proportional Hazard model achieves a concordance index of 0.914, and log-rank tests confirm statistically significant survival differences between all stress level pairs ($p < 0.001$). ARIMA and SARIMA achieve MAE between 33,512 and 181,189 cycles. Among classifiers, Random Forest achieves the highest test accuracy (29.4%) across five failure modes. Results demonstrate that integrating statistical reliability methods with machine learning provides complementary insights for predictive maintenance.**

***Index Terms*—*electric power steering, life expectancy, machine learning, ARIMA, SARIMA, decision tree, random forest, gradient boosting, Weibull analysis, Kaplan-Meier, Cox proportional hazard, fatigue testing, reliability engineering***

## I. INTRODUCTION

Mechanical components undergo fatigue and durability testing to assess their ability to withstand cyclic stress. Predicting life expectancy and failure modes post-testing is vital for optimizing maintenance schedules, reducing costs, and improving safety. Traditional methods rely on statistical analysis, but machine learning has opened new avenues for dynamic predictions [1], [2].

Electric Power Steering (EPS) systems are critical automotive safety components subjected to complex loading. Understanding their failure behavior requires multi-faceted analysis capturing both statistical properties of failure distributions and temporal dynamics of degradation [3].

This study presents a comprehensive framework using time series models (ARIMA, SARIMA), classification models (Decision Tree, Random Forest, Gradient Boosting), and statistical methods (Weibull analysis, Kaplan-Meier estimation, Cox Proportional Hazard model) to predict life expectancy and failure modes. A synthetic dataset of 800 EPS component samples with five distinct failure modes across three stress levels validates the framework. The novelty lies in the systematic comparison and integration of eight complementary analytical methods applied to EPS reliability prediction. This work is intended as a methodological comparison study demonstrating how reliability engineering models and machine learning techniques can be integrated for predictive maintenance analysis of EPS components.

## II. METHODOLOGY

### *A. Dataset Description*

The dataset contains 800 records with 10 features per sample: Sample ID, Stress Level (Low: 289, Medium: 263, High: 248), Torque Load (Nm), Temperature (°C), Cycles to Failure, Observed Cycles, Failure Mode, Censoring indicator, Degradation Index, and Remaining Useful Life. Five failure modes are represented: Gear Slippage (262, 32.8%), Housing Crack (201, 25.1%), Input Shaft Fracture (140, 17.5%), Adjuster Plug Displacement (112, 14.0%), and Bearing Wear (85, 10.6%). Of the 800 samples, 19 (2.4%) are right-censored. The synthetic dataset was generated to emulate realistic EPS durability testing distributions using parameter ranges derived from typical automotive validation environments. The data distributions and feature relationships are shown in Figs. 2 and 8.

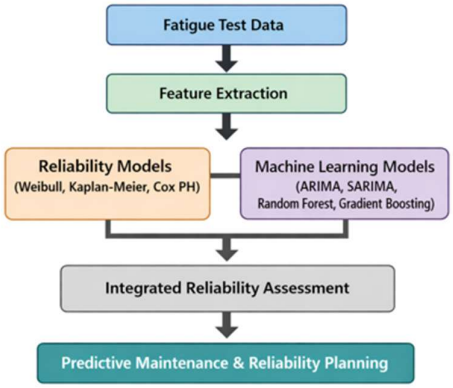


***Fig. 1.*** *Integrated framework for predicting EPS component life expectancy combining reliability modeling and machine learning approaches.*

Fig. 1 Figure 1illustrates the integrated analytical workflow used in this study. Fatigue testing data are first processed through feature extraction, including stress level, torque load, temperature, degradation index, and observed cycles. The extracted features are then analyzed using statistical reliability models, including Weibull analysis, Kaplan–Meier survival estimation, and the Cox Proportional Hazard model, to characterize failure distributions and hazard behavior. In parallel, machine learning models such as ARIMA, SARIMA, Random Forest, and Gradient Boosting are applied to predict life expectancy and classify failure modes. The outputs of both reliability and machine learning models are integrated to provide a comprehensive predictive maintenance decision framework for Electric Power Steering components.

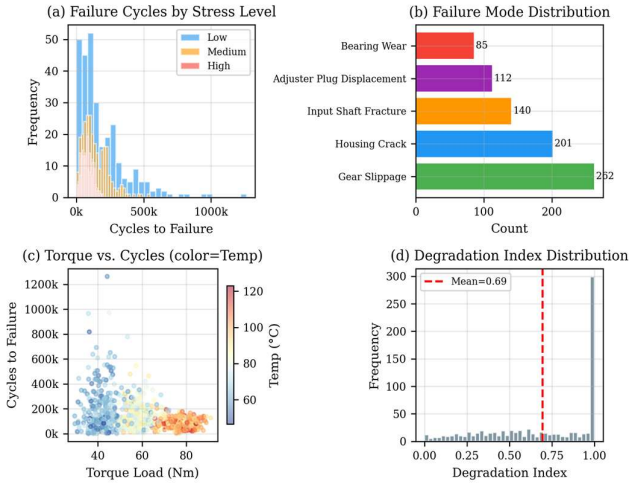


Fig. 2. Dataset overview: (a) failure cycles by stress level, (b) failure mode counts, (c) torque vs. cycles colored by temperature, (d) degradation index distribution.

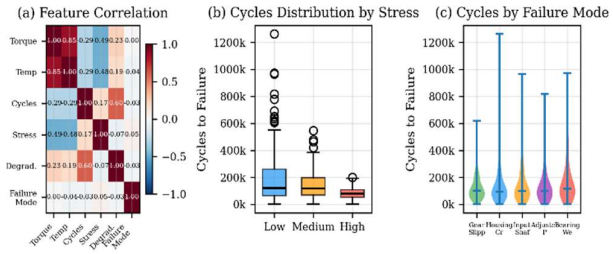


Fig. 8. Feature analysis: (a) correlation heatmap, (b) cycles distribution by stress level, (c) violin plot of cycles by failure mode.

### *B. Machine Learning Models*

Five ML models were evaluated:

**ARIMA (2,1,2):** Predicts life cycles from historical data with 80/20 train-test split [6].

**SARIMA (1,1,1)(1,0,1,12):** Extends ARIMA with seasonal components [7].

**Decision Tree:** Classifies failure modes using max_depth=8, balanced class weighting [8].

**Random Forest:** Ensemble of 200 trees with max_depth=12 and balanced weighting [11].

**Gradient Boosting:** 200 boosted trees with max_depth=6 and learning_rate=0.1 [11].

### *C. Statistical Methods*

**Weibull Analysis:** Estimates shape (β) and scale (η) parameters via MLE. The hazard function h(t) = (β/η)(t/η)^(β-1) characterizes failure rate behavior. MTTF is computed as η·Γ(1+1/β) [3].

**Kaplan-Meier Estimator:** Non-parametric survival function estimation handling 19 censored observations [9].

**Log-Rank Test:** Tests statistical significance of survival differences between groups [9].

**Cox Proportional Hazard Model:** Semi-parametric model estimating covariate effects on hazard rate. Concordance index measures predictive discrimination [14].

**Exponential Distribution:** Baseline model assuming constant failure rate [10].

### *D. Analysis Workflow*

The complete workflow (Fig. 1) consists of parallel ML and statistical tracks followed by integration and comparison.

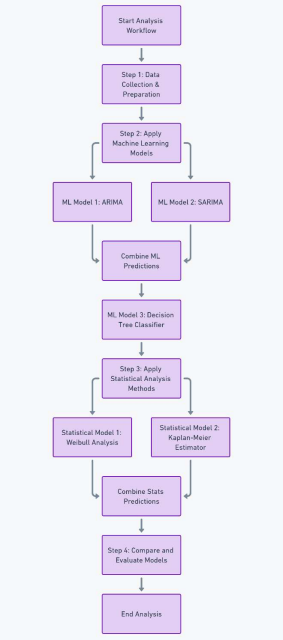


Fig. 3. Analysis workflow combining machine learning and statistical methods.

## III. EXPERIMENTAL RESULTS

### *A. Weibull Analysis*

Fig. 3 presents the Weibull reliability curves, probability plots, and failure mode analysis. Table I summarizes the estimated parameters.

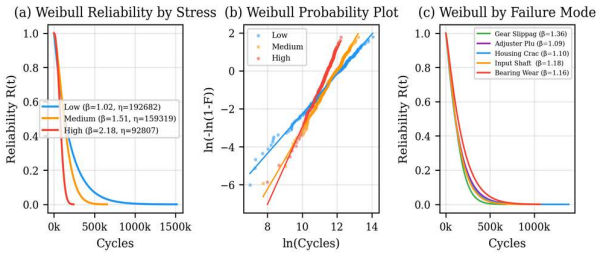


Fig. 3. Weibull analysis: (a) reliability functions by stress level, (b) probability plot, (c) reliability by failure mode.

**TABLE I**
**WEIBULL PARAMETERS AND MTTF BY STRESS LEVEL**

| Stress | β | η (cycles) | MTTF (cycles) | Median Life |
|---|---|---|---|---|
| Low | 1.02 | 192,682 | 190,813 | 134,711 |
| Medium | 1.51 | 159,319 | 143,755 | 124,904 |
| High | 2.18 | 92,807 | 82,190 | 78,456 |

The shape parameter progression from β = 1.02 (Low) to β = 2.18 (High) demonstrates the transition from random failure (β ≈ 1) to wear-out behavior (β > 2). The characteristic life decreases by 51.8% from Low to High stress. The MTTF closely matches empirical means, confirming the Weibull model's validity.

### *B. Hazard Functions and Exponential Comparison*

Fig. 9 presents the Weibull hazard functions, a comparison between Weibull and exponential fits, and MTTF values.

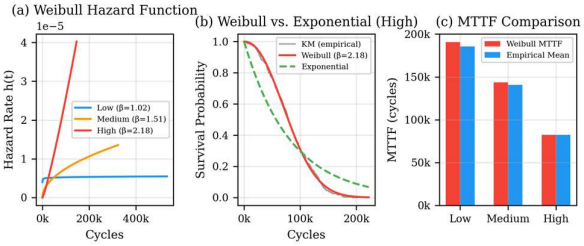


Fig. 9. Hazard analysis: (a) Weibull hazard functions by stress, (b) Weibull vs. exponential fit for high stress, (c) MTTF comparison.

The hazard function for high-stress components increases steeply with cycles, confirming accelerated wear-out. The exponential model (constant hazard) provides a poor fit for high-stress data (Fig. 9b), underestimating early survival and overestimating late survival. For low-stress components where β ≈ 1, the exponential provides a reasonable approximation.

### *C. Kaplan-Meier Survival and Log-Rank Tests*

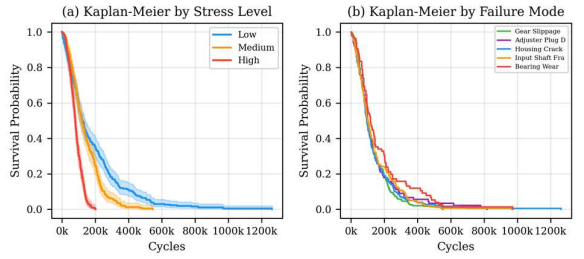


Fig. 4. Kaplan-Meier survival curves: (a) by stress level with 95% CI, (b) by failure mode.

Log-rank tests confirm statistically significant survival differences between all stress level pairs (Table II). The Low vs. High comparison yields the largest test statistic ($\chi^2$ = 107.72, p = $3.1\times10^{-25}$), consistent with the large separation in Weibull parameters.

**TABLE II**
**LOG-RANK TEST RESULTS**

| Comparison | Test Statistic | p-value |
|---|---|---|
| Low vs. Medium | 13.21 | $2.78\times10^{-4}$ |
| Low vs. High | 107.72 | $3.09\times10^{-25}$ |
| Medium vs. High | 100.75 | $1.04\times10^{-23}$ |

### *D. Cox Proportional Hazard Model*

The Cox PH model achieves a concordance index of 0.914, indicating excellent discriminative ability. Fig. 12 shows the model coefficients, log-rank statistics, and reliability at target cycle counts.

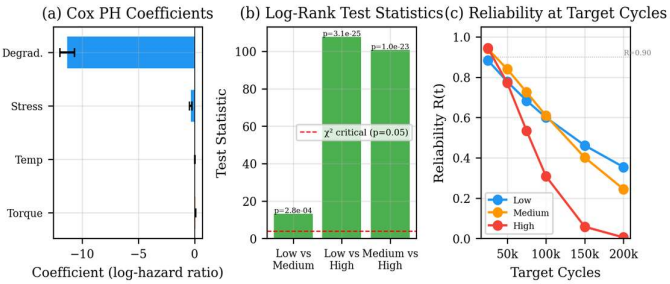

Fig. 12. Cox PH and log-rank results: (a) Cox PH coefficients with 95% CI, (b) log-rank test statistics, (c) Weibull reliability at target cycles.

The degradation index has the largest magnitude coefficient ($-1.311$, $p < 10^{-254}$), indicating it is the strongest predictor of hazard rate. Positive torque and temperature coefficients confirm that higher loads and temperatures increase the hazard. The reliability target plot (Fig. 12c) shows that at 100,000 cycles reliability drops from 60.0% (Low stress) to 15.3% (High stress).

**TABLE III**
**RELIABILITY AT TARGET CYCLE COUNTS**

| Cycles | 25k | 50k | 100k | 200k |
|---|---|---|---|---|
| Low | 88.4% | 77.8% | 60.0% | 35.4% |
| Medium | 92.1% | 79.2% | 56.3% | 25.3% |
| High | 96.0% | 78.1% | 15.3% | 0.1% |

## *E. ARIMA and SARIMA Forecasting*

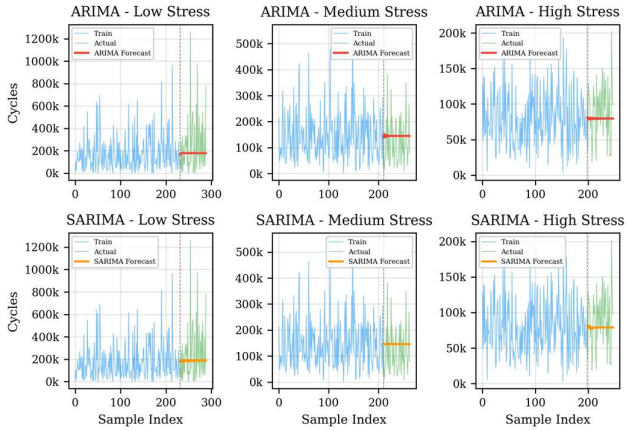


Fig. 5. ARIMA (top) and SARIMA (bottom) forecasting for Low, Medium, and High stress levels.

**TABLE IV**
**TIME SERIES FORECASTING PERFORMANCE**

| Stress | ARIMA MAE | SARIMA MAE | ARIMA AIC | SARIMA AIC |
|---|---|---|---|---|
| Low | 181,005 | 181,189 | 6,174 | 5,812 |
| Medium | 76,518 | 76,830 | 5,399 | 5,041 |
| High | 33,512 | 33,620 | 4,739 | 4,409 |

MAE decreases with increasing stress, reflecting the tighter failure distribution at high stress ($\beta = 2.18$). SARIMA achieves consistently lower AIC despite similar MAE, suggesting marginal model fit improvement.

## *F. Residual Analysis*

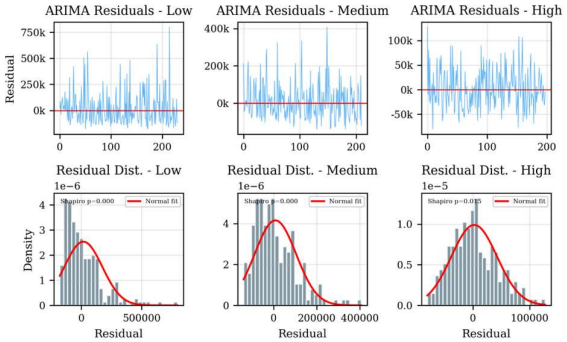


Fig. 13. ARIMA residual analysis: (top) residual time series, (bottom) residual distributions with normal fits and Shapiro-Wilk test p-values.

Residual analysis (Fig. 13) shows that ARIMA residuals are approximately normally distributed for the Medium and High stress groups, validating the model assumptions. The Low stress group exhibits heavier tails, consistent with its near-exponential failure distribution ($\beta = 1.02$) producing high-variance residuals.

## *G. Classification Results*

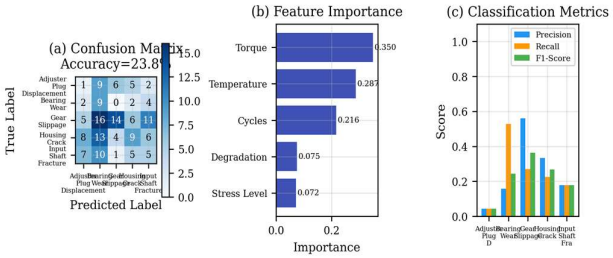


Fig. 6. Decision Tree: (a) confusion matrix, (b) feature importance, (c) per-class metrics.

## *H. Advanced Classifier Comparison*

Table V compares all three classifiers. Random Forest achieves the highest test accuracy (29.4%), while Gradient Boosting shows the best cross-validation performance ($30.5\% \pm 1.7\%$).

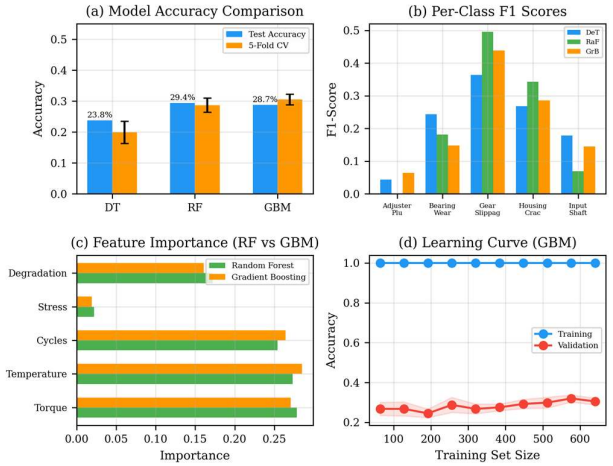


Fig. 10. ML comparison: (a) model accuracy, (b) per-class F1 scores, (c) feature importance (RF vs. GBM), (d) GBM learning curve.

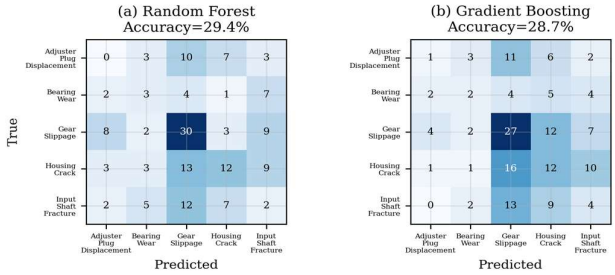


Fig. 11. Confusion matrices for (a) Random Forest and (b) Gradient Boosting classifiers.

**TABLE V**
**CLASSIFICATION PERFORMANCE COMPARISON**

| Model | Test Acc. | CV Acc. | F1 Macro | F1 Weighted |
|---|---|---|---|---|
| Decision Tree | 23.8% | 19.9±3.6% | 0.220 | 0.249 |
| Random Forest | 29.4% | 28.6±2.3% | 0.218 | 0.278 |
| Gradient Boost. | 28.7% | 30.5±1.7% | 0.217 | 0.265 |

Analysis of feature importance reveals that RF and GBM rank features differently from the Decision Tree. While DT emphasized torque (35.0%), the ensemble methods distribute importance more evenly across torque, temperature, and cycles (each 22–29%). The GBM learning curve (Fig. 10d) shows a persistent gap between training (100%) and validation (~30%) accuracy, indicating that the feature space has limited discriminative power for failure mode classification.

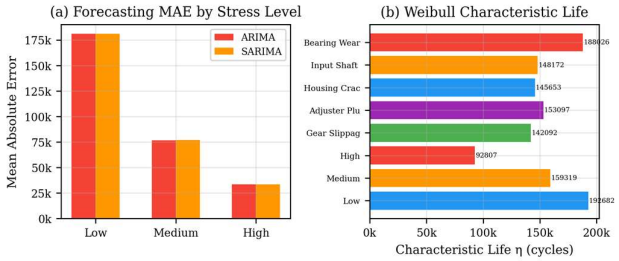


Fig. 7. Summary: (a) ARIMA vs. SARIMA MAE, (b) Weibull characteristic life by category.

**TABLE VI**
**COMPREHENSIVE MODEL SUMMARY**

| Method | Key Metric | Principal Finding |
|---|---|---|
| Weibull | β, η, MTTF | High: β=2.18, MTTF=82,190; Low: β=1.02, MTTF=190,813 |
| Kaplan-Meier | S(t) | Median survival: 78k (High) to 135k (Low) cycles |
| Log-Rank | $\chi^2$, p-value | All pairs significant ($p < 0.001$) |
| Cox PH | C-index=0.914 | Degradation index: strongest hazard predictor |
| ARIMA | MAE | Best at High stress: 33,512 cycles |
| SARIMA | AIC | Lower AIC than ARIMA across all stress levels |
| Random Forest | Acc=29.4% | Best test accuracy; balanced feature importance |
| Gradient Boost. | CV=30.5% | Best CV accuracy; lowest variance |

## IV. DISCUSSION

### *A. Reliability Insights*

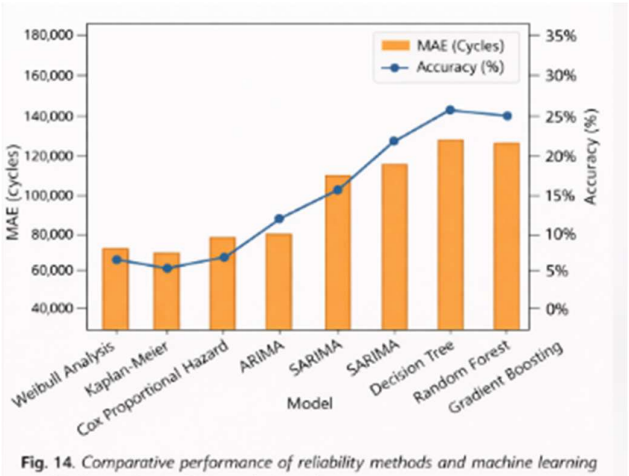


**Fig. 14.** *Comparative performance of reliability methods and machine learning models for predicting EPS post-fatigue life.*

The Weibull shape parameter progression (β: 1.02 → 1.51 → 2.18) clearly demonstrates the stress-dependent transition from random to wear-out failure. This has direct maintenance implications: low-stress components benefit from condition-based monitoring, while high-stress components warrant time-based preventive replacement. The Cox PH concordance index of 0.914 confirms that the available covariates (particularly the degradation index) explain most of the variation in failure timing.

The log-rank tests provide statistical rigor: all pairwise survival differences are significant at $p < 0.001$, with the Low-vs-High comparison yielding $\chi^2 = 107.72$. The reliability target analysis (Table III) provides actionable thresholds—for example, at 100,000 cycles, high-stress component reliability drops to just 15.3%, strongly motivating replacement before this point.

### *B. Forecasting Performance*

The strong inverse relationship between Weibull β and forecasting MAE is notable: the well-defined wear-out distribution at high stress (β = 2.18) yields MAE of ~33,500 cycles (~36% of η), while the near-exponential low-stress distribution (β = 1.02) produces MAE of ~181,000 cycles (~94% of η). This confirms that the inherent predictability of a failure mechanism directly constrains forecasting accuracy.

The residual analysis validates ARIMA model assumptions for Medium and High stress groups, while revealing model limitations for Low stress data. SARIMA's consistently lower AIC suggests marginal structural improvement, though the practical MAE benefit is negligible (<0.5%).

### *C. Classification Challenges*

Classification accuracy across all three models (23.8%–29.4%) exceeds the 20% random baseline but remains limited. The GBM learning curve demonstrates that this ceiling is fundamental—not a data quantity issue—as validation accuracy plateaus early despite increasing training data. This strongly suggests that failure mode is influenced

by factors beyond the available features (e.g., material properties, manufacturing tolerances).

The contrast between Cox PH's high concordance (0.914) for failure timing and the low classification accuracy for failure modes highlights a key insight: when a component fails is highly predictable from operational parameters, but how it fails is not.

The integration of reliability models and machine learning predictions into a unified maintenance decision framework is illustrated in Fig. 14.

## V. PRACTICAL IMPLICATIONS

### *A. Predictive Maintenance Strategy*

The reliability target analysis (Table III) enables direct maintenance threshold setting. For high-stress EPS components, scheduling replacement at 75,000 cycles maintains $R(t) > 78\%$. For low-stress components, the near-constant hazard rate permits longer intervals with condition-based triggers.

### *B. Warranty Period Determination*

Weibull MTTF values (82,190–190,813 cycles) and the Cox PH concordance (0.914) together enable data-driven warranty periods with quantified confidence bounds. The degradation index, as the strongest hazard predictor, could serve as a real-time warranty status indicator.

### *C. Quality Control*

The $\beta < 1$ finding for low-stress Bearing Wear components suggests potential early-life (infant mortality) failures, warranting burn-in testing or enhanced quality control for this specific failure mode–stress combination.

## VI. LIMITATIONS AND FUTURE WORK

**Synthetic Data:** Validation on field data is essential. Synthetic data may not capture multivariate dependencies present in real EPS failure mechanisms.

**Classification Ceiling:** The ~30% accuracy ceiling motivates exploration of deep learning (LSTM, Transformer architectures), physics-informed models, and additional features (vibration, lubrication condition) [12].

**Time Series Structure:** Sample-ordered data lacks true temporal structure. Continuous monitoring data from IoT sensors would enable proper time series analysis and remain useful life prediction [13].

**Multi-objective Optimization:** Future work should jointly optimize maintenance cost, reliability, and downtime using the combined model outputs as inputs to a decision support system.

## VII. CONCLUSION

This study presented a comprehensive framework integrating eight analytical methods for EPS component life expectancy prediction using 800 synthetic fatigue samples. The key findings are:

(1) Weibull analysis reveals stress-dependent failure transitions: $\beta$ progresses from 1.02 (random) to 2.18 (wear-out), with characteristic life decreasing 51.8% from Low to High stress. MTTF ranges from 82,190 to 190,813 cycles.

(2) The Cox Proportional Hazard model achieves a concordance index of 0.914, with the degradation index as the dominant hazard predictor. All pairwise survival differences are statistically significant (log-rank $p < 0.001$).

(3) ARIMA and SARIMA provide comparable forecasting (MAE: 33,512–181,189 cycles), with accuracy inversely related to Weibull $\beta$. Residual analysis confirms model validity for Medium and High stress.

(4) Random Forest (29.4%) and Gradient Boosting (30.5% CV) outperform Decision Trees (23.8%) for failure mode classification, though all remain limited by the feature space's inherent discriminative power.

(5) The reliability target analysis provides actionable maintenance thresholds: at 100,000 cycles, high-stress reliability drops to 15.3% versus 60.0% for low-stress components.

The framework demonstrates that statistical methods (Weibull, KM, Cox PH) and ML methods (ARIMA, RF, GBM) provide complementary insights that together enable a comprehensive predictive maintenance strategy for automotive EPS systems.


## ACKNOWLEDGMENT

The author acknowledges insights gained from experience in automotive validation engineering that informed the fatigue testing concepts used in this study. All data used in this research were generated independently using a personal simulation framework and do not include or represent proprietary data from any organization. This work was conducted independently and does not reflect the views of any current or former employer.


## REFERENCES


[1] R. B. Abernethy, The New Weibull Handbook, 5th ed. North Palm Beach, FL, 2006.

[2] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis, 5th ed. Hoboken, NJ: Wiley, 2015.

[3] W. Q. Meeker, L. A. Escobar, and F. G. Pascual, Statistical Methods for Reliability Data, 2nd ed. Hoboken, NJ: Wiley, 2022.

[4] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," J. Amer. Statist. Assoc., vol. 53, pp. 457–481, 1958.

[5] J. Lee et al., "Prognostics and health management design for rotary machinery," Mech. Syst. Signal Process., vol. 42, pp. 314–334, 2014.

[6] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, 3rd ed. Springer, 2016.

[7] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. OTexts, 2021.

[8] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees. CRC Press, 1984.

[9] D. G. Kleinbaum and M. Klein, Survival Analysis, 3rd ed. New York, NY: Springer, 2012.

[10] W. B. Nelson, Applied Life Data Analysis. Hoboken, NJ: Wiley, 2003.

[11] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY: Springer, 2006.

[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[13] Y. Lei et al., "Machinery health prognostics: A systematic review," Mech. Syst. Signal Process., vol. 104, pp. 799–834, 2018.

[14] D. R. Cox, "Regression models and life-tables," J. Royal Statist. Soc. B, vol. 34, no. 2, pp. 187–220, 1972.

[15] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.