

Reliability-Informed Life Prediction for New Energy Vehicle Components

WANG, Zengcong ^{1*}

¹ ZF Asia Pacific Group Co., Ltd, CN

* WANG, Zengcong is the corresponding author, E-mail: 122104983@qq.com

Abstract: For new energy vehicle components, remaining service life requires crucial maintenance, but operating conditions and incomplete fault records still limit the sustainable development of models. This study constructs a data-driven framework for batteries and traction motors by integrating fault analysis, reliability parameter estimation, and machine learning-based prediction. This framework can support maintenance planning and spare parts planning to some extent. However, these results should be interpreted with caution because component type, brand coverage, and data quality may introduce unobserved biases. More extensive cross-brand datasets, prediction ranges that account for uncertainty, and real-time validation are needed before reliable large-scale deployment.

Keywords: New Energy Vehicle Components, Remaining Useful Life, Reliability Analysis, Machine Learning, Condition-based Maintenance, .

Disciplines: Intelligent Systems.

Subjects: Other.

DOI: <https://doi.org/10.70393/6a696574.343332>

ARK: <https://n2t.net/ark:/40704/JIET.v1n3a02>

1 INTRODUCTION

1.1 RESEARCH BACKGROUND AND SIGNIFICANCE

The expansion of new energy vehicles is often narrated as a relatively direct transition from internal-combustion powertrains to electrified mobility, yet the reliability problem that accompanies this transition is considerably less linear. Batteries, traction motors, power electronics, and electronic control units operate as a tightly coupled system in which thermal exposure, charging behavior, road load, software control, manufacturing variability, and maintenance practice may interact over long periods. One study situates part of this problem upstream, arguing through an integrated smart-manufacturing and quality-control framework that production automation, sensor-based inspection, and information-flow integration influence the initial quality of batteries, motors, and electronic-control components ^[1]. Because that study concentrates on manufacturing performance rather than longitudinal failure observation, it cannot by itself establish service-life relationships; it nevertheless suggests that field degradation data should not be interpreted independently of process variation. A component can remain functional while its degradation trajectory changes, and an apparently stable health indicator may conceal a local defect or an operating regime not represented in historical data. Reliability, in this setting, is not only the probability of avoiding a discrete failure; it also concerns the preservation of an acceptable performance

margin under uncertain and changing conditions.

This problem has practical consequences. Long-term battery research roadmaps have treated durability, safety, testing, and deployment as connected development problems rather than isolated performance targets ^[2], although a programmatic roadmap does not substitute for component-level field evidence. Maintenance based exclusively on fixed intervals may replace usable components prematurely, whereas maintenance delayed until an observable fault appears can increase safety risk, unplanned downtime, and secondary damage. Remaining useful life (RUL) prediction offers a possible middle path by linking condition information to maintenance decisions, although a numerical prediction does not become operationally useful merely because its average error is low. The prediction must remain sufficiently calibrated across operating environments, and the uncertainty surrounding it must be understandable to engineers who decide whether to inspect, derate, repair, or replace a component.

The present study examines reliability analysis and life prediction for batteries and traction motors, which were selected because they represent different degradation structures. Battery aging is strongly influenced by electrochemical and thermal history, while motor degradation may emerge through bearing wear, insulation deterioration, electromagnetic disturbance, or cooling-system weakness. Treating both components with one undifferentiated model would be convenient but potentially misleading. The

proposed framework consequently combines component-specific failure reasoning with data-driven estimation, allowing common analytical stages while retaining room for different features, distributions, and model classes.

1.2 RESEARCH OBJECTIVES AND QUESTIONS

The study seeks to construct a reliability-informed framework that can organize heterogeneous operational and maintenance data, identify plausible failure mechanisms, and estimate component life without presenting data-driven prediction as detached from engineering knowledge. Four questions guide the analysis: which failure modes and covariates appear most relevant to battery and motor reliability; how can reliability models and machine-learning methods be connected without forcing incompatible assumptions; how stable are the reported predictions under holdout and cross-validation procedures; and what additional evidence would be required before the framework could guide maintenance or supply-chain decisions at scale?

These questions are intentionally framed as open empirical problems. The available records permit model development and case-level comparison, but they do not establish that the same relationships will hold across every manufacturer, climate, chemistry, motor architecture, or usage pattern. The contribution is thus better understood as an integrated and testable analytical framework, supported by preliminary component cases, rather than as a universally validated prognostic solution.

2 THEORETICAL BASIS AND CRITICAL REVIEW

2.1 RELIABILITY CONCEPTS FOR DEGRADING COMPONENTS

Reliability engineering conventionally defines reliability $R(t)$ as the probability that an item performs its required function up to time t under specified conditions, with $F(t)=1-R(t)$ representing cumulative failure probability. Mean time to failure and mean time between failures remain useful fleet-level summaries, but their interpretation becomes less straightforward when repair policies, censored observations, and gradual performance loss are involved. An exponential distribution assumes a constant hazard rate and may be convenient during a limited useful-life phase; a Weibull model is more flexible because its shape parameter can represent decreasing, constant, or increasing hazard. Even so, distributional flexibility does not guarantee physical adequacy, particularly when a fleet mixes operating regimes or when failure labels are defined inconsistently.

Reliability engineering conventionally defines reliability $R(t)$ as the probability that an item performs its required function up to time t under specified conditions, with $F(t)=1-R(t)$ representing cumulative failure probability.

Let T denote the component failure time. The reliability and cumulative failure functions are formally defined as

$$R(t)=P(T > t) = 1 - F(t)$$

This general definition does not require a particular lifetime distribution, although its empirical interpretation remains dependent on the criteria used to identify failure and censored observations.

A Weibull model is more flexible because its shape parameter can represent decreasing, constant, or increasing hazard.

$$R(t)=\exp \left[-\left(\frac{t}{\eta} \right) ^{\beta} \right]$$

For condition-monitoring applications, survival quantities and RUL estimates should be considered alongside the process by which observations enter the dataset. A maintenance record may identify the time of replacement rather than the time at which degradation became functionally unacceptable. Preventive replacement creates right-censoring, while delayed reporting may shift the recorded failure time in the opposite direction. These issues make parameter estimates sensitive not only to the component but also to service policy, data governance, and the practical definition of failure.

2.2 FAILURE MECHANISMS IN BATTERIES AND TRACTION MOTORS

Battery capacity fade and resistance growth may reflect calendar aging, cycling intensity, high state of charge, thermal exposure, fast charging, or combinations of these factors. Thermal runaway is a different event class: it is rare, safety-critical, and not necessarily predictable from the same features used to estimate gradual capacity loss. Electrochemical impedance measurements have been used to examine low-temperature behavior^[3], a method that provides mechanism-sensitive information but does not, by itself, resolve the transfer from controlled tests to irregular field use. From a materials perspective, advanced-material applications in battery systems, traction motors, electronic controls, lightweight structures, and thermal-management systems have been linked to component performance^[4]. This work broadens the mechanism discussion, although its experimental and case-based evidence is more suitable for identifying plausible material pathways than for quantifying RUL. A critical comparison of state-of-charge and state-of-health indicators further shows the value of combining indicators^[5]; it also reveals that indicator performance depends on sensing quality and battery context, which limits the case for a single universal health variable.

Traction-motor reliability involves another set of interacting processes. Bearing wear, lubrication condition, thermal cycling, winding insulation, cooling performance, and load transients can generate overlapping signatures.

Vibration and frequency-domain features may reveal mechanical defects, while current, temperature, and efficiency trends may be more informative for electrical or thermal degradation. A high-amplitude feature may represent a developing fault, an abrupt operating-condition change, or a sensor artifact. Failure analysis must therefore precede, and later be revisited after, statistical feature selection rather than being treated as a one-time preliminary exercise.

2.3 PROGNOSTICS LITERATURE: RELEVANCE AND LIMITS

A foundational review organized machinery prognostics around data acquisition, diagnostics, and prognostics, providing a durable conceptual basis for condition-based maintenance [6]. That review predates many contemporary deep-learning architectures, however, and the practical difficulties it identifies - sparse failure examples, uncertain thresholds, and the integration of predictions into maintenance decisions - have not disappeared with increased model complexity. A broad account of electric and hybrid vehicle technology helps situate component reliability within vehicle architecture [7], although its breadth necessarily limits detailed treatment of field-level prognostic validation.

Data-driven battery studies commonly report gains from nonlinear learning, but comparisons are often made on curated laboratory cycles, where load patterns and end-of-life definitions are more controlled than in fleet records. A review of prognostics and health management for lithium-ion batteries points toward the promise of learning-based methods [8]; a review synthesis, however, cannot remove publication bias or guarantee comparability between datasets. Physical models provide interpretability but may require parameters that are difficult to observe in service. Statistical models are transparent but can impose restrictive distributional assumptions. Machine-learning models can capture interactions, though they may learn fleet-specific maintenance practices instead of degradation mechanisms. The present work draws on all three traditions while treating their tension as analytically useful rather than attempting to erase it.

2.4 RESEARCH GAP AND POSITIONING

A recurring gap is the separation of failure-mechanism analysis from predictive modeling. Some studies construct detailed physical explanations without testing whether they improve out-of-sample prediction; others optimize prediction error while leaving uncertain whether the learned relationship is stable, causal, or operationally meaningful. Cross-brand transfer is also frequently claimed from limited evidence, even though differences in sensor calibration, component design, software strategy, and warranty policy can produce dataset shift.

The framework developed here addresses this gap to some extent by using fault tree analysis and reliability estimation to structure feature development, followed by

candidate-model comparison and robustness checks. Its innovation lies less in a single new algorithm than in the disciplined connection between failure logic, heterogeneous field data, model evaluation, and maintenance interpretation. Because the current dataset remains limited in documented brand and component diversity, cross-domain generalization is treated as a future hypothesis requiring further research.

3 RESEARCH DESIGN AND DATA

3.1 ANALYTICAL FRAMEWORK

The research design consists of four connected activities: reliability and failure-mode analysis; data preparation and feature construction; candidate-model development; and evaluation against held-out observations. Although this sequence appears orderly on paper, the actual analytical process was iterative. Early feature sets were revised after missingness patterns were examined, and the distinction between an anomalous sensor value and a rare but credible operating event could not always be made mechanically. Fault-tree assumptions were consequently revisited after exploratory analysis, while model residuals were used to identify operating regions in which the initial engineering representation appeared incomplete.

This iterative structure matters because an apparently clean pipeline can conceal consequential judgment. Removing every extreme temperature, for example, may improve average prediction error while eliminating the operating conditions most relevant to accelerated aging. Retaining every extreme observation may instead allow corrupted sensor values to dominate the fit. The adopted process used statistical screening together with engineering plausibility checks, although the absence of a fully documented independent adjudication panel remains a limitation.

3.2 DATA SOURCES AND PREPROCESSING

The dataset combined vehicle-operation records, maintenance and repair histories, and environmental or operating-condition variables such as temperature, humidity, driving mode, and load. The initial collection contained 10,000 records. After duplicate, missing, inconsistent, and anomalous entries were addressed, 8,500 records remained. This reduction improved internal consistency, but it also removed 15% of the initial records; such a change cannot be treated as neutral because missingness may be associated with severe faults, sensor interruption, or service events.

Numeric variables were scaled for model classes sensitive to magnitude, and categorical variables were encoded using a consistent training-data vocabulary. Preprocessing transformations were estimated on the training partition and then applied to validation or test data to reduce information leakage. Where interpolation was considered for short gaps in time-series signals, it was restricted to gaps whose duration was small relative to the components

sampling interval and degradation timescale. Longer gaps were not filled as if they were directly observed. The source material does not provide sufficient detail to reproduce every threshold, so the reported findings should be read as framework-level evidence; a future archival package should include the data dictionary, exclusion logic, software versions, and preprocessing code.

TABLE 1. DATASET CONDITION BEFORE AND AFTER PREPROCESSING

Measure	Before	After	Interpretive note
Total records	10,000	8,500	15% of records removed or unresolved
Outlier proportion	5%	0% flagged	May improve consistency but can suppress rare events
Missing-value proportion	10%	0% retained	Mechanism of missingness remains relevant
Data consistency	Low	Higher	Depends on harmonized definitions and sensors

3.3 FEATURE ENGINEERING AND SELECTION

Feature construction combined descriptive statistics, temporal trends, and frequency-domain information. Battery candidates included temperature exposure, charge-discharge cycle count, usage pattern, and health-related trajectories. Motor candidates included load variation, temperature, maintenance frequency, and signal features potentially associated with mechanical or electromagnetic behavior. Rolling means and variances summarized local condition, while trend estimates represented slower degradation. Fourier-derived features were considered when periodic or vibration-like signals were available, although frequency-domain interpretation depends strongly on sampling stability and operating speed.

Feature selection combined correlation screening and recursive feature elimination. Correlation was used cautiously: a weak marginal relationship does not imply irrelevance when interactions are present, and a strong relationship may arise from maintenance policy or leakage from a near-failure indicator. Recursive elimination can improve parsimony, but its selections are conditional on the estimator used. Features were therefore assessed for stability across resamples and for engineering plausibility, not solely for their rank in one fitted model. This did not remove subjectivity; it made the points of judgment more visible.

3.4 CANDIDATE MODELS AND PARAMETER ESTIMATION

The candidate set included Weibull reliability models, linear or support-vector regression, decision trees, random

forests, gradient boosting, and long short-term memory networks for sequential data. These models were not assumed to answer exactly the same question. Weibull analysis characterizes population-level failure-time behavior under a chosen distribution, whereas regression and sequence models estimate RUL conditional on observed features. Reliability parameters were considered through maximum-likelihood estimation, Bayesian updating where defensible prior information existed, and non-parametric estimation when a fixed distribution appeared too restrictive.

Hyperparameters were explored through grid or random search, with Bayesian optimization and genetic search identified as possible extensions for larger parameter spaces. Model complexity was constrained by validation performance and by the need for interpretable maintenance use. Deep learning was considered only where sequence length and sample support appeared adequate; otherwise, its additional flexibility could generate a convincing fit without stable external performance. The available paper-level evidence does not identify one universally superior final architecture, so the revised account presents a candidate-model framework rather than attributing all reported errors to an unspecified hybrid model.

3.5 VALIDATION AND PERFORMANCE MEASURES

Data were separated into training and holdout sets using proportions in the 70:30 to 80:20 range, and K-fold cross-validation was used to examine variability across resamples. For time-dependent records, random splitting may allow later information from the same component to influence training; group-aware or chronological partitioning is preferable when component identifiers and timestamps permit it. Mean squared error (MSE), mean absolute error (MAE), and R^2 were used because they emphasize different aspects of error. MSE is sensitive to large deviations, MAE is easier to interpret in the prediction unit, and R^2 measures explained variation but does not establish calibration or safety.

Prediction intervals were reported in the component cases, although their statistical construction is not fully specified in the source study. Future evaluation should report interval coverage and width, error stratified by operating regime, and performance near end of life, where maintenance consequences are most acute. A model with slightly lower overall accuracy may be preferable if it is better calibrated for high-risk cases or less sensitive to dataset shift.

4 RELIABILITY ANALYSIS

4.1 FAILURE DATA AND EVENT DEFINITION

Failure records were grouped by component, observed symptom, probable cause, timing, location, and maintenance response. Descriptive analysis established event frequencies and exposure patterns, while Kaplan-Meier estimation was considered for failure-time data containing censoring. Event definition required particular care. Battery replacement may

follow capacity loss, a diagnostic code, warranty policy, or customer preference; motor repair may be triggered by noise before functional failure. Combining these outcomes into one label would increase sample size but weaken interpretability.

The analysis therefore distinguished, where records allowed, gradual performance degradation from abrupt or safety-critical events. This separation reduces the risk that a model optimized for common degradation predicts rare hazardous events with unjustified confidence. It also limits the number of observations available for each outcome, illustrating a persistent trade-off between label specificity and statistical power.

4.2 FAULT TREE ANALYSIS

Fault tree analysis (FTA) was used as a top-down representation of how basic and intermediate events might lead to component-level failure. For a battery thermal event, candidate branches included abnormal heat generation, ineffective heat removal, control failure, and external abuse, with subevents such as overcharge, internal short circuit, sensor error, or cooling-system malfunction. For motor degradation, branches included bearing damage, insulation deterioration, cooling weakness, abnormal load, and electromagnetic disturbance.

FTA provided a transparent bridge from engineering reasoning to measurable features, but its apparent clarity should not be confused with causal proof. Logical gates simplify dependencies, basic-event probabilities may be poorly estimated, and common-cause failures can violate independence assumptions. Quantitative results were consequently interpreted as prioritization aids. When residual analysis suggested errors concentrated under high temperature or changing load, the corresponding branches were reconsidered rather than accepting the initial tree as final.

4.3 RELIABILITY PARAMETERS AND INFLUENCING FACTORS

Maximum-likelihood estimates provide efficient parameter estimates under a correctly specified distribution, but their stability can deteriorate with small failure counts or heavy censoring. Bayesian estimation may incorporate prior engineering knowledge and express parameter uncertainty; poorly justified priors, however, can produce an appearance of precision that the data do not warrant. Non-parametric approaches relax distributional assumptions but may become unstable in sparsely observed tails. The study therefore treats agreement across methods as supportive evidence and disagreement as a diagnostic signal.

Regression and sensitivity analysis indicated that temperature, cycling or usage intensity, load variation, and maintenance history were plausible reliability drivers. Several explanations remain possible. A temperature effect may represent direct thermal acceleration, harsher driving, geographic differences, or sensor placement. Maintenance

frequency may be protective because inspection detects faults, or it may be associated with already problematic vehicles. Considering these alternatives, the results are more appropriately interpreted as predictive associations that guide investigation, rather than as isolated causal effects.

5 LIFE-PREDICTION MODEL DEVELOPMENT

5.1 MODEL CONSTRUCTION AND ITERATION

Model development began with transparent baselines before more flexible estimators were considered. This ordering made it possible to assess whether nonlinear or sequential models added predictive value beyond simple age, exposure, and condition trends. The initial feature matrix was revised when highly correlated variables produced unstable importance estimates and when several time-series features were unavailable for a material subset of records. Rather than filling all missing sequences, the analysis compared a reduced common-feature model with richer component-specific models.

Random forests and gradient boosting were suitable for nonlinear interactions and mixed feature scales; support-vector regression offered a different bias-variance profile; LSTM models were relevant where ordered histories were sufficiently complete. Weibull-derived reliability quantities could enter the decision framework as population-level context, but combining them with machine-learning output requires care because the resulting quantities may be conditioned on different information. The study therefore avoids describing the framework as a single mathematically unified hybrid unless the exact fusion rule is documented.

5.2 ROBUSTNESS AND GENERALIZATION

Robustness was assessed through cross-validation, sensitivity analysis, and outlier review. Stability across folds provides some evidence against a result driven by one partition, although observations from related vehicles or repeated measurements can still create dependence between folds. Sensitivity analysis examined whether modest changes in influential features produced disproportionate shifts in predicted life. Outlier detection was used diagnostically rather than as an automatic deletion rule, since rare observations may contain either corrupted values or early evidence of an uncommon failure mode.

Generalization across brands, models, climates, and duty cycles remains less certain. Sensor definitions can change while retaining similar names, and component-management software may alter degradation without being represented as an explicit covariate. Transfer learning and domain adaptation are possible responses, but they should be evaluated against a strong strategy of recalibration on local data. Further research is needed to determine whether shared representations capture transferable degradation structure or

merely smooth over meaningful design differences.

5.3 FROM POINT PREDICTION TO MAINTENANCE EVIDENCE

A point estimate of RUL is insufficient when the cost of early replacement differs sharply from the cost of late intervention. Maintenance use requires an uncertainty range, a decision threshold, and a statement of the consequences of error. Prediction intervals can support this transition if their empirical coverage is tested. A narrow interval that excludes the true failure time is less useful than a wider but well-calibrated interval, especially for safety-relevant components.

Prediction intervals can support this transition if their empirical coverage is tested.

The empirical prediction interval coverage probability can be expressed as

$$PICP = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(L_i \leq y_i \leq U_i)$$

Where L_i and U_i are the lower and upper interval bounds and $\mathbb{I}(\cdot)$ is the indicator function. A narrow prediction interval is not necessarily informative if its empirical coverage is substantially below the intended confidence level.

Interpretability is similarly contextual. Feature importance may show which variables influenced a model, but it does not necessarily explain a degradation mechanism. Local explanations can assist engineers in checking whether a prediction aligns with observed temperature, load, or health trends. Their role should be advisory: they can reveal implausible dependencies and possible data leakage, yet they do not replace physical inspection or causal validation.

6 CASE RESULTS AND DISCUSSION

6.1 BATTERY LIFE PREDICTION

For the reported battery cases, the lithium-ion battery prediction error was approximately 5%, with a prediction interval of 1,000-1,500 hours, while the nickel-metal hydride case showed an error of approximately 8% and an interval of 800-1,200 hours. These values suggest that the framework can recover useful degradation information under the evaluated conditions, but the difference between battery types should not be attributed automatically to chemistry. Sample composition, sensor coverage, operating variability, maintenance thresholds, and the amount of training data available for each battery type could also explain part of the error gap.

The relatively lower lithium-ion error may reflect stronger feature availability or a more coherent operating population, rather than inherent predictability. Conversely, the nickel-metal hydride result may have been affected by smaller or more heterogeneous records. Without confidence intervals around the error metrics, repeated external tests, and

a documented baseline comparison, “high precision” is best treated as a provisional description. The interval widths also deserve attention: an acceptable percentage error at fleet level may still translate into a maintenance window too broad for a particular operator.

To clarify the percentage errors reported in the component cases, the absolute percentage error for observation i is defined as:

$$e_i = \frac{|y_i - \hat{y}_i|}{y_i} \times 100\%$$

while its sample-level average can be written as:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Percentage-based errors may become unstable as the observed RUL approaches zero, meaning that the metric should be interpreted together with MAE and the absolute maintenance window.

TABLE 2. REPORTED BATTERY CASE RESULTS

Battery type	Prediction error	Prediction interval	Cautious interpretation
Lithium-ion	5%	1,000-1,500 h	Promising within the evaluated data; external calibration not established
Nickel-metal hydride	8%	800-1,200 h	Higher error may reflect chemistry, sample size, or operating heterogeneity

6.2 MOTOR LIFE PREDICTION

For motors, the permanent-magnet synchronous motor case produced an error of approximately 6% and a 5,000-6,000-hour interval; the asynchronous motor case produced an error of approximately 10% and a 4,000-5,000-hour interval. One interpretation is that the available features represented the degradation of the permanent-magnet system more effectively. Another possibility is that the asynchronous-motor sample included wider load variation, different maintenance practices, or more mixed failure mechanisms. A third explanation is methodological: a model selected on aggregate error may underfit a subgroup whose signal-to-noise ratio differs from the dominant sample.

The motor results illustrate why component labels alone are not sufficient stratification variables. Bearing supplier, cooling design, control strategy, vehicle application, and duty cycle may be at least as influential as motor class. Future analysis should report errors by failure mechanism and operating regime, and should test whether residuals increase near threshold conditions. Such evaluation may reveal that a model is useful for routine wear but unreliable for abrupt insulation or cooling failures.

TABLE 3. REPORTED MOTOR CASE RESULTS

Motor type	Prediction error	Prediction interval	Cautious interpretation
Permanent-magnet synchronous	6%	5,000-6,000 h	Potentially useful, subject to failure-mode and fleet stratification
Asynchronous	10%	4,000-5,000 h	Further optimization and subgroup validation are needed

6.3 COMPARATIVE INTERPRETATION AND POTENTIAL BIAS

Across the four cases, reported errors ranged from 5% to 10%. The pattern is consistent with a framework that offers some predictive value, but it does not establish superiority over traditional methods because the source evidence does not provide matched baseline metrics, uncertainty around the estimates, or an external validation cohort. The reduction from 10,000 to 8,500 records may have improved signal quality, yet it could also have reduced representation of rare and severe events. Selection bias may arise if vehicles with complete service histories differ systematically from those with fragmented records.

Measurement bias is another concern. Temperature recorded at one sensor location may not represent local cell or winding exposure, while maintenance records may encode organizational practice as much as physical condition. Model comparisons can also be optimistic when hyperparameter selection and performance assessment are not cleanly separated. Nested cross-validation, group-aware splitting, and a prospective fleet study would provide stronger evidence. Considering the current results together, the most defensible claim is that the framework is promising for decision support under conditions similar to those represented in the dataset, while dependable transfer remains unresolved.

7 OPERATIONAL APPLICATION AND RESEARCH PROSPECTS

7.1 CONDITION-BASED MAINTENANCE

RUL estimates can support condition-based maintenance by identifying components for inspection before a predicted failure window and by reducing unnecessary replacement when a component retains an adequate health margin. Implementation should use tiered actions rather than a single automated replacement threshold. A moderately shortened RUL may trigger enhanced monitoring, while a short RUL accompanied by abnormal temperature or vibration could trigger immediate inspection. Human review remains important when the model encounters an unfamiliar

operating regime or produces a prediction inconsistent with diagnostic evidence.

Economic evaluation should include the cost of false alarms, missed failures, inspection, downtime, and secondary damage. A model with lower MAE is not necessarily economically preferable if its remaining errors are concentrated in safety-critical cases. Maintenance policy can also change the future dataset, because interventions prevent some failures from being observed. Continuous evaluation must account for this feedback rather than assuming that deployed data have the same distribution as development data.

7.2 SUPPLY-CHAIN AND FLEET PLANNING

Aggregated life predictions may inform spare-parts inventory and production planning, particularly when interval forecasts are combined across a fleet. A supply-chain resilience study using multidimensional time-series data from 15 manufacturing enterprises compared a hybrid temporal architecture with conventional forecasting approaches [9]. Its resilience index concerns order fulfillment, inventory turnover, capacity utilization, and transportation delay rather than component degradation, so the reported prediction errors cannot be transferred to the present RUL task. The study is nonetheless methodologically relevant to the downstream use of life estimates because it illustrates how incomplete, noisy, and temporally dependent operational indicators can be incorporated into recovery-oriented planning. Yet aggregation does not eliminate uncertainty. Correlated exposure, such as a heat wave or software update, can shift many components simultaneously and invalidate independence assumptions used in demand planning. Scenario-based forecasts are therefore more appropriate than a single deterministic demand figure.

At the vehicle-allocation level, one study used a year of operational data and a high-fidelity simulator to formulate shared-electric-vehicle rebalancing as a multi-agent reinforcement-learning problem [10]. Its objective was demand satisfaction and net revenue rather than maintenance reliability, and fleet-rebalancing performance should not be interpreted as evidence for RUL accuracy. Even so, the work indicates a possible decision layer in which predicted component condition, charging constraints, vehicle location, and service demand could eventually be considered together. For suppliers and fleet operators, model governance should specify data ownership, update frequency, performance monitoring, and responsibility for intervention decisions. Privacy and cybersecurity also become relevant when vehicle telemetry is transmitted for remote prognostics. These institutional dimensions are not peripheral: a technically adequate prediction model may fail operationally if its inputs are inconsistent, its alerts are not actionable, or accountability is unclear.

7.3 TECHNICAL CHALLENGES AND FUTURE WORK

Future research should first improve evidence quality rather than simply increasing algorithmic complexity. Cross-brand and cross-climate datasets, standardized failure labels, documented censoring, and prospective validation would allow more credible assessment of transferability. Uncertainty-aware methods, including conformal prediction or probabilistic sequence models, may provide calibrated intervals, although their coverage under distribution shift requires separate testing.

Physics-informed learning could constrain predictions toward plausible degradation behavior while retaining flexibility for field variation. Transfer learning may reduce data requirements for a new vehicle platform, but negative transfer is possible when designs differ materially. Model compression can support edge deployment, yet computational feasibility depends on more than parameter count. One lightweight scheduling framework separates offline resource analysis from real-time decisions to reduce inference overhead on heterogeneous edge hardware [11]; it does not evaluate prognostic models, but it draws attention to shared-memory contention and deadline instability that an offline RUL benchmark would not reveal. A related energy-latency analysis coordinates workload allocation with processor power-state management, suggesting that continuous monitoring may require explicit consideration of both response time and energy cost rather than treating deployment as a final engineering detail. Research on autonomous-vehicle edge devices further examines task affinity, cross-core migration, cache disruption, and computational jitter. These three scheduling studies address general edge-AI workloads rather than battery or motor degradation and should accordingly be treated as deployment analogues, not validation evidence for the present model. Federated approaches may permit multi-organization learning without centralizing raw data, although they introduce new validation and governance problems. This leads us to further thinking about whether the central research objective should be the most accurate isolated model or a learning system that recognizes when its own evidence is insufficient.

8 CONCLUSION

8.1 THEORETICAL AND PRACTICAL IMPLICATIONS

This study reframes component life prediction as an interaction between reliability theory, imperfect field evidence, and maintenance judgment. Fault trees and survival concepts provide structure, while data-driven models capture nonlinear patterns that may not be expressible through a single degradation equation. Their combination is valuable not because it removes uncertainty, but because it makes several forms of uncertainty - mechanism uncertainty, sampling uncertainty, model uncertainty, and decision uncertainty - more visible within one analytical process.

The battery and motor cases, with reported errors between 5% and 10%, indicate that condition data may support useful RUL estimation to some extent. They also show why prediction accuracy should not be separated from interval calibration, subgroup performance, and the consequences of error. In practical terms, the framework is most defensible as an aid to inspection, maintenance scheduling, and inventory planning, rather than as an autonomous replacement for engineering assessment.

8.2 LIMITATIONS AND FORWARD-LOOKING RESEARCH

The study is constrained by incomplete documentation of brand coverage, sample composition, final hyperparameters, interval construction, and matched baseline performance. Data cleaning removed 15% of the initial records, and the effect of this removal on rare-failure representation remains uncertain. The reported cases are informative but insufficient for universal claims about batteries or motor types. These limitations do not nullify the framework; they define the evidence still required for it to mature.

A more consequential next stage would connect prospective multi-fleet validation with explicit maintenance decisions and post-deployment monitoring. Such work should test not only whether predictions are accurate, but also whether they reduce unplanned failures, avoid premature replacement, remain calibrated after software or hardware changes, and communicate uncertainty in a form that engineers can use. If this broader standard is adopted, life prediction may develop from a narrowly optimized forecast into a reliability-governance instrument - one that supports safer electrified mobility while remaining open to revision as components, data, and operating environments evolve.

ACKNOWLEDGMENTS

Not Applicable.

FUNDING

Not Applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not Applicable.

INFORMED CONSENT STATEMENT

Not Applicable.

DATA AVAILABILITY STATEMENT

Not Applicable.

CONFLICT OF INTEREST

Not Applicable.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not application.

ABOUT THE AUTHORS

WANG, Zengcong

ZF Asia Pacific Group Co., Ltd, CN,
122104983@qq.com.

signal processing, 20(7), 1483-1510.

- [7] Chan, C. C. (2002). The state of the art of electric and hybrid vehicles. *Proceedings of the IEEE*, 90(2), 247-275.
- [8] Zhang, J., & Lee, J. (2011). A review on prognostics and health monitoring of Li-ion battery. *Journal of power sources*, 196(15), 6007-6014.
- [9] Huang, S. (2025). Measuring supply chain resilience with foundation time-series models. *European Journal of Engineering and Technologies*, 1(2), 49-56.
- [10] Luo, M., Du, B., Zhang, W., Song, T., Liu, K., Zhu, H., ... & Wen, H. (2023). Fleet rebalancing for expanding shared e-mobility systems: A multi-agent deep reinforcement learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 24(4), 3868-3881.
- [11] Hao, Z. (2026). Low-Overhead Scheduling for Real-Time AI Workloads on Multi-Core Edge Chips. *International Journal of Advance in Applied Science Research*, 5(3), 15-25.

REFERENCES

- [1] Wang, Z. (2024). Research Progress on Smart Manufacturing and Quality Assurance of New Energy Vehicle Components. *JOURNAL OF PROGRESS IN ENGINEERING AND PHYSICAL SCIENCE* Учредители: Aurora Publishing House Limited, 3(4), 56-65.
- [2] DOE, U. (2010). Multi-year research, development and demonstration plan: planned program activities for 2005-2015. US department of energy, office of energy efficiency and renewable energy, hydrogen, fuel cells and infrastructure technologies program (HFCIT), 1-34.
- [3] Zhang, S. S., Xu, K., & Jow, T. R. (2004). Electrochemical impedance study on the low temperature of Li-ion batteries. *Electrochimica acta*, 49(7), 1057-1061.
- [4] Duan, X., & Wang, Y. (2025). A Study on the Impact of Local Policy Response on the Technological Innovation of the New Energy Vehicle Industry. *Sustainability*, 17(19), 8873.
- [5] Xiong, R., Cao, J., Yu, Q., He, H., & Sun, F. (2017). Critical review on the battery state of charge estimation methods for electric vehicles. *Ieee Access*, 6, 1832-1843.
- [6] Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and*