A Hybrid Physics and Machine Learning Digital Twin for Remaining Useful Life Prediction of SiC Power Modules in Traction Inverters

Kooros Moabber[1], Hamed Talebi[2], and Mahmoud Rahat[3]

[1] *Research and Development Department, Volvo Cars Corporation, Gothenburg, Sweden*
kmoabber@volvocars.com

[2] *Department of Computer Science and Media Technology, Malmö University, Malmö, Sweden*
hamed.talebi@mau.se

[3] *Center for Applied Intelligent Systems Research (CAISR), Halmstad University, Sweden*
mahmoud.rahat@hh.se

Abstract

Silicon Carbide (SiC) power modules (PMs) are the most failure-prone component in automotive traction inverters, responsible for the large majority of reported inverter faults. There are two established strategies for estimating power module lifetime: a physics-based approach and a data-driven approach. Each strategy has complementary weaknesses: the physics-based approach cannot adapt to real-world degradation and is computationally heavy while pure ML approaches require large high-quality datasets, extrapolate poorly outside their training range and can produce physically unrealistic and hard-to-justify predictions. In contrast to prior hybrid SiC prognostics studies, this paper proposes and evaluates a hybrid physics and ML digital twin architecture that corrects a physics-based remaining useful life (RUL) estimate by using a machine-learned residual model trained on power-cycling data from 20 SiC power modules. Input features (on-state voltage, junction temperature swing, junction temperature, thermal resistance, accumulated fatigue damage and their moving averages) are standardized and reduced via Principal Component Analysis (PCA) to a single health indicator, which together with the physics-based damage estimate feeds a residual RUL model. Six candidate regression architectures, a multilayer perceptron (MLP), Extreme Gradient Boosting (XGBoost), linear regression, ridge regression, lasso regression and a quadratic response-surface model (QRSM) are compared by using 5-fold cross-validation. XGBoost and the quadratic model achieve the best trade-off between accuracy (mean $R^2 \approx 0.87$) and embedded implementability although the MLP achieves the highest raw accuracy ($R^2 = 0.94$) at higher computational cost. We further demonstrate online RUL tracking and residual-based anomaly detection under two synthetic field operating scenarios (physics-pessimistic and physics-optimistic) that emulate contrasting field-like degradation behaviors and finally discuss a path toward embedded deployment and further hybrid blending strategies.

1. Introduction

The traction inverter converts the DC voltage supplied by the traction battery into a three-phase variable-frequency AC current that drives the electric motor via pulse-width modulation (PWM), realized by a power module (PM) consisting of six power switches (Fig. 2a) where a real power module is illustrated in Fig. 2b from (Semikron Danfoss, 2024)).

The degradation of the power module illustrated in Fig. 1 is mainly driven by (1) thermal cycling (characterized by the mean junction temperature $T_{j,mean}$ and the delta junction temperature ΔT_j) which causes solder-layer fatigue and the lift-off of the bond-wire or the bond-wire cracking, (2) high current stress which can cause chip metallization cracking, and (3) mission-profile variability arising from driver behavior and environmental conditions which produces degradation that is difficult to predict in advance. These mechanisms manifest electrically as an increase in the on-state voltage ($V_{DS,on}$), an increase in the junction temperature and its swing for a given load and an increase in thermal resistance (R_{th}) where these three signatures are used throughout this paper as condition-monitoring features.

The power module is the most fragile component of the traction inverter: Original Equipment Manufacturer (OEM) field and reliability data indicate that approximately 82% of inverter failures originate in the power module itself. Accurate,

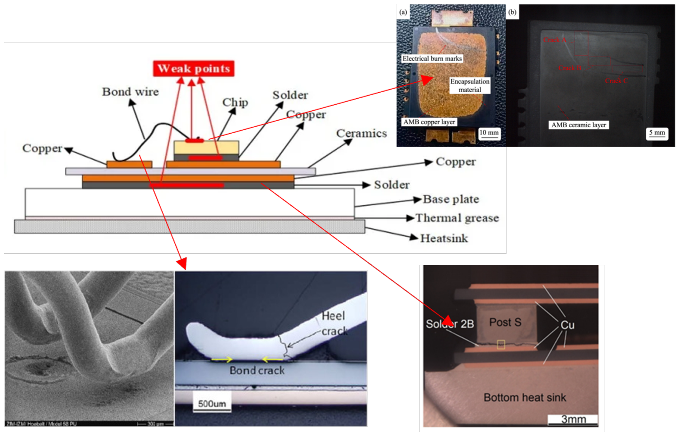


Figure 1. Power module layers and typical failures

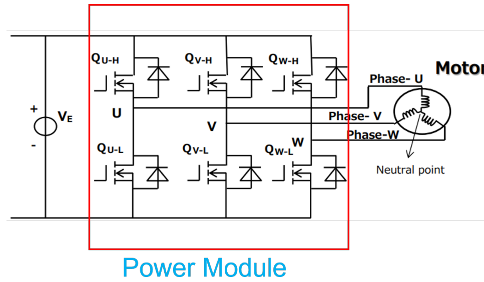
adaptive lifetime and health estimation of the power module is therefore a first-order concern for automotive OEMs, both at the design stage and during in-service predictive maintenance.

Previous research has described two dominant predictive maintenance philosophies for power modules: a physics-based lifetime estimation workflow used by OEMs (Pino et al., 2023) and a data-driven condition-monitoring approach reported in the literature (Di Nuzzo, Lewitschnig, Tuellmann, Rzepka, & Otto, 2023). These studies indicate that neither approach alone is fully sufficient: the physics-based approach is interpretable and requires no field data but cannot adapt to real degradation, whereas the data-driven approach adapts well to measured degradation, but risks unrealistic extrapolation and offers weaker physical grounding.

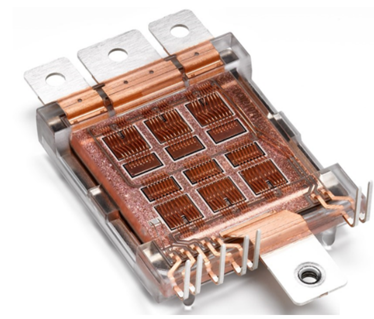
This paper builds directly on that review and presents our main contribution: a hybrid physics-ML digital twin that fuses both approaches into a single, embeddable RUL estimation framework, together with quantitative modeling results, cross-validated performance comparisons, and online tracking behavior under representative field-data scenarios. Compared with prior hybrid SiC prognostics studies, the contribution of this work lies in explicitly modeling the physics–ML interaction as a residual correction to a Coffin–Manson/Miner baseline. This formulation is tailored to traction-inverter operating profiles and emphasizes embedded implementability.

2. RELATED WORK

This section situates our hybrid approach within four strands of prior work: physics-based lifetime modeling, data-driven and ML-based RUL estimation including PCA-based health indicators, hybrid and physics-informed prognostics, and condition-monitoring approaches for traction inverters.



(a)



(b)

Figure 2. Six-switch SiC power module for a three-phase motor drive: (a) inverter topology with three upper and three lower switches; (b) Semikron Danfoss Shower Power physical module.

2.1. Physics-Based Lifetime Models

The Coffin-Manson law and its extensions remain widely used physics-based approaches for estimating the power module lifetime. (Bayerer, Herrmann, Licht, Lutz, & Feller, 2008) proposed an influential multi-parameter lifetime model incorporating peak junction temperature, pulse duration, current, and bond-wire diameter (Bayerer et al., 2008) and these and related Coffin-Manson-Arrhenius and Norris-Landzberg variants are surveyed in a recent review of junction-temperature-based lifetime estimation methods (Morel & Morel, 2024). A related Transactions on Power Electronics study combines Coffin-Manson-based analytical models with finite-element simulation and a metamodel-based surrogate to streamline lifetime assessment for automotive power electronics converters (Bhoi et al., 2025), illustrating the same finite element analysis (FEA)-plus-fatigue-law pipeline summarized in Section 3 of this paper.

2.2. Data-Driven and ML-Based RUL Estimation

Beyond the condition-monitoring method of (Di Nuzzo et al., 2023) discussed in Section 3, a companion study by the same group analyzes complementary electro-thermal degradation indicators for chip-solder-layer monitoring in SiC power switches (Di Nuzzo, Tuellmann, Methfessel, & Rzepka, 2022). For SiC MOSFETs more broadly, RUL prediction has been approached using an Extended Kalman Particle Filter applied to on-state voltage degradation, reducing prediction error relative to a standard particle filter (W. Wu, Gu, Yu, Gao, & Chen, 2023) and using a non-linear Wiener-process model that explicitly accounts for uncertainty in the failure threshold (Q. Wu, Xu, Xiao, & Wang, 2024). A recent survey of prognosis approaches for SiC MOSFETs finds that the majority of reported methods rely on neural-network-based, purely data-driven models and highlights out-of-sample validation and uncertainty quantification as open challenges (Kumar, Allard, Hologne-Carpentier, Clerc, & Auger, 2025), both of which motivate the cross-validation and residual-based anomaly detection used in our hybrid model. For IGBT modules, PCA-based feature reduction combined with a feedforward neural network has been used to construct a trend/health indicator for RUL estimation (Ismail, Saidi, Sayadi, & Benbouzid, 2020) and a broader review of data-driven IGBT prognostics techniques is given by Fang et al. (Fang et al., 2018). Similar data-driven RUL prediction methods have also been investigated for other vehicle components, particularly turbochargers, using autoencoder-based dimensionality reduction, domain adaptation, and survival analysis (Revanur, Ayibiowu, Rahat, & Khoshkangini, 2020; Rahat et al., 2022; Rahat, Kharazian, Mashhadi, Rögnvaldsson, & Choudhury, 2023).

2.3. Hybrid and Physics-Informed Approaches

Several recent studies combine physics-based and ML models in a manner conceptually aligned with our approach. A PHM Society study proposes a simple hybrid RUL model for SiC MOSFETs that uses junction-temperature swing and temperature-compensated on-state resistance as inputs to a damage-accumulation scheme based on Paris' crack law, explicitly framed as a first step toward physics-informed neural network (PINN) prognostics (Eng et al., 2023). Outside power electronics, a fatigue-life study integrates a particle-swarm-optimized XGBoost model with a physical fatigue model for aluminum alloys, reporting that the hybrid formulation improves on both a purely physical model and a purely data-driven XGBoost baseline (Li et al., 2024), supporting our choice of XGBoost as a residual-correction learner. More broadly, physics-informed machine learning (ML) has been reviewed as a general strategy for improving interpretability and data efficiency in digital twins for electrical machines and power converters (Nyangon, 2026).

2.4. Condition Monitoring for Traction Inverters

Closer to our specific application, an AC power-cycling test setup and associated condition-monitoring toolchain for SiC-based traction inverters has been reported, integrating on-resistance-based temperature-sensitive electrical parameter (TSEP) monitoring directly into the gate-driver desaturation-protection circuit for low-cost, in-vehicle implementation (Farhadi, Vankayalapati, Sajadi, & Akin, 2023), a system-level integration strategy consistent with the embedded architecture discussed in the earlier work (Pino et al., 2023).

Taken together, this body of work confirms that (i) physics-based models remain widely used for design-stage lifetime estimation but have limited ability to adapt to degradation observed during operation; (ii) PCA-derived pure ML health indicators have been widely used as a practical basis for data-driven RUL estimation but it needs qualified input data and can sometimes produce physically unrealistic results due to small or non-qualifies training set or poor extrapolation outside of the observed stress envelope; and (iii) explicit hybrid physics-ML fusion is an active but still comparatively young research direction, particularly for SiC power modules in traction applications, which is the gap this paper addresses. Table 1 summarizes key differences between representative hybrid SiC prognostics studies and the present work. Unlike approaches that tune physics models using data or directly regress RUL from measurements, our method retains the physics-based lifetime estimate as an explicit baseline and learns only the residual correction, which is particularly suited to small-data traction-inverter applications.

Table 1. Comparison of Hybrid Prognostics Approaches

Study	Eng et al. (2023)	Li et al. (2024)	This work (2026)
Hybrid Strategy	Physics-informed NN	Physics + XGBoost	Residual physics–ML
Data Source	MOSFET cycling	Alloy fatigue	SiC PM cycling
Focus	Conceptual PINN	Material testing	Automotive inverter

3. METHOD

3.1. Background

3.1.1. Predictive Maintenance of Power Module

Predictive maintenance of power modules can be approached from two complementary directions: (1) physics-based lifetime estimation, in which a mission profile is translated into losses, thermal cycles and a fatigue-based lifetime via the Coffin-Manson model, and (2) data-driven ML RUL estimation, in which degradation-sensitive electrical quantities (features) are measured directly, reduced via PCA and mapped to RUL through a regression model or neural network. The following sections summarize each approach and its trade-offs and finally introduce our hybrid contribution.

3.1.2. OEM Physics-Based Lifetime Estimation Workflow

The conventional workflow, illustrated in Fig. 3, proceeds as follows: different mission profiles (phase current, modulation index, power factor, DC bus voltage) that presents driving vehicles under the different realistic conditions and the module's electrical/thermal characteristics are used to compute losses, yielding the junction temperature T_j, the mean junction temperature $T_{j,mean}$, the max junction temperature $T_{j,max}$ and junction temperature swing ΔT_j, finite element analysis (FEA) provides the resulting thermal-mechanical stress, rainflow counting extracts discrete thermal cycles from the junction temperature profile and the Coffin-Manson fatigue model estimates the number of cycles to failure:

$$N_f = A_F(\Delta T_j)^{-\alpha}, \tag{1}$$

$$A_F = A \cdot \exp\left(\frac{E_g}{k}\left(\frac{1}{T_{j,ref}} - \frac{1}{T_j}\right)\right) \tag{2}$$

In the present study, junction temperature T_j and its swing ΔT_j are estimated from temperature-sensitive electrical parameters (TSEPs) using a calibrated thermal-network model together with the available module-level temperature sensor, rather than any direct junction-temperature measurement. This reflects typical traction-inverter and gate-driver implementations, where only module temperature is sensed and the junction temperature is obtained through thermal-network estimation.

Cumulative fatigue damage D is computed from the same Coffin–Manson/Miner model (Eq. 3) that yields RUL_{Phy}. Although this introduces a potential dependency between the physics baseline and the residual target, we mitigate leakage by grouping training and validation strictly at module level (Section 4.2) and by treating D as a compact summary of stress history rather than a direct proxy for the residual.

$$D = \sum_i \frac{n_i}{N_i} \tag{3}$$

where n_i is the number of cycles experienced and N_i the number of cycles allowable at stress level i. The module's physics-based RUL at a given cycle is defined as the number of remaining cycles until the accumulated damage reaches unity (Eq. 4).

$$RUL_{Phy} = \text{Failure Cycle} \cdot (1 - D) \tag{4}$$

3.1.3. Strengths and Limitations of Physics-Based Models

Physics-based lifetime models are physically interpretable, require no field data, are well suited to design-stage validation, and provide predictable, deterministic, worst-case-oriented estimates. However, their limitations are significant: they require accurate thermal models, are sensitive to mission-profile assumptions, cannot adapt to real-world degradation, offer no real-time RUL prediction, lose accuracy specifically in the late-life region where accuracy matters most, and are computationally heavy (not implementable in embedded traction inverter software). These weaknesses motivate the use of ML models.

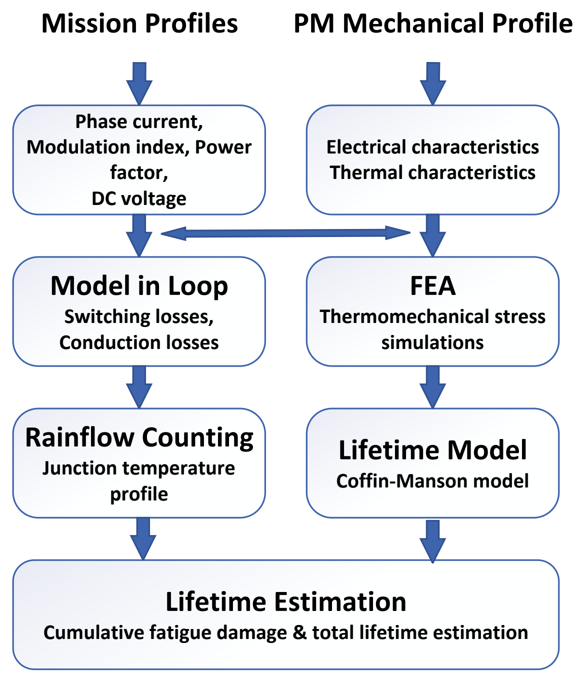


Figure 3. OEM Physics-Based Lifetime Estimation Workflow

3.1.4. Data-Driven RUL Prediction (Background)

As reviewed in earlier work (Pino et al., 2023), Nuzzo et al. (Di Nuzzo et al., 2023) that is depicted in Fig. 4 proposed estimating RUL from measured power module parameters, notably $V_{DS,hot}$ and $T_{j,max}$, collected during the accelerated power cycling test (PCT). The pipeline standardizes the measured signals, extracts a PCA-based health indicator (PC1: State of Health) and fits a dynamic ML-based regression model $Y = X\beta + \epsilon$, with $\beta = (X^TX)^{-1}X^TY$, evaluated via the coefficient of determination

$$R^2 = 1 - \frac{(y - X\beta)^T(y - X\beta)}{(y - \bar{y})^T(y - \bar{y})}. \tag{5}$$

As summarized in Table 2, this ML-based approach achieves lower complexity and higher reported accuracy than the Coffin-Manson approach and crucially offers online capabilities. Qualitative levels ('Low', 'Medium', 'High') reflect relative deployment effort rather than absolute metrics.

Table 2. Comparison of Physics-Based and ML-Based RUL Methods

Method	Complexity	Accuracy	Online Capability
Physics-based (Coffin-Manson)	Medium	Medium	No
ML-based (Regression)	Low	High	Yes

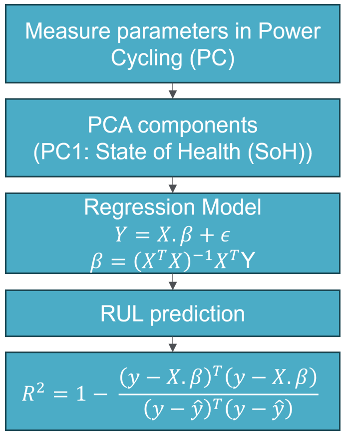


Figure 4. Regression method RUL prediction Workflow

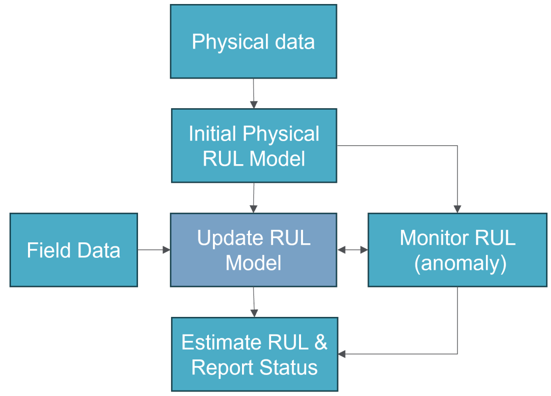


Figure 5. Hybrid digital twin architecture

3.1.5. Strengths and Limitations of ML-Based Models

Pure ML models learn real degradation patterns, adapt to real-world mission profiles, are computationally cheap, support fast online RUL prediction, and can handle stochastic module-to-module variability. However, they require large, high-quality datasets, extrapolate poorly outside their training range, can produce unrealistic predictions in the absence of physical constraints, are sensitive to measurement noise and sensor drift, and are harder to interpret and justify to safety reviewers. These complementary weaknesses of the physics-based and ML-based approaches motivate combining them into a hybrid model, which is the core contribution of this paper.

3.2. Proposed Hybrid Digital Twin Architecture

We propose a hybrid digital twin architecture in which a physics-based model predicts the expected RUL from the mission profiles, and a ML-based model predicts the actual degradation from field measurements. These two models are fused so that the hybrid model corrects the physics-based model in real time while the physics-based model constrains the ML predictions to remain physically plausible.

The architecture proceeds through the following stages, illustrated conceptually in Fig. 5: (i) generation of initial physical data from power cycling test, (ii) initial RUL modeling based on physical data, (iii) reception of field data yielding a true (observed) RUL, (iv) update of the hybrid model using the residual between true and physics-based RUL, (v) continuous monitoring of the RUL residual for anomalies and (vi) RUL estimation and status reporting.

3.2.1. Physical Dataset

The hybrid model is trained on power-cycling data from 20 SiC power modules. For each module and cycle, the dataset records the on-state drain to source voltage $V_{DS,on}$, thermal resistance R_{th}, junction temperature T_j, the junction temperature swing ΔT_j, the accumulated Coffin-Manson damage D (Eq. 3) and the resulting RUL, defined per module as Eq. 4. Based on the dissipated power (P_{diss}) and the temperature gradient ($T_j - T_{j,a}$ where $T_{j,a}$ is ambient temperature) provided, R_{th} is calculated (6) and the accumulated damage and junction temperature are considered to extract A_F (Eq. 1 and Eq. 2).

$$R_{th} = \frac{T_j - T_{j,a}}{P_{diss}}. \tag{6}$$

A sample dataset of power module has been illustrated in Fig. 6. The measured parameters have been noisy; therefore, the moving average of each parameter (orange line) has been calculated and shown in these figures.

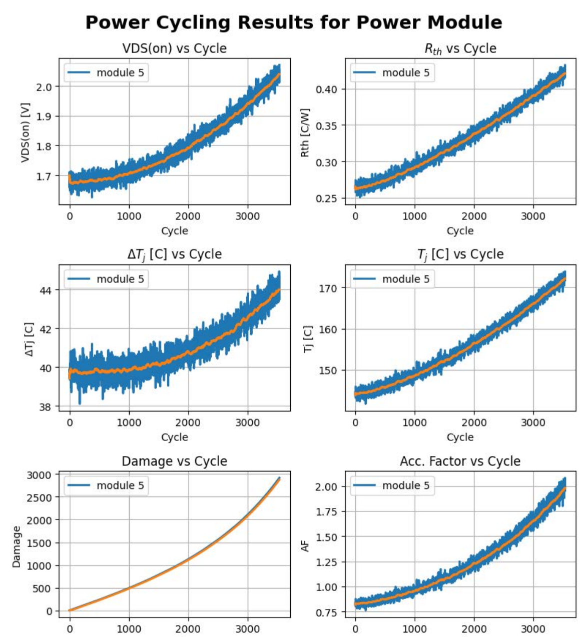


Figure 6. Typical power cycling test parameters

3.2.2. Hybrid Model Input Data and Processing

The regression input feature set consists of $V_{DS,on}$, ΔT_j, T_j, R_{th} and the accumulated damage D, together with their moving averages (moving average for D excluded since it is already a calculated cumulative calculated quantity). All features are standardized and reduced via PCA to a single principal health indicator, PC1, representing the module's State of Health (SoH). In the considered dataset, PC1 and PC2, shown in Fig. 7 for two sample power modules across all 20 modules, together explain approximately 88.6% of the feature variance (PC1: 71.7%, PC2: 16.9%), indicating that the original degradation-sensitive features are highly correlated and that most of the module's health information can be captured by a low-dimensional, effectively one-dimensional representation. This supports the use of PC1 alone as a compact, physically significant State of Health indicator for the subsequent RUL regression.

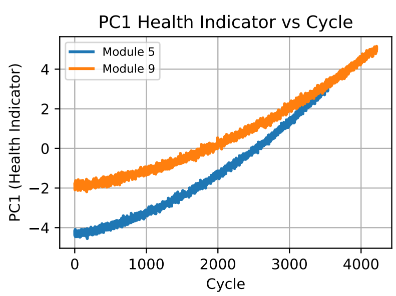


Figure 7. Typical PCA (SoH) for two SiC PM

The regression target is defined as the residual RUL (Eq. 7), i.e. the difference between the true (data-derived) RUL and the physics-based RUL to improve regression modeling so that the learned model corrects rather than replaces the physics-based estimate. We adopt a residual formulation in which the physics-based RUL estimate is treated as a baseline and the ML model predicts a correction, and therefore, the risk of extrapolating to implausible absolute RUL values when operating conditions deviate from the training set reduces, and moreover, the residual is learned around a physically meaningful baseline, which empirically limits extreme predictions compared to direct non-linear RUL regression. In this paper, nonnegative RUL and zero RUL at failure have been considered as constraints to avoid illogical results. Future work will incorporate other explicit constraints, e.g., monotonic decrease with cumulative damage in implementation. Additional constraints such as monotonic RUL decrease and bounded correction magnitude will be incorporated in future work.

$$RUL_{res} = RUL - RUL_{Phy} \tag{7}$$

3.2.3. Candidate Regression Models

Six candidate model architectures were trained and compared. The power modules were divided into 70% training and 30% test sets:

- **Multilayer Perceptron (MLP)**: a multilayer perceptron with two hidden layers (32, 16 units).
- **XGBoost**: Extreme Gradient Boosting with maximum tree depth 6, optimized on RMSE.
- **Linear Regression**: $\min \sum (y_i - \hat{y}_i)^2$.
- **Ridge Regression**: $\min \sum (y_i - \hat{y}_i)^2 + \alpha \sum \beta_i^2$.
- **Lasso Regression**: $\min \sum (y_i - \hat{y}_i)^2 + \alpha \sum |\beta_i|$.
- **Quadratic Response Surface Model (QRSM)**: a second-order polynomial expansion of the input features, including squared and pairwise interaction terms.

3.2.4. Validation Method

All these models are validated based on the K-Folding method (5 folds) and compared for 3 main Key Performance Indicators (KPIs) (Eq. 5, Eq. 8 and Eq. 9) that are explained in the result section in detail.

- **MAE**: Mean Absolute Error (Eq. 8).
- **RMSE**: Root Mean Squared Error (Eq. 9).
- $\mathbf{R^2}$: Coefficient of determination (Eq. 5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{9}$$

4. RESULT

4.1. Modeling Results

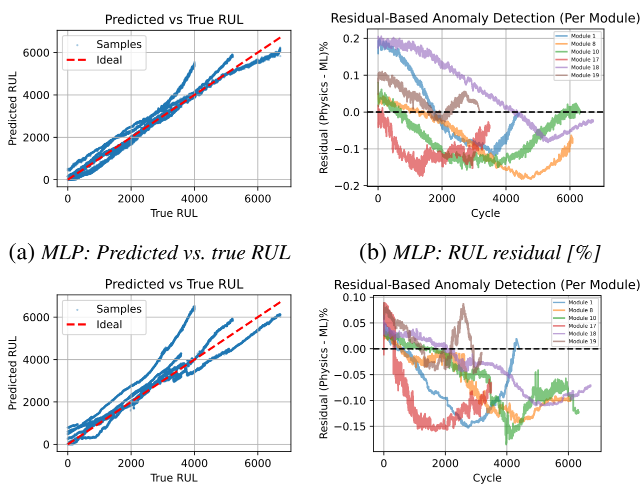
Table 3 summarizes the hold-out test performance for each model, evaluated using KPIs: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R^2 (Eq. 5, Eq. 8 and Eq. 9).

On this single 70/30 hold-out split, the MLP achieves the lowest error and highest R^2 by a clear margin, while XGBoost performs worst of all six models, notably worse than even plain linear regression. The relationship between the PCA-derived health indicator and RUL is smooth and continuous over a module's operating life, with curvature that becomes more dropped only near end of life. The MLP, as a universal function approximator, can fit this smooth nonlinear trajectory directly. The quadratic response-surface model captures the same curvature through explicit second-order terms, which explains why it outperforms the strictly linear models (linear regression, ridge, lasso) that can only fit a straight-line trend and therefore systematically underestimate degradation acceleration near failure. By contrast XGBoost, partitions the feature space into piecewise-constant leaf regions, does not contain enough distinct degradation trajectories (with only 20 power modules) for the boosted trees to interpolate smoothly between them, so XGBoost's step-wise predictions unexpectedly show poorly to the specific unseen module(s) in this hold-out split. By changing 70/30 hold-out split on dataset, the XGBoost performs differently, which confirms the reason of its strange results; therefore, in the 5-fold cross-validation results, with 5 sets hold-out splits, the mean values of these results are more meaningful.

Table 3. Hold-Out Test Performance of Candidate RUL Models

Model	MAE (cycles)	RMSE (cycles)	R^2
MLP	270.79	367.01	0.9397
XGBoost	473.18	685.59	0.7897
Linear Regression	426	634	0.8201
Ridge	424.56	640.13	0.8167
Lasso	426	634	0.8201
Quadratic	421.70	563.35	0.8580

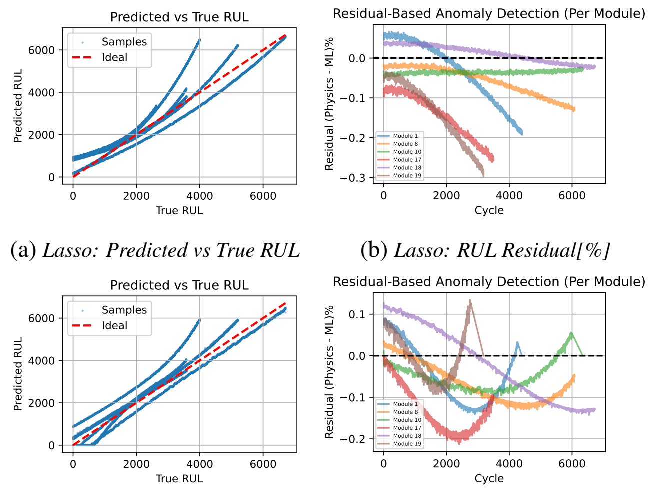
The initial results, shown in Fig. 8, according to the initial dataset split (70/30) demonstrate that among the more complex models (MLP and XGBoost), the MLP is closer to the 45-degree line in the predicted vs true RUL figure and lower residual RUL values; therefore, MLP achieves better raw hold-out accuracy in this split. These results from figures confirm what has been presented and discussed in Table 3.



(a) *MLP: Predicted vs. true RUL*
(b) *MLP: RUL residual [%]*
(c) *XGB: Predicted vs. true RUL*
(d) *XGB: RUL residual [%]*

Figure 8. Test performance of the complex RUL models.

Among the simpler models (linear regression and quadratic RSM), shown in Fig. 9, which are more embedding-friendly models for deployment on an automotive microcontroller, the quadratic response surface model (QRSM) is closer to 45-degree line in predicted vs true RUL figure and lower residual RUL values; therefore, the QRSM outperforms the linear, ridge and lasso variants, achieving the lowest RMSE and highest R^2 of the simpler-model group.



(a) *Lasso: Predicted vs True RUL*
(b) *Lasso: RUL Residual[%]*
(c) *QRSM: Predicted vs True RUL*
(d) *QRSM: RUL Residual[%]*

Figure 9. Test performance of simpler RUL models

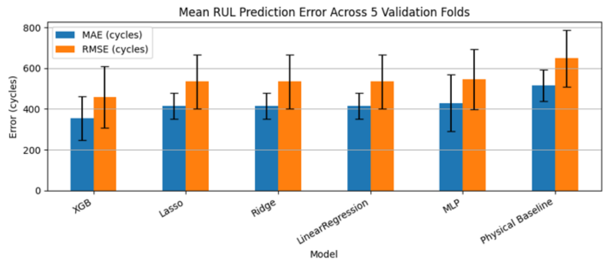
4.2. Models Validation

The quality of the models was assessed using three complementary methods: (i) standard regression metrics (MAE, RMSE, R^2 mean and standard deviation), (ii) 5-fold cross-validation and (iii) engineering judgment on the physical plausibility of predicted RUL trajectories. All preprocessing, training, and validation are performed with strict module-level grouping: each fold contains complete power modules, and no cycle-level samples from a given module appear simultaneously in both training and test sets. This avoids leakage between highly correlated measurements from the same unit. Table 4 and Fig. 10 report the mean and standard deviation of each metric in five folds.

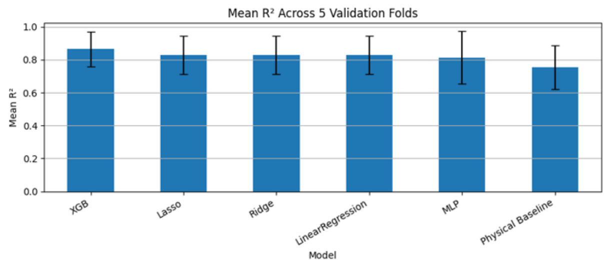
Table 4. 5-Fold Cross-Validation Results

Model	MAE (std)	RMSE (std)	R^2 (std)
MLP	409.8 (132.0)	515.9 (133.5)	0.828 (0.118)
XGBoost	352.6 (95.2)	454.3 (132.2)	0.868 (0.091)
Linear Regr.	415.0 (56.5)	534.3 (118.9)	0.828 (0.103)
Ridge	415.0 (56.5)	534.3 (118.9)	0.828 (0.103)
Lasso	414.9 (56.5)	534.3 (118.9)	0.828 (0.103)
Quadratic	357.4 (114.9)	440.9 (128.5)	0.869 (0.090)

In cross-validation, XGBoost and the quadratic model achieve the best and most consistent performance (highest mean R^2, competitive MAE/RMSE and among the lowest variance between folds). This suggests that, despite the MLP's advantage on a single hold-out split, XGBoost and the quadratic response-surface model are the most robust and most readily implementable candidates for embedded deployment. From an embedded perspective, the XGBoost



(a)



(b)

Figure 10. Model K-Fold validation: (a) MAE & RMSE, (b) R^2.

and QRSM models used in this study are compact: typical configurations require on the order of tens of kilobytes of memory and can be evaluated within a few hundred microseconds on standard automotive microcontrollers. Real-time rainflow counting and thermal-network updates add computational overhead, but remain compatible with current inverter-control cycles; a detailed deployment study is left for future work.

The regression coefficients for linear-family models offer physical interpretability for engineering judgment. For example, in Table 5, in the linear regression model, the negative coefficients of $V_{DS,on}$, R_{th}, T_j and ΔT_j show their negative effect on RUL when they increase and their magnitudes of these quantities indicate their effectiveness (weak or strong) on RUL which are fully interpretable and physically correct (Section 2).

Table 5. Regression Coefficients by Feature

Feature	LinearRegression	Ridge	Lasso
Vce_on	-115.773	-61.468	-99.939
Vce_on_ma	-15.873	-59.392	-0.007
Rth	-30.766	81.934	50.663
Rth_ma	266.988	101.590	362.555
Tj	-46.323	66.642	-104.745
Tj_ma	242.036	86.446	165.738
Delta_Tj	-58.496	-3.972	-18.177
Delta_Tj_ma	57.358	5.913	101.322
Damage	-211.433	-122.474	-194.788

Coefficients are interpreted with respect to standardized input features; absolute magnitudes should be read in that context.

4.3. Field Data Scenarios and Online Tracking

4.3.1. Field Data Scenarios

To evaluate the hybrid model under realistic field conditions, two synthetic scenarios designed to emulate field-like degradation trajectories were considered:

- Scenario 1: *physics-pessimistic* scenario, in which the field like RUL is better than the physics-based RUL estimate
- Scenario 2: *physics-optimistic* scenario, in which the physics-based RUL is better than the field like RUL

Based on inverter field experience, Scenario 1 (the field RUL is better than the physics-based RUL estimate) is estimated to be valid approximately 90% of the time and a held-out test power modules were used in the field for online tracking evaluation in both scenarios.

4.3.2. Online RUL Tracking and Anomaly Detection

In this section, the hybrid RUL estimate was tested for each candidate model to evaluate if the model can track the online field like RUL against the physics-only RUL over the test module's operating life.

In Scenario 1 for complex models, illustrated in Fig. 11, XGBoost, which also has the best validation performance, adapts to the field like RUL trajectory more accurately than MLP.

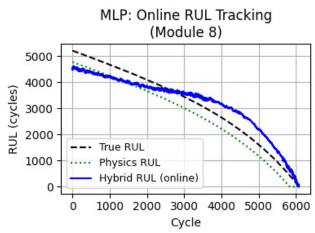
An anomaly is identified when the hybrid-physics RUL residual deviates by more than $\pm 3\sigma$ from its healthy baseline (μ_r), which has been formulated in Eq. 10 and Eq. 11 (Montgomery, 2020)(Brown, 2011). The healthy baseline is computed over the first 40% of module life when modules work without significant deviations in the specifications. This residual-based monitoring scheme allows the hybrid model to flag deviations from expected degradation behavior in addition to producing a corrected RUL estimate, which is valuable both for early fault detection and for triggering model re-calibration.

$$|r(t) - \mu_r| > 3\sigma \tag{10}$$

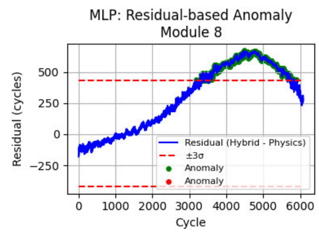
$$r(t) = RUL_{Hybrid}(t) - RUL_{Phy}(t) \tag{11}$$

In addition, shown in Fig. 11, XGBoost has had better performance for anomaly detection with a lower σ value and faster detection.

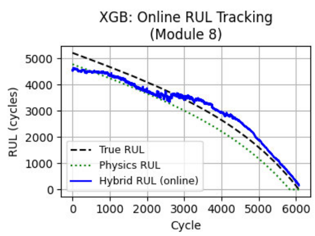
In Scenario 1 for simpler models, presented in Fig. 12, the quadratic model provides the best adaptation speed and stable anomaly detection although the σ value is higher due to its fast adaptation to the field like RUL. linear regressions, by contrast, adapt too slowly and trigger a disproportionately high number of anomaly flags although the σ value is lower due to its slow tracking.



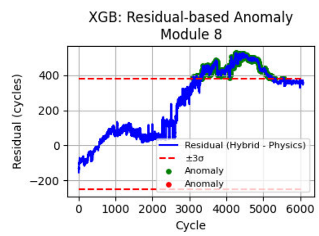
(a) *MLP: RUL tracking*



(b) *MLP: Anomaly detection*

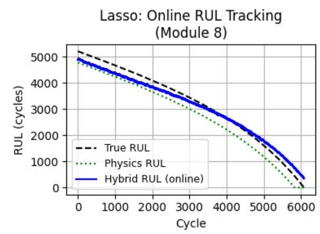


(c) *XGB: RUL tracking*

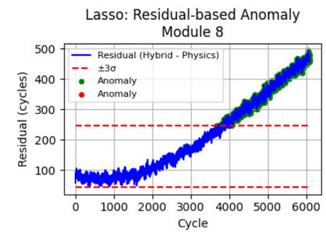


(d) *XGB: Anomaly detection*

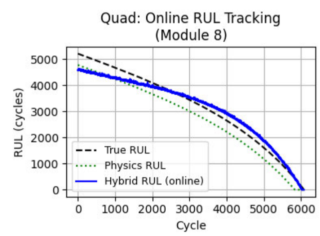
Figure 11. Hybrid model field like RUL tracking and anomaly detection



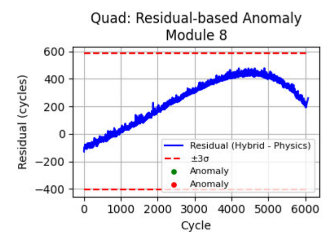
(a) *Lasso: RUL tracking*



(b) *Lasso: Anomaly detection*



(c) *QRSM: RUL tracking*



(d) *QRSM: Anomaly detection*

Figure 12. Hybrid model field like RUL tracking and anomaly detection

Scenario 2 follows a similar but less critical pattern and is omitted for brevity. These scenarios are constructed from the available power-cycling data and physics model; they do not represent independent fleet measurements, and therefore the conclusions regarding field robustness and safety relevance should be interpreted as indicative rather than definitive.

5. CONCLUSION & FUTURE WORK

5.1. Conclusion

The physics-based Coffin-Manson approach, while physically interpretable and useful for design-stage validation, is isolated from real-world degradation and computationally heavy. This approach does not offer a mechanism to incorporate measured field degradation and, as summarized in Table 2, does not provide online estimation capability at all.

A purely data-driven approach, while adaptive and microcontroller low-cost implementable, is limited by the availability of high-quality training data, poor extrapolation behavior, and non-interpretable or physically unrealistic predictions.

These two approaches make them insufficiently reliable and safe for driveline use on its own, particularly in battery electric vehicles (BEVs) therefore a new mixed approach is necessary.

The hybrid model proposed in this paper that combines a physics-based RUL baseline with a machine-learned residual correction addressed both limitations simultaneously. Reducing the five raw degradation-sensitive features to a single PCA health indicator (PC1) retained 71.7% of the variance of the features (88.6% jointly with PC2), confirming that the underlying degradation process is effectively low-dimensional and can be tracked with a compact, physically motivated indicator rather than a black-box high-dimensional model.

In this study, the MLP, the most flexible model tested, achieved the highest single-split accuracy ($R^2 = 0.940$, Table 3) but the lowest and least stable cross-validated accuracy (mean $R^2 = 0.828$, the highest fold-to-fold standard deviation of all six models) in 5-fold cross-validation (Table 4). These results indicates that with limited availability of power-module data (high-quality training data), a purely data-driven model risks overfitting to a specific train/test split rather than learning a generalizable degradation trend.

This study shown that among the six candidate residual-correction models, XGBoost and the quadratic response-surface model achieved the best and most consistent cross-validated performance (mean $R^2 = 0.868$ and 0.869, respectively, with the lowest fold-to-fold standard deviation in folds), making them the preferred choice for embedded deployment over the other models with higher standard deviations and lower mean R^2.

In online tracking under the physics-pessimistic field scenario, the residual-based anomaly detector (Eq. 10, Eq. 11, $\pm 3\sigma$ threshold) successfully flagged deviations between the hybrid and physics-only RUL trajectories, demonstrating that the hybrid architecture retains a usable safety-relevant monitoring capability that a pure regression model does not inherently provide.

Taken together, these results support the central claim of this paper: a physics-anchored, ML-corrected hybrid architecture is both more robust than a purely data-driven model, especially under limited training data, and more adaptive than a purely physics-based model, while remaining lightweight enough (via XGBoost or the quadratic model) for realistic deployment on a traction inverter microcontroller.

5.2. Future Work

Several directions are planned to extend this work. First, in this work, we focus on comparing candidate residual learners (MLP, XGBoost, linear, ridge, lasso, and quadratic response-surface models) under a fixed hybrid formulation. A full ablation against a physics-only baseline, a direct ML RUL model, and simple linear physics–ML blends is left for future work.

Moreover, we intend to replace the current residual-correction formulation with an explicit hybrid blending scheme.

$$RUL_{hybrid} = (1-\alpha)\cdot RUL_{Phy} + \alpha \cdot RUL_{ML}, \quad (12)$$

allowing the relative trust placed in the physics-based and ML-based estimates to be tuned or learned. Second, we plan to explore non-linear blending functions, such as an exponential form,

$$RUL_{hybrid}(x) = \alpha - \beta \cdot e^{\lambda x}, \quad (13)$$

which may better capture the accelerating degradation observed near end of life. Third, we intend to validate the hybrid model against real field data from a larger selected fleet of field vehicles rather than only the data from the test vehicle data. Finally, we plan to deploy the lightweight validated candidate models (quadratic and XGBoost) on an available embedded automotive microcontroller and send inverter status over the air (OTA) to monitor the real-time inverter state of health and optimize maintenance.

ACKNOWLEDGMENT

The authors thank their respective institutions for supporting this work. This research was supported by KK Synergy.

NOMENCLATURE

T_j	junction temperature
$T_{j,mean}$	mean junction temperature
$T_{j,max}$	maximum junction temperature
$T_{j,a}$	ambient (heatsink) temperature
ΔT_j	junction temperature swing
$T_{j,ref}$	reference junction temperature
V_{CE}/V_{DS}	collector-emitter/drain-source voltage
$V_{DS,on}$	on-state drain-source voltage
R_{th}	thermal resistance
P_{diss}	dissipated power
N_f	number of cycles to failure
A_F	acceleration factor
E_g	activation energy
k	Boltzmann constant
α	Coffin-Manson exponent weight
D	cumulative fatigue damage
n_i	number of cycles at stress level i
N_i	allowable cycles at stress level i
RUL	remaining useful life
RUL_{Phy}	physics-based RUL
RUL_{ML}	ML-based RUL
RUL_{Hybrid}	hybrid RUL estimate
RUL_{res}	residual RUL
$PC1, PC2$	principal components
X	input feature matrix
β	regression coefficient vector
Y	target output vector
ϵ	model error term
$y, \hat{y}$	true and predicted output
$\bar{y}$	mean of true output
R^2	coefficient of determination
MAE	mean absolute error
$RMSE$	root mean squared error
$r(t)$	hybrid-physics RUL residual
μ_r	mean residual
σ	standard deviation of residual

REFERENCES

Bayerer, R., Herrmann, T., Licht, T., Lutz, J., & Feller, M. (2008). Model for power cycling lifetime of igbt modules-various factors influencing lifetime. In *5th international conference on integrated power electronics systems* (pp. 1–6).

Bhoi, S. K., Chakraborty, S., Aghdaei, A., Mathew, A., Horn, T. D., Liebig, S., … Hegazy, O. (2025). A streamlined test-oriented lifetime assessment framework for automotive power electronics converters. *IEEE Transactions on Power Electronics*, *41*(4), 5862–5883.

Brown, . e. a. (2011). Power curve analytic for wind turbine performance monitoring and prognostics. In *Annual conference of the prognostics and health management society.*

Di Nuzzo, G., Lewitschnig, H., Tuellmann, M., Rzepka, S., & Otto, A. (2023). A data-driven condition monitoring method to predict the remaining useful life of sic power modules for traction inverters. In *2023 ieee international conference on prognostics and health management (icphm)* (pp. 312–319).

Di Nuzzo, G., Tuellmann, M., Methfessel, T., & Rzepka, S. (2022). Condition monitoring indicators for si and sic power modules. *Microelectronics Reliability*, *138*, 114614.

Eng, M. P., Lövberg, A., Misiorny, M., Vermelin, W. S., Brinkfeldt, K., & Mishra, M. (2023). A simple hybrid model for estimating remaining useful life of sic mosfets in power cycling experiments. In *Phm society asia-pacific conference* (Vol. 4).

Fang, X., Lin, S., Huang, X., Lin, F., Yang, Z., & Igarashi, S. (2018). A review of data-driven prognostic for igbt remaining useful life. *Chinese Journal of Electrical Engineering*, *4*(3), 73–79.

Farhadi, M., Vankayalapati, B. T., Sajadi, R., & Akin, B. (2023). Ac power cycling test setup and condition monitoring tools for sic-based traction inverters. *IEEE Transactions on Vehicular Technology*, *72*(10), 12728–12743.

Ismail, A., Saidi, L., Sayadi, M., & Benbouzid, M. (2020). A new data-driven approach for power igbt remaining useful life estimation based on feature reduction technique and neural network. *Electronics*, *9*(10), 1571.

Kumar, S., Allard, B., Hologne-Carpentier, M., Clerc, G., & Auger, F. (2025). *Review of prognosis approaches applied to power SiC MOSFETs for health state and remaining useful life prediction.* (Preprint)

Li, Z., Yue, H., Zhang, C., Dai, W., Guo, C., Li, Q., & Zhang, J. (2024). Fatigue life prediction of 2024-t3 Al alloy by integrating particle swarm optimization–extreme gradient boosting and physical model. *Materials*, *17*(21), 5332.

Montgomery, D. (2020). *Introduction to statistical quality control*. John Wiley & Sons, Incorporated.

Morel, C., & Morel, J.-Y. (2024). Power semiconductor junction temperature and lifetime estimations: A review. *Energies*, *17*(18), 4589.

Nyangon, J. (2026). Hybrid physics-informed artificial intelligence for high-fidelity modeling and optimization of electrical systems. *Frontiers in Artificial Intelligence*, *9*, 1751785.

Pino, C., Sitta, A., Castagnolo, G., Messina, A. A., Coffa, S., Calabretta, M., ... others (2023). Intelligent traction inverter in next generation electric vehicles: The health monitoring of silicon-carbide power modules. *IEEE Transactions on Intelligent Vehicles*, *8*(12), 4734–4753.

Rahat, M., Kharazian, Z., Mashhadi, P. S., Rögnvaldsson, T., & Choudhury, S. (2023). Bridging the gap: A comparative analysis of regressive remaining useful life prediction and survival analysis methods for predictive maintenance. In *Proceedings of the asia pacific conference of the prognostics and health management society.*

Rahat, M., Mashhadi, P. S., Nowaczyk, S., Rögnvaldsson, T., Taheri, A., & Abbasi, A. (2022). Domain adaptation in predicting turbocharger failures using vehicle's sensor measurements. In *Proceedings of the european conference of the prognostics and health management society* (Vol. 7, pp. 432–439).

Revanur, V., Ayibiowu, A., Rahat, M., & Khoshkangini, R. (2020). Embeddings based parallel stacked autoencoder approach for dimensionality reduction and predictive maintenance of vehicles. In *Iot streams for data-driven predictive maintenance and iot, edge, and mobile for embedded machine learning* (pp. 127–141). Springer.

Semikron Danfoss. (2024). *Shower power.* `https://www.semikron-danfoss.com/`. (Accessed: 2026-07-11)

Wu, Q., Xu, B., Xiao, L., & Wang, Q. (2024). A remaining useful life prediction method of sic mosfet considering failure threshold uncertainty. *IET Power Electronics*, *17*(12), 1594–1606.

Wu, W., Gu, Y., Yu, M., Gao, C., & Chen, Y. (2023). Remaining useful lifetime prediction based on extended kalman particle filter for power sic mosfets. *Micromachines*, *14*(4), 836.

BIOGRAPHIES

Kooros Moabber received the M.S. degree in Communication Systems Engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2014. He is currently a Principal System Architect and Electronics Engineer with Volvo Cars, Gothenburg, Sweden, where he leads the technical and system architecture of electrical and electronic modules, including traction inverters, on-board chargers, and DC-DC converters, and serves as functional safety lead engineer. He has authored several conference publications and holds multiple U.S. patents in EMI filtering and multi-level inverter design. His research interests include power electronics, functional safety, and ML-based predictive maintenance. Mr. Moabber is a Member of the IEEE and a Principal Partner of the European Center for Power Electronics (ECPE).

Hamed Talebi is a researcher at the Department of Computer Science and Media Technology, Malmö University, Sweden. He holds a B.Sc. in Software Development, an M.Sc. in Software Engineering from Blekinge Institute of Technology, and an M.Sc. in Internet of Things from Malmö University. His research interests include prognostics and health management, predictive maintenance, machine learning, edge artificial intelligence, computer vision, and intelligent monitoring systems. His current work focuses on data-driven and hybrid approaches for remaining useful life estimation, condition monitoring, and the deployment of intelligent systems on resource-constrained edge devices.

Mahmoud Rahat is a Senior Lecturer in Machine Learning at Halmstad University, Sweden. He has many years of experience as a researcher and developer in both academia and industry. He has collaborated with companies and industrial partners in the domains of predictive maintenance, intelligent mobility, and healthcare. From 2015 to 2016, he was a visiting research scholar at Northern Illinois University. Since 2019, he has been a member of the Center for Applied Intelligent Systems Research (CAISR) at Halmstad University. His research interests include machine learning, robotics, natural language processing, and predictive maintenance.