Physics-Informed Condition Monitoring of SiC Power Modules

Mattia Scarpa[1,2,4], Evgeny Kusmenko[2], Francesco Toso[4], Mattia Bruschetta[1], Ruggero Carli[1], and Simon Achatz[3]

[1] *Department of Information Engineering, University of Padova, Padova, Italy*
scarpamatt@dei.unipd.it
carlirug@dei.unipd.it
bruschet@dei.unipd.it

[2] *Infineon Technologies Dresden AG & Co. KG, Dresden, Germany*
Evgeny.Kusmenko@infineon.com

[3] *Infineon Technologies AG, Munich, Germany*
Simon.Achatz@infineon.com

[4] *Newtwen S.r.l., Padova, Italy*
francesco.toso@newtwen.com

Abstract

Silicon carbide (SiC) power modules are increasingly deployed in automotive traction inverters, where reliable condition monitoring is essential to prevent in-service failures and reduce maintenance costs. Despite extensive qualification procedures standardized under AQG 324, no consolidated approach exists for in-field health state estimation of SiC devices. Existing methods range from physics-of-failure lifetime models, which lack real-time applicability, to purely data-driven architectures that require large labeled datasets and offer limited generalization, to physics-informed machine learning frameworks that, while promising, remain computationally demanding for embedded deployment.

This work addresses condition monitoring of SiC MOSFET power modules assembled with sintered packaging technology, which prevents solder degradation and thereby produces aging behavior qualitatively distinct from previously studied devices. In solder-based modules, the forward voltage drop V_{DS} follows smooth quasi-exponential trajectories driven by progressive solder delamination. With sintered packaging this mechanism is suppressed, and V_{DS} instead exhibits multi-regime degradation profiles; modules are additionally subject to wirebond liftoff events that introduce abrupt, non-monotonic perturbations directly onto V_{DS}, posing new challenges for condition monitoring approaches.

To address these challenges, we propose a condition monitoring framework integrating three complementary elements. First, physics-informed feature engineering replaces raw sensor signals with physically grounded inputs, including cumulative damage indicators derived from junction temperature swing and mean junction temperature, combined with a Miner rule accumulator, which encode degradation history in a form directly interpretable by the model. Second, a monotonicity constraint enforced via gradient penalty regularization embeds the expected degradation direction as a physics-guided prior, improving generalization and cross-validation stability. Third, the model output is parameterized as a heavy-tailed distribution rather than a point estimate, providing calibrated uncertainty quantification inherently robust to the out-of-distribution variance introduced by liftoff events.

The framework is evaluated on an industrial power cycling dataset provided by Infineon Technologies, acquired under multiple operating conditions. Multiple neural network architectures of varying complexity and sequential context are compared within a rigorous cross-validation protocol. The full physics-informed feature set combined with gradient penalty regularization consistently outperforms purely data-driven baselines, reducing mean absolute error by approximately 70% and maintaining stable performance across all cross-validation folds, including conditions where baseline models degrade significantly. These results confirm that integrating physical prior knowledge at the feature, constraint, and output distribution levels is both necessary and sufficient to handle the complexity of SiC power module degradation, while remaining lightweight enough for practical embedded deployment.

1. INTRODUCTION

Silicon carbide (SiC) power metal-oxide-semiconductor field-effect transistors (MOSFETs) are now widely deployed in the high-voltage stage of automotive traction inverters, where their higher switching frequency and elevated junction-temperature limits allow for smaller passive components and higher converter efficiency. Reliability of the power-semiconductor stage has long been identified by industry as one of the most pressing reliability concerns among the components that make up a power converter (Yang et al., 2011), and accelerated power-cycling tests under the European Center for Power Electronics (ECPE) Automotive Qualification Guideline (AQG) 324 qualification (ECPE European Center for Power Electronics e.V., 2025) provide the bridge between bench characterization and field deployment. The in-service condition monitoring (CM) problem is distinct from the qualification one: rather than counting cycles to a fixed threshold under a known stress profile, on-board electronics must keep an updated health estimate from electrical measurements taken while the module is operating under variable load. Reviews of SiC reliability converge on the dominant failure modes, but the electrical signals proposed for tracking those modes online are still treated as candidates rather than as a consolidated indicator set, and their combination into a per-cycle health estimate consistent with the lifetime models obtained from accelerated tests is identified as a central open methodological problem (Ni et al., 2020).

The chip-solder degradation mechanism has long been recognized as the leading source of package-level failure in conventional power modules under thermo-mechanical cycling (Clech, 1997; Kovacevic et al., 2010; Kovacevic-Badstuebner et al., 2015), with a physical signature on the on-state voltage drop that is well described as a smooth quasi-exponential rise driven by progressive crack growth in the die-attach. This profile contrasts sharply with the abrupt step-like behavior of wirebond fatigue, where each individual bond-wire lift-off has been observed to produce a stepwise increase in the on-state voltage drop (Kumar et al., 2026). For this reason, data-driven SiC CM has predominantly grown on modules where solder degradation is the dominant failure mode, and has gradually drifted away from the closed-form parametric route that fits an electro-thermal indicator to a regression law and inverts it to recover a lifetime fraction (Di Nuzzo et al., 2022, 2023). The analytical transparency of that route comes at the cost of a tight coupling to the assumed indicator shape, and more flexible machine-learning formulations have also been explored in parallel, including sequence models that forecast the relative on-state voltage drop forward to the qualification threshold (Olschewski et al., 2025; Villalobos et al., 2025) and end-to-end snapshot models that map the current measurement directly to an expected lifetime fraction (Achatz et al., 2026). The physics of failure (PoF) tradition that underpins lifetime modeling in this domain combines Coffin–Manson estimates with Miner's rule to handle variable stress profiles (Ceccarelli et al., 2019; Barbagallo et al., 2021), and physics-informed machine learning (PIML) frameworks have started to bridge it with data-driven approaches (Fassi et al., 2024). From a broader prognostics and health management (PHM) standpoint, three persistent issues have been repeatedly flagged in the literature: data-driven prognostics are typically evaluated on training-distribution-matched test sets, the domain shift between accelerated tests and field operation is seldom assessed, and predictive uncertainty is rarely reported alongside the point estimate (Zio, 2022; Fink et al., 2020), and these gaps are confirmed to remain open in our SiC domain (Kumar et al., 2026).

In this work we examine an accelerated dataset whose degradation pattern departs from those most frequently analyzed in the literature. The modules under test are 1.2 kV automotive SiC half-bridges assembled with a sintered silver die-attach technology, which suppresses solder degradation as a failure mode and leaves wirebond fatigue as the dominant aging mechanism. On the on-state voltage drop, this manifests as a continuous drift driven by progressive crack propagation in the bond-pad metallization, with rapid rises that occurs stocastically whenever a wire lifts off. Four stress conditions, obtained by combining heating durations of 1.5 s and 30 s with different load-current levels, all exhibit this profile, to get module-to-module variability within each condition.

To handle this setting, we explore three complementary additions to a standard data-driven baseline, one acting on the input, one on the training loss, and one on the output. On the input side, the raw indicator vector is augmented with physics-informed features that summarize cumulative thermo-mechanical stress along the module lifetime: the per-cycle junction-temperature swing and mean junction temperature, their integrals over time, and a Miner-rule damage term obtained from a per-cycle Coffin–Manson estimate. The Miner-rule construction is the canonical physics-of-failure aggregator for SiC under realistic mission profiles (Ceccarelli et al., 2019; Barbagallo et al., 2021); here it enters the network as a learned input rather than as the prediction itself. On the loss side, a monotonicity prior is added as a gradient-penalty regularizer on the features that are physically expected to grow with degradation in line with the loss-based formulation used in our earlier work (Scarpa et al., 2025). Compared with approaches that enforce monotonicity by construction, through custom activation functions inside the network (Zhao & Fink, 2025), this soft formulation introduces no inference overhead, remains compatible with simple feed-forward networks, and matches the constraints of embedded deployment. On the output side, the network produces the parameters of a Student-t distribution rather than a single point estimate, so that location, scale, and degrees-of-freedom are learned jointly and the heavy tails accommodate the larger residuals that lift-off events produce in the regression target.

The framework is evaluated under a four-fold cross-validation stratified at the module level. Three neural architectures (MultiLayer Perceptron (MLP), Contextual MLP (CMLP), Neural ODE (NODE)) are trained with different input-feature configurations: a baseline using only the line current and the relative on-state voltage drop, and the same set augmented with the cumulative features and the monotonicity penalty. The combination of cumulative features and monotonicity penalty reduces the mean absolute error (MAE) on the lifetime estimate by up to 60% relative to the baseline, at a cross-fold standard deviation of 0.019 on the health index that holds also on the folds whose validation modules exhibit lift-off.

The remainder of this paper is structured as follows: Section 2 describes the failure physics and the indicator dataset. Section 3 introduces the feature pipeline, the loss formulation, and the output head. Section 4 details the experimental protocol. Section 5 reports the comparative evaluation. Sections 6 and 7 close with the discussion and the open questions.

2. BACKGROUND AND PROBLEM FORMULATION

2.1. Module and Power-Cycling Dataset

The dataset analyzed in this work consists of accelerated power-cycling tests performed at Infineon Technologies on a 1.2 kV automotive SiC half-bridge module. Its sintered silver die-attach removes solder fatigue at the chip–substrate interface as a failure mode, leaving wirebond fatigue at the chip-side bond pads as the dominant package-level mechanism. As thermal cycles accumulate, mechanical stress drives crack propagation in the bond-pad metallization, and individual wire lift-offs translate into the rapid rises of the on-state voltage drop described in Section 1.

Each device under test (DUT) is repeatedly heated by a load current I_{load} flowing for a heating duration t_{heat}, then cooled to the inlet coolant temperature over a duration t_{cool}. During each cycle, the per-cycle on-state voltage drop V_{DS}, the thermal resistance R_{th}, the junction-temperature swing $\Delta T_{\text{vj}} = T_{\text{vj,max}} - T_{\text{vj,min}}$, together with the load current and the heating/cooling durations, are recorded. Following the AQG 324 qualification standard (ECPE European Center for Power Electronics e.V., 2025), a module is considered failed when at least one of

$$\frac{\Delta V_{DS}}{V_{DS,\text{nom}}} \geq 5\%, \qquad \frac{\Delta R_{\text{th}}}{R_{\text{th,nom}}} \geq 20\% \tag{1}$$

is reached, where $V_{DS,\text{nom}}$ and $R_{\text{th,nom}}$ are the reference values measured on the pristine module. In the present dataset, crossing of the V_{DS} threshold is the primary end-of-life (EOL) event observed, while the thermal-resistance excursion remains bounded throughout the test.

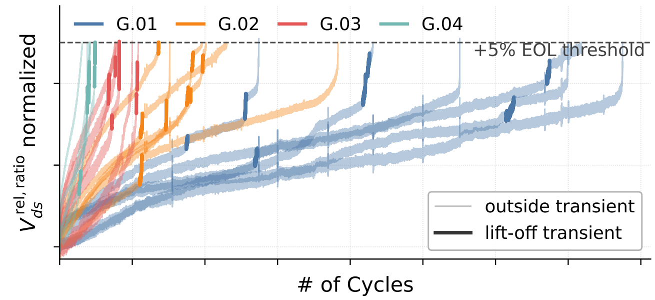

Figure 1. Relative on-state voltage drop of each DUT along its cycle counter. The light curves show the underlying drift; the solid-colour overlays mark the cycle ranges annotated as wire lift-off transients during data preprocessing. Numeric tick labels are suppressed; the dashed line marks the $+5\%$ EOL threshold.

Table 1. Operating conditions of the four power-cycling groups. Current and temperatures are reported as mean $\pm$ standard deviation over all recorded cycles of the six DUTs of the group, both switches pooled for $T_{\text{vj,max}}$; each group contains $N = 6$ DUTs.

Group	t_{heat} [s]	t_{cool} [s]	I_{load} [A]	$T_{\text{vj,max}}$ [°C]	T_{inlet} [°C]
G.01	1.5	3.5	694.5 ± 0.05	177.2 ± 1.2	73.3 ± 0.09
G.02	1.5	3.5	752.6 ± 0.05	178.2 ± 2.0	53.6 ± 0.02
G.03	30	30	669.5 ± 0.07	176.6 ± 1.6	74.1 ± 0.17
G.04	30	30	776.7 ± 0.07	176.1 ± 3.0	33.1 ± 0.06

Twenty-four DUTs are distributed across four operating-condition groups, summarized in Table 1. The groups span two heating-pulse durations (short, $t_{\text{heat}} = 1.5$ s, and long, $t_{\text{heat}} = 30$ s) and four load-current levels, with the inlet coolant temperature trimmed condition-by-condition to keep the junction-temperature peak close to $T_{\text{vj,max}} \approx 175\,°\text{C}$ across conditions. The short-pulse groups compress the same heating energy into a shorter cycle, inducing larger ΔT_{vj} excursions than the long-pulse groups. The same power-cycling protocol on a related solder-based SiC module has been previously analyzed (Olschewski et al., 2025); the experimental framework is therefore consistent across studies, and only the packaging technology differs.

Figure 1 reports the relative on-state voltage drop of each DUT along its cycle counter, with numeric tick labels intentionally suppressed and only the $+5\%$ EOL threshold marked. Most trajectories drift smoothly under the continuous crack propagation in the bond-pad metallization; on top of this drift, the cycle ranges in which wire lift-off transients have been manually identified during data preprocessing are overlaid in solid colour. Out of the 24 DUTs, 19 exhibit at least one such transient (27 events in total across the four groups).

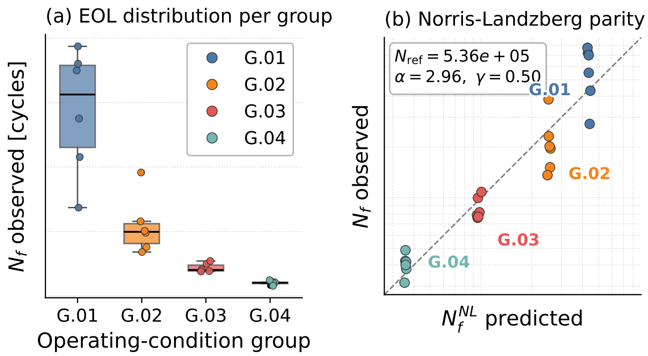

Figure 2. (a) Empirical EOL distribution per operating-condition group as per Table 2 on the dataset (six DUTs per group). (b) Parity plot of observed and NL-predicted cycles to failure for the same DUTs; the dashed line marks ideal agreement.

2.2. Lifetime Modeling and Damage Aggregation

The PoF framework that traditionally underpins lifetime estimation for power modules combines a Coffin–Manson-type estimate of cycles to failure under constant stress with Miner's rule for aggregating variable stress histories (Ceccarelli et al., 2019; Barbagallo et al., 2021). For wirebond-driven fatigue, we adopt the Norris–Landzberg (NL) parameterization, which expresses the cycles to failure under a constant operating point as

$$N_f^{\text{NL}}(\Delta T_{\text{vj}}, t_{\text{heat}}) = N_{\text{ref}} \left(\frac{\Delta T_{\text{vj}}}{100}\right)^{-\alpha} t_{\text{heat}}^{-\gamma}, \qquad (2)$$

where the three coefficients N_{ref}, α, γ are fitted to the EOL data of the four groups in Table 1. Log-linear least-squares fitting on the 24 DUTs yields

$$N_{\text{ref}} = 5.32 \times 10^5, \quad \alpha = 2.94, \quad \gamma = 0.50, \qquad (3)$$

with ΔT_{vj} averaged over the early-stable cycle window $[100, 600]$ to remove degradation-induced drift from the operating-point estimate. Figure 2(a) reports the empirical EOL distribution per group as a boxplot, and Figure 2(b) compares the observed cycles to failure of each DUT with the NL prediction evaluated at Eq. (3).

To aggregate variable stress histories at the DUT level, Miner's rule is applied cycle-by-cycle. For a module that has completed n cycles under stress profile $\{(\Delta T_{\text{vj},k}, t_{\text{heat},k})\}_{k=1}^{n}$, the accumulated damage is

$$D(n) = \sum_{k=1}^{n} \frac{1}{N_f^{\text{NL}}(\Delta T_{\text{vj},k}, t_{\text{heat},k})}, \qquad (4)$$

with $D = 1$ corresponding to the predicted end of life. Evaluated at the recorded EOL of each DUT, the Miner-damage distribution is summarized in Table 2.

Table 2. Miner-rule damage Eq. (4) evaluated at the recorded EOL of each DUT, reported per group and aggregated.

Experiment	n	$\overline{D}$	σ_D	CV
G.01	6	1.323	0.419	32%
G.02	6	0.855	0.324	38%
G.03	6	0.898	0.116	13%
G.04	6	1.132	0.146	13%
Global	24	1.052	0.338	32%

The global aggregate $\overline{D} = 1.052$ confirms that the NL fit is on average well-centred on the dataset. The per-group coefficient of variation (CV), however, ranges from 13% in the long-pulse groups (G.03, G.04) to 38% in the short-pulse group G.02, signalling that a substantial fraction of the module-to-module variability remains unexplained at the operating-condition level.

2.3. Problem Formulation

The NL/Miner framework provides a bulk estimate of the cycles to failure at the population level and a scalar damage index D at the DUT level, both convenient for accelerated-test characterisation. It does not provide a per-cycle estimate of the current health state, nor a quantification of predictive uncertainty around that estimate, nor a mechanism to track module-specific deviations from the empirical fit. These three quantities are the natural targets of an in-service CM tool.

We therefore introduce a dimensionless health index

$$r(n) = \frac{n}{N_f^{\text{EOL}}}, \qquad r \in [0,\, 1 + \varepsilon], \qquad (5)$$

where n is the current cycle count and N_f^{EOL} is the cycle index at which the EOL criterion Eq. (1) is first satisfied. By construction, $r = 0$ corresponds to a pristine module, $r = 1$ to the EOL event, and a small overshoot $\varepsilon > 0$ is allowed to absorb the discretisation of EOL detection. The CM task addressed in the next section is to estimate $r(n)$ from the per-cycle measurements available at cycle n, together with a predictive distribution that captures both module-to-module residual variability and the larger residuals produced by individual lift-off events.

3. PROPOSED FRAMEWORK

3.1. Network Architectures

We compare three different neural architectures that differ in how the cycle history is exposed to the input: a snapshot MLP, which uses only the current measurement; a contextual MLP (CMLP), which appends a short window of past on-state voltage values; and a NODE, which carries a hidden state across cycles. All three map a per-cycle input vector

$x_n \in \mathbb{R}^d$ to the parameters of the predictive distribution introduced in Section 3.2.

The MLP is the simplest of the three. It treats every cycle independently, mapping a single snapshot x_n through L fully-connected layers with non-linear activation (Rumelhart et al., 1986),

$$h_1 = \phi(W_1 x_n + b_1), \quad h_\ell = \phi(W_\ell\, h_{\ell-1} + b_\ell), \tag{6}$$

with the output read from the last layer provided as $[\mu_n, \tilde{\sigma}_n, \tilde{\nu}_n] = W_{\text{out}}\, h_L + b_{\text{out}}$. The stateless formulation is straightforward to train and fits comfortably within the parameter budget of an embedded target, but it is blind to the cycle history: any signal that requires comparing the current snapshot with the past must already be encoded in x_n.

The contextual MLP extends the input vector with a size limited rolling window of past on-state voltage measurements,

$$\tilde{x}_n = \left[\, x_n,\, V_{DS,n-K+1}, \ldots, V_{DS,n} \right] \in \mathbb{R}^{d+K}, \tag{7}$$

and feeds $\tilde{x}_n$ through an MLP with the same forward form as Eq. (6). This restores access to a finite local context without introducing a recurrent state. The trade-off is in the choice of K, which is fixed by hand and cannot be learned: anything outside the window is invisible to the model.

The NODE replaces the discrete stacking of layers with a continuous-time vector field on a hidden state (Chen et al., 2018),

$$\frac{dh(t)}{dt} = f_\theta\big(h(t),\, x(t)\big), \tag{8}$$

which is integrated forward in time by an explicit Euler step,

$$h_{n+1} = h_n + \Delta t_n \cdot f_\theta(h_n,\, x_n), \tag{9}$$

and produces the output by reading the hidden state, $[\mu_n, \tilde{\sigma}_n, \tilde{\nu}_n] = W_{\text{out}}\, h_n + b_{\text{out}}$. The step size Δt_n is set proportional to the per-cycle heating duration $t_{\text{heat},n}$, so that the dynamics learned by f_θ are decoupled from the operating-point-dependent cycle duration; the details of this scaling are deferred to Section 4. Compared with the previous two architectures, the NODE carries unbounded temporal memory in its hidden state and is the only one whose recurrence is conditioned explicitly on t_{heat}, at the cost of harder hyperparameter tuning. Training is performed on subsequences rather than on the full trajectory, so the hidden state at each subsequence start is set to zero; at inference, however, the state at any cycle results from the integration of the dynamics from cycle zero onward, and the zero-initialised training state is therefore inconsistent with the state actually reached during inference. To reduce this mismatch, we periodically run a full-trajectory simulation on the training set and use the resulting hidden state at each subsequence boundary as the new initialisation, in a manner analogous to (Scarpa et al., 2025). The refresh frequency and the subsequence spacing are reported in Section 4.

3.2. Probabilistic Output Head

In place of a Gaussian point-estimate trained against the mean squared error, the network produces the parameters of a Student-t distribution over the health index r,

$$p(r \mid x_n) = \mathcal{T}\big(\mu_n,\, \sigma_n,\, \nu_n\big), \tag{10}$$

where the location μ_n, the scale $\sigma_n > 0$ and the degrees-of-freedom $\nu_n > 0$ are all produced by the network for each cycle. Written out, the density is

$$p(r \mid x_n) = \frac{\Gamma\left(\frac{\nu_n+1}{2}\right)}{\Gamma\left(\frac{\nu_n}{2}\right)\sqrt{\pi\nu_n}\,\sigma_n} \left[1 + \frac{1}{\nu_n}\left(\frac{r-\mu_n}{\sigma_n}\right)^2\right]^{-\frac{\nu_n+1}{2}}, \tag{11}$$

which recovers the Gaussian in the limit $\nu_n \to \infty$ and grows heavier tails as ν_n decreases.

The three parameters are learned end-to-end minimizing the negative log-likelihood of Eq. (10),

$$\mathcal{L}_{\text{NLL}} = -\frac{1}{N}\sum_{n=1}^{N} \log p(r_n \mid x_n). \tag{12}$$

A Gaussian likelihood, equivalent to the mean squared error up to constants, penalises large residuals quadratically. In our setting the cycles in which an individual wire lift-off occurs are exactly the cycles where the residual on r is largest, so a quadratic loss amplifies the corresponding gradient contributions and tends to destabilise training. The Student-t replaces the quadratic tail with a $\log\big(1 + (r-\mu)^2/(\nu\sigma^2)\big)$ tail whose curvature decreases as ν shrinks. The learned ν_n then acts as a per-cycle reliability measure rather than a fixed hyperparameter, and the high-residual cycles contribute proportionally less to the gradient than they would under a Gaussian assumption.

3.3. Monotonicity Prior as a Soft Constraint

The health index $r(n) = n/N_f^{\text{EOL}}$ is by construction non-decreasing in n, and the cumulative features introduced in Section 2.3 are non-decreasing in the relevant inputs as well. We encode this physical prior as a gradient-penalty regulariser added to the training objective. For the stateless architectures (MLP and CMLP), the constraint is imposed pointwise as a non-negativity penalty on the partial derivative of the central estimate with respect to each cumulative input,

$$\mathcal{L}^{\text{pw}}_{\text{mono}} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i\in\mathcal{I}_{\text{cum}}} \left[-\frac{\partial \mu_n}{\partial x_{n,i}}\right]_+^2, \tag{13}$$

where $[\cdot]_+$ denotes the positive part and $\mathcal{I}_{\text{cum}}$ collects the indices of the physically monotone features. For the recurrent architecture (NODE), the same prior can also be imposed temporally on the predicted trajectory,

$$\mathcal{L}_{\text{mono}}^{\text{t}} = \frac{1}{N-1} \sum_{n=1}^{N-1} \Big[-(\mu_{n+1} - \mu_n) \Big]_+^2. \tag{14}$$

In both cases the training objective combines the two terms with a non-negative weight λ,

$$\mathcal{L} = \mathcal{L}_{\text{NLL}} + \lambda \, \mathcal{L}_{\text{mono}}, \tag{15}$$

and the prior acts on the loss only at training time, so the inference graph is left unchanged and no overhead is added at deployment.

4. EXPERIMENTAL SETUP

4.1. Input Features

Each architecture is fed a per-cycle vector x_n built from two groups of features. The *base* group contains the line current and the relative on-state voltage drop,

$$x_n^{\text{base}} = \big[I_{\text{load},n},\, V_{DS,n}^{\text{rel}} \big] \in \mathbb{R}^2, \tag{16}$$

which are the two quantities measured at every cycle in the power-cycling test. The *cumulative* group collects four physics-informed quantities that summarise the stress history up to cycle n,

$$\begin{aligned} S_{T_j}(n) &= \textstyle\sum_{k=1}^{n} \int_{\text{t}_k} T_{\text{vj,ss}}(\tau)\, d\tau, \\ S_{\Delta T_j}(n) &= \textstyle\sum_{k=1}^{n} \Delta T_{\text{vj},k}, \\ S_I(n) &= \textstyle\sum_{k=1}^{n} \int_{\text{t}_k} I_{\text{load}}(\tau)\, d\tau, \\ D(n) &= \text{Miner damage of } Eq. \text{ (4)}, \end{aligned} \tag{17}$$

all of which are by construction non-decreasing in n and form the set $\mathcal{I}_{\text{cum}}$ on which the pointwise monotonicity penalty of Eq. (13) acts. The baseline configuration uses only x_n^{base}; the full configuration appends Eq. (17), so that $x_n \in \mathbb{R}^6$.

4.2. Network Sizes and Training

The MLP and CMLP share a three-layer feed-forward network with ReLU activation and a hidden width of about 15 units per layer; the last layer outputs the three Student-t parameters $(\mu_n, \tilde{\sigma}_n, \tilde{\nu}_n)$. The CMLP appends $K = 10$ past values of V_{DS}^{rel} to the per-cycle input as in Eq. (7). The NODE uses a hidden state of comparable dimension and a two-layer MLP as the vector field f_θ. The hidden widths of all three networks are chosen so that, on the full feature set, the total parameter count remains below 10^3 trainable weights, keeping the three models within the same order of magnitude so that the comparison isolates the effect of the temporal context exposed to the input from the effect of model size. The integration step in Eq. (9) is

$$\Delta t_n = \frac{t_{\text{heat},n}}{\text{DT}_{\text{scale}}}, \tag{18}$$

which is the explicit mechanism that decouples the dynamics learned by f_θ from the per-group cycle duration: under the power-cycling protocol of Table 1 the heating duration varies by a factor of 20 between the short and the long groups, so that a constant integration step would force the network to absorb this variability into the weights rather than into the recurrence. We set $\text{DT}_{\text{scale}} = 1000$, the same value across all groups, so that Δt_n remains in the $[10^{-3},\, 10^{-2}]$ range throughout the dataset.

The Student-t scale and degrees-of-freedom are passed through a softplus to keep them strictly positive, $\sigma_n = \text{softplus}(\tilde{\sigma}_n) + \sigma_{\text{floor}}$ and $\nu_n = \text{softplus}(\tilde{\nu}_n) + \nu_{\text{min}}$, with $\sigma_{\text{floor}} = 0.02$ to prevent vanishing scale and $\nu_{\text{min}} = 2$ to keep the second moment of the heavy-tail likelihood finite. Optimisation uses Adam with a learning rate of 10^{-3} scheduled piecewise with a 0.1 decay, gradient norm clipped at 1.0, and a monotonicity weight of $\lambda = 10$ when the penalty is active. All three architectures are trained for 500 epochs with a batch size of $2{,}048$ samples; for the NODE the batch is composed of subsequences of $5{,}000$ cycles drawn from the training DUTs with a stride of 500 cycles, while the MLP and CMLP are fed batches of independent per-cycle snapshots. The full-trajectory simulation that refreshes the hidden-state cache at each subsequence boundary is re-run every 30 epochs. All networks are fed a $10\times$ subsampled version of the cycle counter to keep the per-epoch compute budget manageable on embedded-class hardware.

4.3. Cross-Validation Protocol and Baseline

The 24 DUTs are split with a four-fold cross-validation stratified at the operating-condition group level: each fold reserves one DUT per group for validation and uses the remaining 20 DUTs for training. The folds are mutually exclusive at the DUT level, so every DUT is used exactly once as validation across the four folds. Performance is reported as mean and standard deviation across folds.

The reference against which the proposed configurations are compared is the parameter-free estimator

$$\hat{r}_n^{\text{base}} = 20\, V_{DS,n}^{\text{rel}}, \tag{19}$$

obtained by inverting the 5% AQG 324 criterion of Eq. (1) into a linear map from $V_{DS}^{\text{rel}} \in [0, 0.05]$ to $r \in [0, 1]$. This estimator requires no training and uses only the indicator that the qualification standard already monitors. Any improvement of the proposed networks over Eq. (19) therefore quantifies the information contributed by the cumulative features and the monotonicity prior on top of the raw V_{DS} trace.

4.4. Evaluation Metrics

Four metrics are reported per fold, each computed over all validation cycles. The first two assess the central estimate $\hat{r}_n = \mu_n$ against the true health index r_n,

$$\text{MAE} = \frac{1}{N}\sum_{n=1}^{N} |\hat{r}_n - r_n|, \tag{20}$$

$$R^2 = 1 - \frac{\sum_n (r_n - \hat{r}_n)^2}{\sum_n (r_n - \bar{r})^2} \tag{21}$$

where $\bar{r}$ is the mean of r on the validation set. The remaining two are standard prognostic metrics that weigh the estimate against the residual life (Saxena et al., 2010),

$$\alpha\text{-acc} = \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}\big[\,|\hat{r}_n - r_n| \leq \alpha\,(1 - r_n)\,\big], \tag{22}$$

which measures the fraction of cycles inside an α-cone that narrows toward the EOL (we use $\alpha = 0.2$), and the relative accuracy

$$\text{RA} = 1 - \frac{|\hat{r}_n - r_n|}{r_n}, \tag{23}$$

averaged over the cycles in which $r_n \geq 0.1$ to keep the denominator away from zero in the pristine region.

5. RESULTS

5.1. Aggregate Cross-Validation Performance

The four metrics of Eqs. (20) to (23), aggregated as mean and standard deviation across the four cross-validation folds, are reported in Table 3. The trivial baseline of Eq. (19) sets a non-trivial floor that any learned estimator must clear. Training on the base features alone — the telemetric signals available from the inverter without any explicit history encoding — yields only a modest MAE improvement and keeps α-accuracy essentially at the level of the trivial reference. Appending the cumulative features of Eq. (17), which encode the stress history experienced by the module up to cycle n, translates the input into a sharper estimate: the gap closes on every metric and the recurrent NODE becomes consistently best. The monotonicity penalty of Eq. (13) adds a smaller increment — a few percentage points on α-accuracy and a fraction of a percent on MAE — consistent with the hypothesis that the cumulative features themselves already encode a substantial degree of physical monotonicity, partly carrying out the role of the explicit regularizer.

Beyond the best mean performance, the NODE also exhibits the lowest fold-to-fold variance on every metric, as shown by the standard deviation columns of Table 3. This stability is consistent with the recurrent hidden state acting as an implicit memory of the trajectory, which absorbs the cumulative-damage signal across regular cycles as well as across anomalous events such as the wirebond lift-off transients that disrupt $V_{DS,n}^{\text{rel}}$ on individual DUTs.

5.2. Health Index and Relative Accuracy Across Life

The aggregate behaviour of the four series along the life fraction $\tau = n/N_f^{\text{EOL}}$ is reported in Figure 3, with one column per architecture and one row per metric. The top row overlays the ideal $\hat{r} = \tau$ diagonal and the $\alpha = 20\%$ cone of Eq. (22); the bottom row marks the RA $= 1$ ideal.

The trivial baseline, drawn in gray in all six panels, captures the slow V_{DS} drift correctly in mid-life but lags below the diagonal for the first three quarters of the trajectory and only catches up close to the EOL threshold. As a direct consequence its RA dives below zero for $\tau \in [0.05,\, 0.15]$: in this regime $V_{DS,n}^{\text{rel}} \approx 0$ on most modules, so $\hat{r}$ is comparable in magnitude to the absolute error itself and the denominator of Eq. (23) is too small to absorb it. This regime is precisely where a learned estimator has the most to add: from $\tau = 0.2$ onward all three architectures stay above RA $= 0.8$ on the median curve, while the baseline only recovers to RA ≈ 0.7 around $\tau = 0.5$.

The base configuration (blue curve) is the one in which the choice of architecture matters most. On the MLP panel the base median strays visibly outside the cone for $\tau \lesssim 0.2$ and shows the widest 10-90 band of the three columns, consistent with the largest cross-fold variance reported in Table 3. The CMLP corrects part of this through the past-V_{DS} context window; the NODE closes most of the gap through the recurrent state alone, with the base median already inside the cone almost everywhere.

The cumulative variants (orange and green) compress these differences across architectures: once S_{T_j}, $S_{\Delta T_j}$, S_I and D are available, the three median curves essentially overlap. The monotonicity penalty (green vs. orange) brings a slight visible tightening of the band on the CMLP and NODE panels for $\tau > 0.3$ but is barely distinguishable on the MLP. The bands of the NODE cumulative variants are the narrowest among the nine network configurations, consistent with the lowest cross-fold variance of Table 3: the recurrent architecture pairs the best central tendency with the most consistent inter-DUT response.

5.3. Predictive Uncertainty Calibration

We check whether the Student-t output of Eq. (10) delivers calibrated uncertainty. For each validation cycle we form the Student-t central interval at nominal coverage $p \in \{68.27\%,\, 95.45\%\}$ from the predicted (σ_n, ν_n) and count the fraction of cycles in which the true r_n falls inside that interval. Table 4 reports the result. The base NODE is undercovered at both levels: with only the raw electrical inputs the network reports tighter bands than the residuals justify.

Table 3. Cross-validation metrics on the dataset: mean ± standard deviation across the four DUT-stratified folds described in Section 4.3. Arrows indicate the direction of improvement; the smaller line below each value reports the relative improvement over the trivial baseline (positive = better, irrespective of the metric's direction). The best-performing configuration is highlighted in bold.

Architecture	Features	MAE ↓	R^2 ↑	α-acc ↑	RA ↑
$20\,V_{DS}^{\mathrm{rel}}$ (baseline)	—	0.120 ± 0.013 —	0.738 ± 0.078 —	0.444 ± 0.037 —	0.716 ± 0.017 —
MLP	base	0.093 ± 0.020 (+23%)	0.816 ± 0.049 (+11%)	0.489 ± 0.039 (+10%)	0.695 ± 0.015 (-3%)
	cum (no penalty)	0.073 ± 0.047 (+40%)	0.846 ± 0.171 (+15%)	0.667 ± 0.211 (+50%)	0.836 ± 0.083 (+17%)
	cum + monotonicity	0.072 ± 0.048 (+40%)	0.850 ± 0.173 (+15%)	0.673 ± 0.219 (+51%)	0.839 ± 0.085 (+17%)
CMLP	base	0.097 ± 0.030 (+19%)	0.723 ± 0.193 (-2%)	0.413 ± 0.174 (-7%)	0.653 ± 0.071 (-9%)
	cum (no penalty)	0.068 ± 0.035 (+43%)	0.893 ± 0.072 (+21%)	0.629 ± 0.183 (+42%)	0.817 ± 0.065 (+14%)
	cum + monotonicity	0.063 ± 0.028 (+48%)	0.911 ± 0.049 (+23%)	0.662 ± 0.166 (+49%)	0.819 ± 0.061 (+14%)
NODE	base	0.077 ± 0.010 (+36%)	0.850 ± 0.026 (+15%)	0.493 ± 0.081 (+11%)	0.717 ± 0.018 (+0%)
	cum (no penalty)	0.050 ± 0.019 (+58%)	0.944 ± 0.028 (+28%)	0.768 ± 0.111 (+73%)	0.860 ± 0.043 (+20%)
	cum + monotonicity	0.049 ± 0.019 **(+59%)**	0.947 ± 0.029 **(+28%)**	0.792 ± 0.126 **(+78%)**	0.859 ± 0.045 **(+20%)**

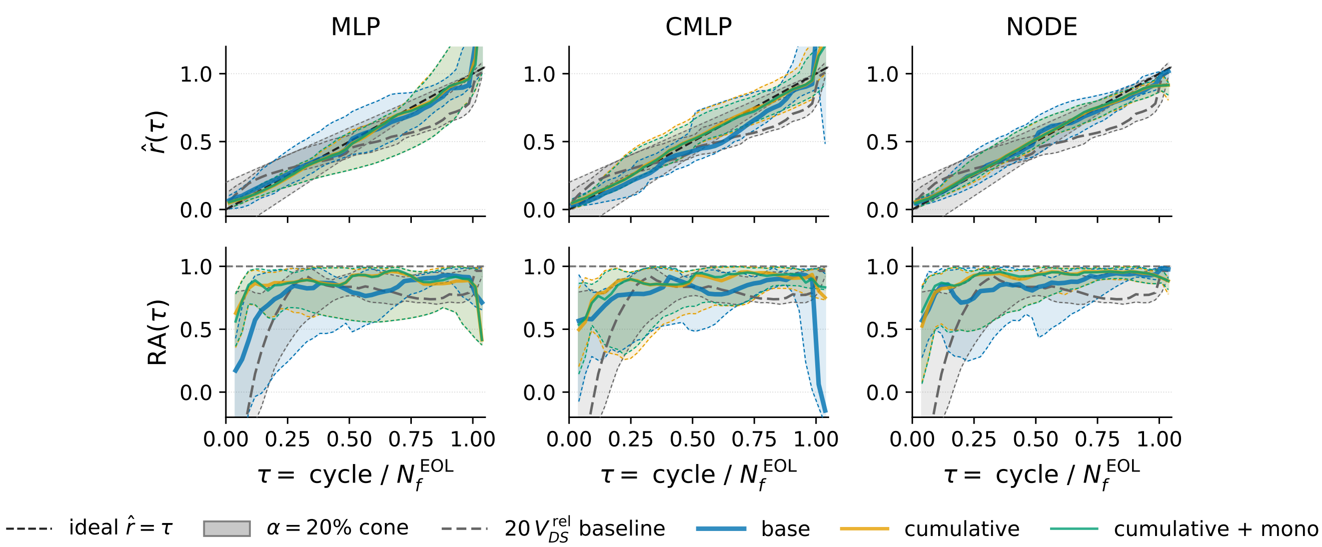

Figure 3. Aggregate validation behaviour along the normalised life $\tau = n/N_f^{\mathrm{EOL}}$. *Top:* predicted health index $\hat{r}(\tau)$ with the ideal $\hat{r} = \tau$ diagonal and the $\alpha = 20\%$ cone of Eq. (22). *Bottom:* relative accuracy RA(τ) of Eq. (23). Curves are the cross-DUT median at each τ bin; dashed envelopes mark the 10th and 90th percentiles over the validation DUTs pooled across the four folds of Section 4.3. The parameter-free baseline of Eq. (19) is drawn in gray as reference in every panel.

Adding the cumulative features of Eq. (17) brings coverage within two percentage points of nominal and shrinks the median predictive scale $\tilde{\sigma}$ by roughly 37%. The monotonicity penalty leaves the calibration essentially unchanged, in line with the small effect it already had on the point-estimate metrics of Table 3. The median ν_n sits around six in all three

Table 4. Predictive uncertainty calibration of the NODE on the validation cycles of the four-fold cross-validation. "Cov. $p\%$" is the empirical fraction of cycles inside the Student-t central interval at nominal coverage $p\%$, computed per-cycle using the predicted (σ_n, ν_n); the matching nominal values are 68.27% and 95.45%. Median ν and the fraction of cycles with $\nu_n < 10$ summarise the use of the heavy-tail regime; $\tilde{\sigma}$ is the median predictive scale in percent of the unit lifetime range.

NODE variant	68% cov. [%]	95% cov. [%]	$\tilde{\nu}$	ν<10 [%]	$\tilde{\sigma}$ [%]
base	66.1	88.3	5.0	100.0	7.28
cumulative	70.7	96.6	6.5	100.0	4.61
cumulative + mono	70.9	97.4	6.6	100.0	4.58

variants and $\nu_n < 10$ on every validation cycle: the heavy tails are always active and never collapse to a Gaussian limit.

Figure 4 resolves the same comparison locally along the life fraction τ. The deviation-size statistic adopted at each τ is the median absolute deviation (MAD), measured from the data as the cross-DUT median of the residual,

$$\mathrm{MAD}_n^{\mathrm{emp}} = \mathrm{med}\big|\, r_n - \mu_n \,\big|, \tag{24}$$

and predicted by the model as the analytic MAD of the Student-t distribution at cycle n,

$$\mathrm{MAD}_n^{\mathrm{pred}} = \sigma_n \cdot t^{-1}_{0.75,\,\nu_n}, \tag{25}$$

where $t^{-1}_{0.75,\nu}$ is the standardised Student-t median quantile that rescales the scale parameter σ_n into a deviation and brings the two quantities onto the same axis. For the base configuration the predicted MAD sits visibly below its empirical counterpart over most of the trajectory: the network is not only larger in absolute residual than the cumulative variants but also overconfident on that residual, declaring a tighter spread than the data actually exhibit. The cumulative variant brings the residual MAE from 7.7% to 5.0% of the unit lifetime range (Table 3), a 35% relative reduction, and at the same time aligns its predicted MAD with the empirical one, so the gain on accuracy and the gain on calibrated uncertainty come together. The cumulative with monotonicity variant tightens this picture a step further, with the predicted MAD slightly below the cumulative variant for most of τ without losing the calibration alignment, consistent with the marginal accuracy gain already observed in Table 3. The Student-t output is then a usable per-cycle uncertainty signal in its own right, and might be an input for the probabilistic prognostic metrics flagged as missing in Section 6.

6. DISCUSSION

The cross-validation results support an observation that cuts across the three architectures: appending the cumulative features of Eq. (17) provides a net positive contribution on every metric of Table 3, and the choice of architecture then modulates how much each network makes of this enrichment. The monotonicity penalty adds a smaller, architecture-dependent bump, primarily on MAE and α-accuracy. The broader reading is that, when degradation is dominated by a slow accumulation of stress, encoding that history at the feature level is a robust way to anchor the predictor to the physics; how much benefit each architecture extracts from this anchor is then a function of how well it combines the instantaneous history estimate with the temporal context it provides on its own.

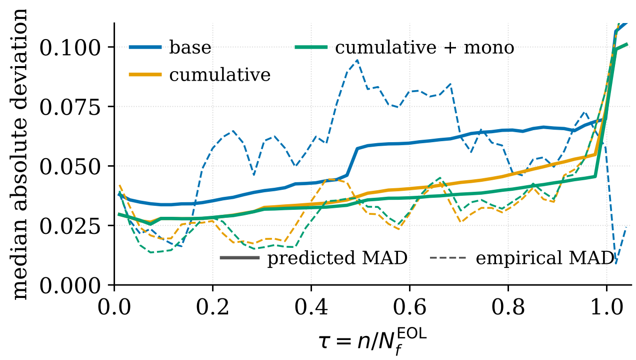

Figure 4. Predicted versus empirical median absolute deviation along the life fraction $\tau = n/N_f^{\mathrm{EOL}}$ for the three NODE variants. The predicted curves (solid) are $\sigma_n \cdot t^{-1}_{0.75,\,\nu_n}$, i.e. each σ_n rescaled by the standardised Student-t median quantile so that it can be compared directly to a deviation measure; the empirical curves (dashed) are the cross-DUT median of $|r - \mu|$ binned at the same resolution and pooled across the four folds. The cumulative variants track their empirical counterparts within one bin width over the central τ range; the base variant sits below its empirical counterpart, consistent with the under-coverage reported in Table 4.

A few aspects of the experimental setting frame the scope of these results. The protocol of Table 1 exercises the heating-pulse duration t_{heat} at the two levels prescribed by the AQG 324 recipe, on disjoint groups of DUTs and without mixing them within a single trajectory. The integration step of Eq. (18) is designed precisely to absorb such a mixing, and the consistent behaviour of the NODE across the two operating groups suggests that the NODE mechanism does carry useful information.

Verifying this directly would require dedicated campaigns with t_{heat} varying within a single DUT, in profiles closer to the realistic mission of a traction inverter.

The metrics of Section 4.4 are point-estimate metrics; they do not use the predictive distribution of the Student-t head. We rely on them because they are consolidated in the prognostics literature (Fink et al., 2020) and informative on different stages of the SiC health-state estimate: MAE and R^2 summarise the global residual, while α-accuracy and RA emphasise the late-life region where the tolerance shrinks and the denominator of the relative error is small. A separate calibration analysis is reported in Section 5.3. The general recommendation to report uncertainty alongside a point estimate is also raised in the same review; the (σ_n, ν_n) of Table 4 and Figure 4 is the form adopted here.

The predictive scale σ_n aggregates aleatoric and epistemic contributions; the present training setup does not separate them. With twenty-four DUTs the data budget for an explicit epistemic spread is limited, and the σ_n reported in Section 5.3 is dominated by the aleatoric component conditional on the input. A future development could refine the decomposition of these two contributions by training an ensemble of NODE variants from different seeds.

7. CONCLUSION

We have presented a condition-monitoring framework for SiC power modules that combines physics-informed cumulative features, a Student-t output head, and an optional monotonicity prior on top of three compact neural architectures of increasing temporal context. Evaluated under AQG 324-compliant power cycling on four operating-condition groups, the framework reduces the mean absolute error on the lifetime-fraction estimate by up to 60% relative to a parameter-free baseline derived from the qualification threshold, with the recurrent variant offering both the best central tendency and the lowest cross-fold variance. The dominant lever in the comparison is the feature engineering step: encoding the cumulative thermo-mechanical history of the module is what closes the gap between the raw per-cycle measurement and a usable health estimate, and we read the result of this study as a quantitative argument in favour of placing physical priors at the input level rather than absorbing them entirely into architectural complexity. A companion study (Scarpa et al., 2026) benchmarks the same cumulative NODE against five reference estimators from the prognostics literature: a parametric fit, a convolutional neural network (CNN), a long short-term memory network (LSTM), a snapshot MLP and a transformer forecaster. The five are tested on two campaigns with different failure mechanisms, the wire lift-off studied here and solder degradation. The cumulative NODE is the only method whose accuracy transfers across both mechanisms, while the reference estimators degrade for sintered modules compared to soldered ones.

This confirms the weight of the input representation, and of the junction-temperature history in particular, in making a health estimate portable across failure modes. Broader operating envelopes that mix mission profiles within a single trajectory, probabilistic metrics that complement the point-estimate ones, and an online adaptation phase that refreshes the network during normal converter operation are the natural next steps to consolidate the framework toward in-service deployment.

ACKNOWLEDGMENT

In Germany this work of Infineon Technologies AG and Infineon Technologies Dresden AG & Co. KG is funded in the frame of the Important Project of Common European Interest on Microelectronics and Communication Technologies (IPCEI ME/CT). The IPCEI ME/CT is funded by the German Federal Ministry for Economic Affairs and Energy, the Bavarian Ministry for Economic Affairs, Regional Development and Energy, the Ministry of Economic Affairs, Industry, Climate Action and Energy of the State of North Rhine-Westphalia, the Saxon State Ministry for Economic Affairs, Labour, Energy and Climate Action and the European Union within "NextGenerationEU".

REFERENCES

Achatz, S., Hoppe, M., Knobloch, K., Kusmenko, E., Olschewski, J., Tüllmann, M., & Pfost, M. (2026, March). End-to-end snapshot models for RUL prediction of silicon carbide power modules. In *Proceedings of the international conference on integrated power electronics systems (CIPS).* Dresden, Germany. (Available via VDE-Verlag; forthcoming on IEEE Xplore)

Barbagallo, C., Rizzo, S. A., Scelba, G., Scarcella, G., & Cacciato, M. (2021). On the lifetime estimation of SiC power MOSFETs for motor drive applications. *Electronics*, *10*(3), 324. doi: 10.3390/electronics10030324

Ceccarelli, L., Kotecha, R. M., Bahman, A. S., Iannuzzo, F., & Mantooth, H. A. (2019). Mission-profile-based lifetime prediction for a SiC MOSFET power module using a multi-step condition-mapping simulation strategy. *IEEE Transactions on Power Electronics*, *34*(10), 9698–9708. doi: 10.1109/TPEL.2019.2893636

Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2018). Neural ordinary differential equations. In *Advances in neural information processing systems (NeurIPS).*

Clech, J.-P. (1997). Solder reliability solutions: A PC-based design-for-reliability tool. *Soldering & Surface Mount Technology*, *9*(2), 45–54. doi: 10.1108/09540919710800719

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 *Proceedings of the ieee international conference on prognostics and health management (icphm)* (pp. 312–319). doi: 10.1109/ICPHM57936.2023.10194131

Di Nuzzo, G., Tuellmann, M., Methfessel, T., & Rzepka, S. (2022). Condition monitoring indicators for Si and SiC power modules. *Microelectronics Reliability*, *138*, 114614. doi: 10.1016/j.microrel.2022.114614

ECPE European Center for Power Electronics e.V. (2025).

Qualification of power modules for use in power electronics converter units in motor vehicles (Tech. Rep. Nos. AQG 324, Release 04.1/2025). ECPE. (Available: `https://www.ecpe.org`)

Fassi, Y., Heiries, V., Boutet, J., & Boisseau, S. (2024). Toward physics-informed machine-learning-based predictive maintenance for power converters—a review. *IEEE Transactions on Power Electronics*, *39*(2), 2692–2723. doi: 10.1109/TPEL.2023.3328438

Fink, O., Wang, Q., Svensén, M., Dersin, P., Lee, W.-J., & Ducoffe, M. (2020). Potential, challenges and future directions for deep learning in prognostics and health management applications. *Engineering Applications of Artificial Intelligence*, *92*, 103678. doi: 10.1016/j.engappai.2020.103678

Frusque, G., Nejjar, I., Nabavi, M., & Fink, O. (2024). Semisupervised health index monitoring with feature generation and fusion. *IEEE Transactions on Reliability*, 1–15. doi: 10.1109/TR.2024.3496076

Hanif, A., Yu, Y., DeVoto, D., & Khan, F. (2019). A comprehensive review toward the state-of-the-art in failure and lifetime predictions of power electronic devices. *IEEE Transactions on Power Electronics*, *34*(5), 4729–4746. doi: 10.1109/TPEL.2018.2860587

Kovacevic, I. F., Drofenik, U., & Kolar, J. W. (2010). New physical model for lifetime estimation of power modules. In *Proceedings of the international power electronics conference (ipec)* (pp. 2106–2114). Sapporo, Japan. doi: 10.1109/IPEC.2010.5542019

Kovacevic-Badstuebner, I., Kolar, J. W., & Schilling, U. (2015). Modelling for the lifetime prediction of power semiconductor modules. In H. S.-h. Chung, H. Wang, F. Blaabjerg, & M. Pecht (Eds.), *Reliability of power electronic converter systems* (pp. 103–140). London, UK: Institution of Engineering and Technology (IET). doi: 10.1049/PBPO080E_ch5

Kumar, S., Allard, B., Hologne-Carpentier, M., Clerc, G., & Auger, F. (2026). Review of prognosis approaches applied to power SiC MOSFETs for health state and remaining useful life prediction. *Entropy*, *28*(2), 234. doi: 10.3390/e28020234

Ni, Z., Lyu, X., Yadav, O. P., Singh, B. N., Zheng, S., & Cao, D. (2020). Overview of real-time lifetime prediction and extension for SiC power converters. *IEEE Transactions on Power Electronics*, *35*(8), 7765–7794. doi: 10.1109/TPEL.2019.2962503

Olschewski, J., Hoppe, M., Kusmenko, E., Knobloch, K., Tüllmann, M., Martin, A., & Pfost, M. (2025). Machine learning-based condition monitoring from power cycling data for silicon carbide power inverter modules. In *Proceedings of the ieee international reliability physics symposium (irps).* doi: 10.1109/IRPS48204.2025.10983172

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. doi: 10.1038/323533a0

Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and Health Management*, *1*(1). doi: 10.36001/ijphm.2010.v1i1.1336

Scarpa, M., Kusmenko, E., Toso, F., Bruschetta, M., Carli, R., & Achatz, S. (2026). Failure-mechanism transferability of cumulative-damage features for health state estimation of SiC power modules. In *Proceedings of the 52nd annual conference of the ieee industrial electronics society (iecon).* (Accepted. arXiv:2608.08365)

Scarpa, M., Pase, F., Carli, R., Bruschetta, M., & Toso, F. (2025). Data-driven power loss identification through physics-based thermal model backpropagation. In *Proceedings of the 23rd european control conference (ecc)* (pp. 2758–2763). Thessaloniki, Greece. doi: 10.23919/ECC65951.2025.11186906

Villalobos, A., Barrutia, I., Peña-Alzola, R., Dragicevic, T., & Aizpurua, J. I. (2025). Comparative analysis and evaluation of ageing forecasting methods for semiconductor devices in online health monitoring. *Engineering Applications of Artificial Intelligence*, *150*, 110545. doi: 10.1016/j.engappai.2025.110545

Yang, S., Bryant, A., Mawby, P., Xiang, D., Ran, L., & Tavner, P. (2011). An industry-based survey of reliability in power electronic converters. *IEEE Transactions on Industry Applications*, *47*(3), 1441–1451. doi: 10.1109/TIA.2011.2124436

Zhao, M., & Fink, O. (2025). *Disentangling slow and fast temporal dynamics in degradation inference with hierarchical differential models.* (Preprint, revised March 2026. Available: `https://arxiv.org/abs/2509.00639`)

Zio, E. (2022). Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. *Reliability Engineering & System Safety*, *218*, 108119. doi: 10.1016/j.ress.2021.108119