



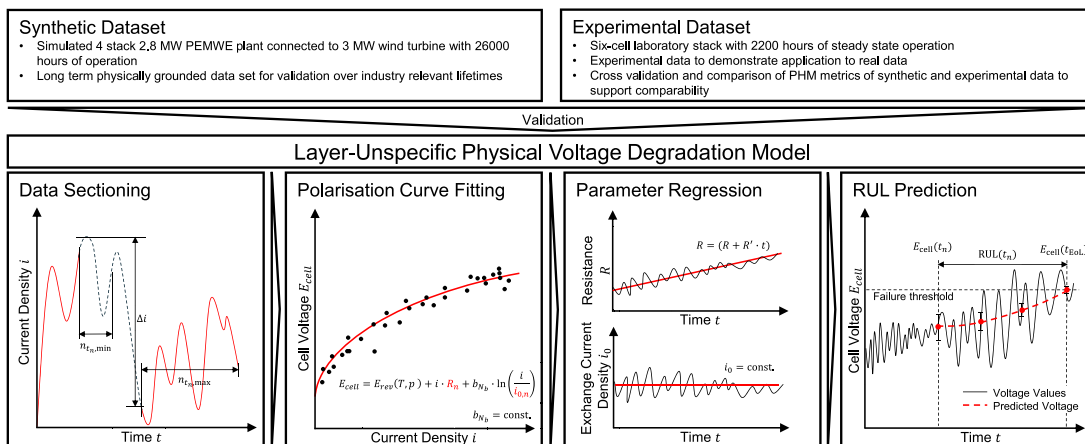
# Development and validation of a layer-unspecific physical model for voltage degradation to predict the remaining useful life of proton exchange membrane water electrolyzers

Felix Dittmar<sup>a</sup> ,\* , Thomas Lickert<sup>b</sup> , Tom Smolinka<sup>b</sup> , Karsten Pinkwart<sup>a</sup> , Jens Tübke<sup>a</sup>

<sup>a</sup> Fraunhofer Institute for Chemical Technology (ICT), Joseph-von-Fraunhofer-Straße 7, Pfinztal (Berghausen), 76327, Baden-Württemberg, Germany

<sup>b</sup> Fraunhofer Institute for Solar Energy Systems (ISE), Heidenhofstraße 2, Freiburg, 79110, Baden-Württemberg, Germany

## GRAPHICAL ABSTRACT



## HIGHLIGHTS

- Introduces in-operation polarization-curve extraction without downtime.
- Applies a layer-unspecific physical model to track voltage degradation.
- Validates with synthetic (26 200 h) and experimental (2200 h) data.
- Demonstrates synthetic data use for validation over industry-relevant lifetimes.

## ARTICLE INFO

Dataset link: <https://doi.org/10.5281/zenodo.18268174>

### Keywords:

PEM water electrolysis (PEMWE)  
Voltage decay  
Degradation model

## ABSTRACT

The deployment of proton exchange membrane (PEM) water electrolysis plants is accelerating, creating a need for accurate lifetime prediction to support reliable and cost-effective operation. Yet, degradation models suitable for plant-level lifetime estimation remain scarce. To address this gap, this study introduces a layer-unspecific physical model to estimate voltage degradation and the remaining useful life of PEM water electrolysis stacks from current–voltage data. The model extracts quasi-steady polarization curves from operation without diagnostic downtime. As a central novelty, it fits a simplified electrochemical equation

\* Corresponding author.

E-mail address: [felix.dittmar@ict.fraunhofer.de](mailto:felix.dittmar@ict.fraunhofer.de) (F. Dittmar).

<https://doi.org/10.1016/j.jpowsour.2025.239122>

Received 2 October 2025; Received in revised form 28 November 2025; Accepted 13 December 2025

Available online 17 January 2026

0378-7753/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

assuming a fixed Tafel slope, which enables stable tracking of the temporal evolution of resistance and exchange current density by avoiding parameter collinearity. Uncertainty is quantified through Monte Carlo propagation of regression parameters, enabling lifetime forecasts with confidence intervals. Validation is conducted with a 26 200 h synthetic dataset and a 2200 h experimental six-cell stack dataset. The synthetic data enable validation over industry-relevant lifetimes without multi-year testing, while the experimental data confirm practical applicability. Across both datasets, the model reproduced voltages with millivolt-scale error and provided prognostic horizons suitable for maintenance planning. Overall, the approach offers a practical and interpretable model for PEM water electrolysis efficiency prediction and accelerates model development while reducing reliance on extended lifetime experiments.

## 1. Introduction

As of September 2025, the operational global proton exchange membrane water electrolysis (PEMWE) capacity is approximately 605 MW<sub>el</sub> [1]. Based on International Energy Agency projections of high-likelihood projects, the global installed PEMWE capacity could reach about 30 GW<sub>el</sub> by 2030 [2]. This rapid expansion makes long system lifetimes, high reliability, and cost-effectiveness increasingly relevant for PEMWE systems [3,4]. Prognostics and health management (PHM) is an engineering discipline aimed at ensuring reliability and cost-efficiency of complex systems by combining monitoring, diagnostics, and prognostics [5,6]. Applied to PEMWE, PHM offers degradation-informed operation management and predictive maintenance [7]. At its core, PHM depends on accurate degradation models of lifetime-limiting parameters, since these models form the foundation for predicting the future state of health and enabling reliable lifetime forecasting [8].

Dittmar et al. (2025) show in a recent review that research on applying PHM to PEMWE is still scarce and must satisfy several key requirements to be viable for industrial application [7]: (i) Appropriate health indicators must be selected to represent the system's state of health. (ii) Models must operate under the limited data conditions of PEMWE plants, where only a few sensors are installed and additional characterization measurements reduce system availability [9,10]. Furthermore, given the industrial lifetime of PEMWE systems of 10 000 h to 50 000 h, the degradation models must be validated over these periods [3]. (iii) The selection of a model method must not only be able to accurately describe the evolution of the health indicator over time but also match the constraints of PEMWE plant data availability, the characteristics of the chosen health indicator, and the intended prognostic task [7]. (iv) Any prognostic method must provide an accurate estimate of the remaining useful life (RUL) and quantify the confidence associated with the prediction [11], especially as predictions are updated continuously with incoming operational data [8].

However, currently available PEMWE degradation models do not fully meet the requirements for PHM [7]. Among these, the choice of degradation model (iii) is particularly critical, as it integrates aspects of health indicators, data availability, and prediction tasks. The following paragraphs outline the limitations of the current literature in fulfilling these requirements.

For the selection of health indicators (i), a range of candidates have been proposed for PEMWE, such as gas crossover or parameters based on cell components [7]. This study focuses on voltage, as it is widely reported in both experimental degradation and degradation modeling literature and directly reflects efficiency loss [4]. Moreover, both current and voltage measurements are available from online sensors on the plant-level, providing a robust basis to model the voltage-current relationship [9,10].

Regarding the functionality with limited experimental data availability (ii), most reported degradation experiments last less than 1000 h, providing only a limited foundation for validating PEMWE stacks over industry-relevant lifetimes [12]. One potential solution that avoids multi-year experiments, widely used in machine learning and increasingly applied to time-series data, is the generation of synthetic datasets [13]. Yan et al. (2025) applied such an approach to PEMWE degradation modeling by using a synthetically generated dataset with a known ground truth to assess model accuracy [14]. However, their

data generation relied on a purely empirical model that does not capture PEMWE behavior. While suitable for evaluating the accuracy of their specific empirical degradation model, the broader validity of this approach is limited. Extending this concept to physically grounded synthetic data generation could enable more realistic long-term simulations and facilitate cross-study comparisons of PHM approaches over industry-relevant lifetimes.

For the selection of a model (iii), a prognostic approach for PHM consists of two components: a modeling method, which describes the evolution of the HI over time, and a prediction method, which extrapolates this evolution into the future [7]. Modeling methods can be model-based (empirical, semi-empirical, or physical) or data-driven [15]. Prediction methods range from statistical approaches to filtering or machine-learning techniques. In data-driven models, prediction is inherently part of the modeling method, whereas in model-based approaches it is explicitly paired with the chosen model [7]. Given the constraints of plant-level data availability, layer-unspecific voltage degradation models are a promising choice because they can be validated directly with available measurements and do not require resolving the underlying degradation phenomena [7]. Such models capture the integral change of a system variable — such as voltage — across the entire system [16]. Only a few layer-unspecific voltage degradation models have been validated with experimental plant-representative data meeting requirement (ii) [14,17–20]. Existing data-driven models show specific limitations: lack of reported accuracy measures [17], focus on static operation profiles [18], or inability to represent the logarithmic behavior in the low-current density region [19]. While Xu et al. (2024) and Woelke et al. (2025) achieve high predictive accuracy over approximately 1000 h in the high-current density region [18,19], accurate prediction of the low-current regime is essential for mission planning of PEMWE plants operating under partial load [7]. Yan et al. (2024) present an empirical layer-unspecific model combined with a statistical method (transfer linear regression) that predicts voltage degradation over long time periods, but fails to capture the logarithmic low-current behavior [14]. Layer-unspecific physical models incorporating a logarithmic term have, in the related field of PHM for PEM fuel cells, demonstrated the ability to reproduce voltage degradation in this regime [21]. Pape et al. (2025) adapt such an approach, using an electrochemical equation with resistance and Tafel terms to fit polarization curves, and tracking parameter changes via time-interval-based linear regression [20]. While promising, the method is prone to parameter collinearity and risks inaccurate long-term extrapolation. These considerations motivate further exploration of layer-unspecific physical models for voltage degradation modeling.

For requirement (iv), only Yan et al. (2024) among the previously discussed PEMWE degradation modeling studies evaluate the uncertainty of their predictions [14]. None report estimates of the RUL or assess how prediction accuracy evolves when trained on progressively larger datasets. Therefore, leaving this requirement for PHM in PEMWE largely underexplored.

Based on these gaps, two aspects emerge as priorities for advancing PHM of PEMWE. First, enabling long-term model validation under laboratory data constraints, and second developing layer-unspecific physical models that include uncertainty quantification, RUL estimation, and performance assessment as datasets grow. These priorities

motivate the central research question of this study: *How can a layer-unspecific physical model for PEMWE voltage degradation be developed and validated for industry-relevant lifetimes under laboratory data constraints?*

We address this research question by developing a layer-unspecific physical model for predicting PEMWE voltage degradation, designed for minimal data requirements and validated on both synthetic and experimental datasets. The approach extracts polarization curves from operational data and tracks degradation parameters over time, enabling long-term prediction with quantified uncertainty. RUL estimates are derived from these predictions and assessed using established PHM performance metrics.

Our study offers several contributions to the literature. First, we demonstrate how predictive models for voltage degradation can be validated over industry-relevant lifetimes without waiting for multi-year experiments, thereby providing researchers with a method to accelerate the development of PEMWE degradation models. Second, we present a novel physics-based polarization-curve fitting approach that yields a unique solution for extracting parameters associated with ohmic and kinetic losses, enabling researchers to track degradation trends of underlying mechanisms using only current and voltage measurements while avoiding parameter overfitting. Finally, by coupling these parameters with uncertainty-based RUL estimation and established PHM metrics, we provide a framework for assessing and comparing prognostic models, illustrating how PHM approaches for PEMWE can be systematically evaluated. In addition to these contributions to the literature, our study also provides practical contributions. We introduce an in-operation polarization-curve extraction method that removes the need for dedicated diagnostic downtime, making voltage degradation tracking feasible in commercial plants. Moreover, we develop an approach that requires only the minimal data typically available in PEMWE plants, thereby enhancing the feasibility of PHM for industrial application.

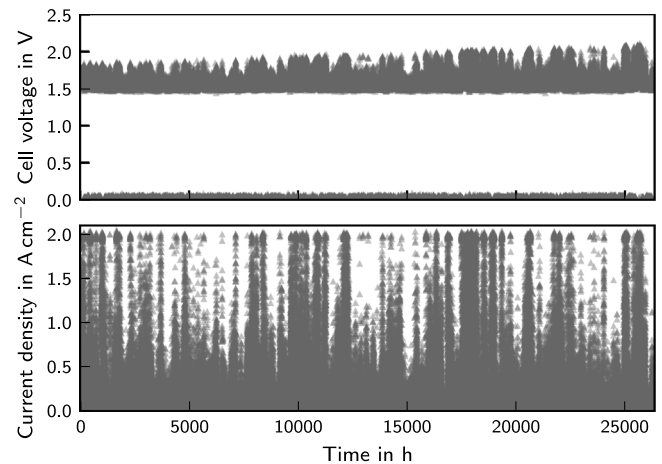
The work is structured as follows. Section 2 presents the utilized datasets and method of the degradation model. In Section 3 we present the results, discuss the findings, and compare PHM performance metrics of the synthetic and experimental datasets.

## 2. Method

### 2.1. Generation of synthetic data

For the evaluation of the voltage degradation model and its computational accuracy over long time periods, a synthetic dataset was utilized providing a defined ground truth. This synthetic dataset simulates a PEMWE plant coupled to a wind turbine. The synthetic data was based on wind velocity records provided by the German weather authorities over a three-year period with a ten-minute resolution corresponding to 26 256 operating hours [22]. Wind velocity values were converted into a power profile assuming a 3 MW Vestas V112-3.075 onshore wind turbine [23].

The assumed PEMWE plant consists of four current state of the art stacks as reported by Hemauer et al. (2023), each with an electrical capacity of  $P_{\text{stack,el}} = 700 \text{ kW}$ , resulting in a total plant capacity of  $P_{\text{plant,el}} = 2.8 \text{ MW}$  [24]. Such configurations are representative of wind-to-hydrogen systems, where electrolyzers are directly coupled to onshore or offshore wind power installations [25,26]. The voltage and current of each stack were calculated using a simplified electrochemical equation for the cell voltage. Constant operating conditions were assumed, with a temperature of  $\vartheta = 55^\circ \text{C}$ , atmospheric pressure on the cathode and a pressure of  $p_a = 20 \text{ bar}$  on the anode side [24]. A linear cell voltage decay rate of  $\delta E_{\text{cell}} = 10 \mu\text{V/h}$  was assumed, consistent with decay rates reported in the literature [12,27]. The assumption of a linear decay transfers to the assumption of an increase in ohmic resistance being the dominant degradation mechanisms as reported in experimental literature for a variety of current density profiles [28,29]. Nevertheless, this simplification does not account for



**Fig. 1.** Synthetic current density and cell voltage profile generated for a four-stack PEMWE plant coupled to a 3 MW wind turbine, assuming a linear cell voltage decay rate of  $\delta E_{\text{cell}} = 10 \mu\text{V/h}$  and a relative inaccuracy of  $f_{\text{rel,max}} = 1\%$ .

degradation mechanisms like catalyst degradation on the anode, which depending on the operation mode displays varying non-linear dynamics [28]. Additionally, normally distributed noise with a maximum relative magnitude  $f_{\text{rel,max}}$  was applied to the calculated current density and voltage values to simulate fluctuations in operating conditions and measurement noise. The relative magnitude was varied between  $f_{\text{rel,max}} = 0\%$ ,  $1\%$ , and  $3\%$  to evaluate the computational accuracy of the model in dependence on the fluctuations in operation data. This type of noise cannot reproduce systematic temperature or pressure dynamics that occur during plant load changes, as these depend on controller response and balance-of-plant characteristics. Nevertheless, modern PEMWE systems equipped with optimized control schemes exhibit only small variations in operating conditions even under dynamic profiles, supporting the reasonableness of this assumption [30]. Fig. 1 shows the generated current density and cell voltage profile with a relative inaccuracy of  $f_{\text{rel,max}} = 1\%$ . While these simplifications limit the generalization of the synthetic dataset as a representation of real PEMWE operation, the dataset provides a controlled ground truth for assessing long-term numerical stability and parameter tracking, which cannot be obtained from experimental data alone. Further details on the assumptions and parameters for synthetic data generation are provided in the Supplementary Information.

### 2.2. Experimental setup

The experimental dataset comprises 2200 h of operation of a six-cell PEMWE rainbow stack conducted in 2016, initially intended to evaluate various coated porous transport layers (PTLs) and bipolar plates (BPPs). Each catalyst-coated membrane (CCM) had an active area of  $150 \text{ cm}^2$  and utilized a Nafion 117 membrane. All CCMs were obtained from a commercial manufacturer with catalyst loadings of approximately  $2 \text{ mg}_{\text{Ir}} \text{ cm}^{-2}$  on the anode and  $1 \text{ mg}_{\text{Pt}} \text{ cm}^{-2}$  on the cathode. For all cells, the cathode PTL was a carbon sheet (Spectrecarb 2050A-6060) with a microporous transport layer (SGL GDL 35 DC). The anode PTL was an experimental titanium-mesh with a gradient in porosity, and in cells two, four, and six, it was coated along with the anode and cathode sides of the BPPs. The applied coatings were oxide based metal coatings. The BPP was a sheet made from titanium with no flow field geometry. Further details about the stack configuration are provided in Figure S2 of the Supplementary Information.

Testing was conducted using an in-house automated test station equipped with online measurements of temperature, pressure, gas crossover, water conductivity, stack voltage, and individual cell voltages. A piping and instrumentation diagram is included in Figure S3 of the Supplementary Information.

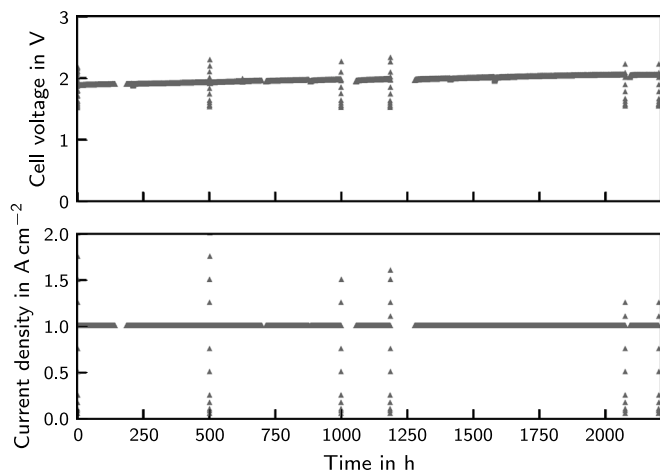


Fig. 2. Mean cell voltage and current density over time of a with constant current operated six-cell laboratory stack with a geometric area of  $150\text{cm}^2$ .

Following initial conditioning and begin-of-test (BoT) characterization, the stack was operated under constant conditions at  $60^\circ\text{C}$ , atmospheric pressure, and a constant current density of  $1\text{A cm}^{-2}$  in four repeating intervals of approximately 500 h. At the beginning of each interval, polarization curves were recorded between  $0.05\text{A cm}^{-2}$  and  $2\text{A cm}^{-2}$  with a hold duration of 3 min per measurement point. Water was circulated at flow rates of 3 kg/min at the anode and 2.4 kg/min at the cathode. The flow rate of the cathode was adjusted after approximately 1090 h to 2 kg/min due to an increase in pressure loss on the cathode. The conductivity of the circulating water remained below  $0.1\ \mu\text{S cm}^{-1}$  throughout constant operation. The mean cell voltage and current density are presented in Fig. 2. Operation interruptions occurred periodically due to technical issues; detailed descriptions of these events can be found in the Supplementary Information (Table S2). Furthermore, impedance measurements at a fixed testing frequency of 1 kHz (with an amplitude of 10 mA and without DC-bias) have been carried out at BoT and EoT with a 4338B Milliohmimeter (Agilent).

## 2.3. Voltage degradation model

### 2.3.1. Data acquisition and processing

The proposed modeling method requires time-series measurements of the mean cell voltage  $E_{\text{cell}}$  and current density  $i$ , along with fixed values for anode pressure  $p_a$ , cathode pressure  $p_c$ , and stack temperature  $T$ . When only the voltage of an entire electrolyzer module (comprising multiple stacks) or of a single stack is available, the cell voltage can be estimated assuming a homogeneous voltage distribution across all stacks and cells. The current density is calculated from the applied current and the active cell area.

A uniform distribution of voltage and temperature across the stack is assumed, consistent with the layer-unspecific nature of the model. If individual cell voltages are available, the method can also be applied at the cell level to analyze degradation. To ensure quasi-steady-state conditions, all data points with  $p_a$ ,  $p_c$ , or  $T$  outside defined bounds are excluded [7]. Excessive fluctuations reduce the number of usable data points and thereby increase predictive uncertainty. Moreover, because temperature and pressure directly affect the cell potential, wider bounds increase the uncertainty of the quasi-steady-state assumption. The feasibility of narrow bounds depends on the balance-of-plant capabilities: even under highly dynamic wind-driven operation, temperature stability within 5 K to 10 K has been achieved with standard PID control, and below 3 K with optimized feedforward controllers [30]. The selected cut-off bounds must therefore reflect the achievable operating stability of the monitored system while preserving a sufficient number of usable data points.

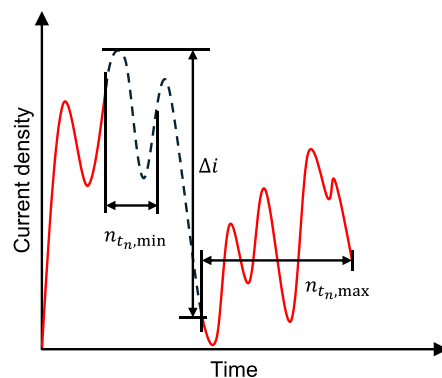


Fig. 3. Visualization of data sectioning into quasi-steady-state time intervals for extracting polarization curves. The current density profile over time is plotted as a continuous line, although in reality it is composed of individual data points.

### 2.3.2. Data sectioning to obtain polarization curves

The underlying method extracts time intervals from time-series data to obtain polarization curves. This so-called “data sectioning” is based on a methodology first proposed by Heyne (2015) for the lifetime estimation of PEM fuel cells [21]. As illustrated in Fig. 3, the segmentation criteria include a current density interval,  $\Delta i$ , a minimum number of data points,  $n_{t_n, \text{min}}$ , and a maximum number of data points,  $n_{t_n, \text{max}}$ . The current density interval  $\Delta i$  defines the range over which voltage and current density data points are selected. The criterion  $n_{t_n, \text{min}}$  ensures a sufficient number of data points for reliable polarization curve fitting. Conversely,  $n_{t_n, \text{max}}$  restricts the segment duration to maintain a quasi-steady-state for every extracted polarization curve. The selection of these criteria is dependent on data characteristics like the measurement frequency and current operating range of the monitored system.

In addition to operating parameter stability, other interfering effects — such as reversible voltage losses and long-term voltage degradation — must also be considered when assuming quasi-steady-state conditions [7]. In this study, no correction was applied to account for in-situ reversible effects. The only available modeling approach in the literature, proposed by Krenz et al. (2024) [31], attributes these effects to improved anode reaction kinetics due to oxidation and reduction processes of the Ir catalyst. However, applying this method over long periods requires a model that explicitly accounts for electrochemically active surface area (ECSA) degradation of the Ir anode catalyst. In practice, reversible effects, typically reported in the range of 10 mV to 60 mV per cell [7], can lead to apparent shifts in segments of the extracted polarization curve. These effects generally relax at lower potentials [31], so the dynamic operation required to extract polarization curves naturally promotes their partial relaxation. Thus, neglecting reversible effects introduces only a limited error and is acceptable for the present modeling approach, although their influence may be more pronounced under steady-state operation. Consequently, the choice of  $n_{t_n, \text{max}}$  only accounts for both the voltage decay rate and the time resolution of the time-series data to balance the trade-off between capturing sufficient data and maintaining quasi-steady-state conditions.

### 2.3.3. Polarization curve fitting

After identifying time intervals that correspond to polarization curves, an electrochemical model is fitted to each interval using least squares. The cell voltage  $E_{\text{cell}}$  can be calculated according to Eq. (1) [32–34].

$$E_{\text{cell}} = E_{\text{rev}} + \eta_{\text{HER}} + \eta_{\text{OER}} + i \cdot (R_{\text{mem}} + R_{\text{el}} + R_{\text{H}^+, \text{a}}^{\text{eff}} + R_{\text{H}^+, \text{c}}^{\text{eff}}) + \eta_{\text{mt}} \quad (1)$$

$E_{\text{rev}}$  is the reversible voltage (open circuit voltage or Nernst potential), a purely thermodynamic property depending on the anode pressure  $p_a$ , cathode pressure  $p_c$ , and cell temperature  $T$ . The kinetic overpotentials of the hydrogen and oxygen evolution reactions (HER, OER) are denoted by  $\eta_{\text{HER}}$  and  $\eta_{\text{OER}}$ , respectively. The ohmic resistance of the membrane is represented by  $R_{\text{mem}}$ , while  $R_{\text{el}}$  accounts for electronic resistances, such as contact resistances between cell components. The effective proton transport resistances in the anode and cathode are given by  $R_{\text{H}^+,a}^{\text{eff}}$  and  $R_{\text{H}^+,c}^{\text{eff}}$ . Mass-transport losses are described by  $\eta_{\text{mt}}$  being particularly relevant at elevated pressures and current densities [34,35].

The reversible voltage can be determined using Eq. (2), assuming ideal gases, an incompressible fluid, neglecting gas crossover, and assuming thermodynamic equilibrium for gas saturation. These assumptions are reasonable for the typical pressure range  $p_c = 1 - 30$  bar and temperature range  $\vartheta = 60 - 80$  °C of a PEMWE plant [9].

$$E_{\text{rev}}(T, p) = E_{\text{rev}}^0(T, p^*) + \frac{\tilde{R}T}{2F} \ln \left[ \left( \frac{p_c - p_{\text{H}_2\text{O}}}{p^*} \right) \left( \frac{p_a - p_{\text{H}_2\text{O}}}{p^*} \right)^{0.5} \right] \quad (2)$$

The standard reversible potential  $E_{\text{rev}}^0(T, p^*)$  at a reference pressure of  $p^* = 1$  atm and temperature of  $\vartheta = 298.15$  K can be obtained from Bratsch et al. (1989) [36] as:

$$E_{\text{rev}}^0(T, p^*) = 1.2291 \text{ V} + (T - 298.15 \text{ K}) \cdot 8.456 \times 10^{-4} \text{ V K}^{-1} \quad (3)$$

The partial pressure of water  $p_{\text{H}_2\text{O}}$ , which is assumed to be equal to the saturated water pressure  $p_{\text{s,H}_2\text{O}}$ , is calculated according to the Antoine equation with coefficients taken from Sattler et al. (2001) [37]:

$$\log \left( \frac{p_{\text{s,H}_2\text{O}}}{\text{mbar}} \right) = 8.19625 - \frac{1730.63}{233.426 + \frac{\vartheta}{\text{[}^\circ\text{C]}}} \quad (4)$$

While Eq. (1) provides a reasonable approximation of the contributions of different mechanisms to the cell voltage, it includes too many variables for a stable fit to a polarization curve, leading to overfitting and infinitely many equivalent solutions. A more practical approach is to simplify Eq. (1) by retaining only the major voltage contributions and consolidating mathematically equivalent terms, such as slopes, into a single variable to reduce the risk of overfitting. This can be achieved by neglecting the kinetic overpotential of the cathode,  $\eta_{\text{HER}}$ , as it contributes only minor voltage losses per cell (approximately 1 mV at 3 A/cm<sup>2</sup>) [33]. The voltage losses due to mass transport  $\eta_{\text{mt}}$  will also be neglected, as modeling them in detail will result in overfitting. This assumption is reasonable because current PEMWE stacks typically operate below 2.5 A cm<sup>-2</sup> [9], and experimental studies show that even up to 10 A cm<sup>-2</sup> no significant mass-transport-related voltage rise is observed [38,39]. Furthermore, the linear slope contributions,  $R_{\text{mem}}$ ,  $R_{\text{el}}$ ,  $R_{\text{H}^+,a}^{\text{eff}}$ , and  $R_{\text{H}^+,c}^{\text{eff}}$ , are combined into a single resistance term,  $R$ . The kinetic overpotential of the anode  $\eta_{\text{OER}}$  can be calculated using the Tafel equation:

$$\eta_{\text{OER}} = b \cdot \log \left( \frac{i}{i_0} \right) \quad (5)$$

$b$  is the Tafel slope and  $i_0$  is the anodic exchange current density. The Tafel slope is given by [40]:

$$b = 2.3026 \cdot \frac{\tilde{R}T}{2\alpha F} \quad (6)$$

Both parameters are temperature-dependent and can be described using an Arrhenius-type relationship for the anodic exchange current density and the charge transfer coefficient  $\alpha$ , respectively [41]. However, since the polarization curve is extracted over a finite time period, only an average temperature can be considered. To mitigate overfitting and reduce computational complexity, the temperature dependence of  $b$  and  $i_0$  is therefore neglected. Instead, temperature variations are controlled by excluding data points with significant fluctuations, as

PEMWE stacks typically maintain stable operating temperatures as described in Section 2.3.1. Under these assumptions, Eq. (1) simplifies to:

$$E_{\text{cell}} = E_{\text{rev}}(T, p) + i \cdot R + b \cdot \log \left( \frac{i}{i_0} \right) \quad (7)$$

Eq. (1) is an often utilized approach in the literature on modeling PEMWE cell voltage [42]. To enhance the stability of the optimization, the first  $N_b$  polarization curves after the beginning of the test (BoT) are used to determine an average Tafel slope  $\bar{b}_{N_b}$ , under the assumption that it remains constant over time. Generally, a change in the Tafel slope would indicate a change in the rate-determining step or a shift in the OER mechanism, which is unlikely for the same catalyst type [43]. Experimental literature supports the assumption of a stable Tafel slope across a range of Ir-based catalysts and loadings [44,45]. Instead, the dominant degradation effects are typically attributed to changes in the exchange current density  $i_0$  and the high-frequency resistance (HFR), which approximates the overall ohmic cell resistance [28,44,45]. The parameter  $N_b$  is chosen such that the evaluation period remains short enough to limit the influence of degradation on the extracted values of  $b$ . Rogler et al. (2023) [43] compared different approaches for fitting Tafel equations to describe OER degradation at low catalyst loadings and arrived at a similar conclusion: fixing the Tafel slope reduces the risk of error propagation when estimating the apparent exchange current density, which can otherwise vary by orders of magnitude. If the Tafel slope were to change over time, its effect would manifest as systematic shifts in the fitted values of  $R$  and  $i_0$ . This would reduce the mechanistic interpretability of these apparent parameters, but would not necessarily impair the accuracy of the predicted cell voltage, as long as the regression (see Section 2.3.4) adequately captures the resulting parameter trends.

$$E_{\text{cell},n} = E_{\text{rev}}(T_n, p_n) + i \cdot R_n + \bar{b}_{N_b} \cdot \log \left( \frac{i}{i_{0,n}} \right) \quad (8)$$

Furthermore, with this constraint, the optimization problem reduces to fitting  $R_n$  and  $i_{0,n}$  in the electrochemical model equation (Eq. (8)) to each polarization curve  $n$ , which was done using a least-square method. The least-squares objective function, defined as the sum of squared residuals, is strictly convex in the fit parameters, ensuring a unique global minimum and robust convergence (see Supplementary Information for a detailed derivation of convexity and the Hessian analysis). The implementation was carried out using Python's SciPy package, specifically the `least_squares` function, which reliably finds the optimal solution.

### 2.3.4. Parameter regression

The polarization curve fitting results in an array of  $N$  values for  $R_n$  and  $i_{0,n}$ , corresponding to  $N$  time intervals and polarization curves. Each value is assigned a timestamp defined as the mean timestamp of the respective interval  $n$ . To predict parameter trends, a statistical model (e.g., linear or exponential regression) is selected and fitted to the obtained parameter arrays using the least squares method. Prediction accuracy depends strongly on selecting a model appropriate for the observed evolution of  $R_n$  and  $i_{0,n}$ .

However, literature on the temporal evolution of these parameters is scarce and, beyond 1000 h, not available at all. Suermann et al. (2019) [46] report an increase in the Tafel slope and apparent exchange current density over a measurement period of 350 h, with no clear trend extractable for two CCMs tested at two constant currents. Rakousky et al. (2017) [28] report a linear trend for resistance and an exponential trend for the exchange current density, based on the fitting of seven polarization curves over 1000 h for five CCMs tested under different constant and intermittent accelerated stress tests. Overall, in literature reported voltage decay rates typically range between 0 and 200  $\mu\text{V/h}$ , depending on experimental conditions, materials, production methods, and test durations [7,27]. Consequently, the degradation trends of  $R$  and  $i_0$  cannot be assumed to always follow the same statistical model.

Beyond these statistical considerations, the interpretability of the extracted parameters is limited by the layer-unspecific model formulation and the simplifications made in Section 2.3.3. The fitted resistance  $R$  reflects the sum of all ohmic contributions and therefore cannot be assigned to specific components, although membrane resistance is typically dominant and an increasing  $R$  is thus most plausibly linked to membrane aging [12,27]. Likewise, the apparent exchange current density  $i_0$  combines the catalyst specific exchange current density, the catalyst loading and the specific ECSA, so that a decreasing  $i_0$  indicates anode catalyst degradation [28] but does not resolve whether agglomeration, dissolution or migration is responsible [47]. Finally, the diagnostic resolution of  $R$  and  $i_0$  decreases at stack or module level, where averaging over heterogeneous cells or stacks obscures the attribution of trends to specific physical modes [48,49].

To account for potential variations in the temporal evolution of these trends, the model is able to calculate multiple fitting approaches: constant, linear (Eq. (9)), and exponential (Eq. (10)).

$$y = \chi' \cdot t + \chi \quad (9)$$

$$y = \gamma \cdot \exp(\gamma' \cdot t) \quad (10)$$

The variable  $y$  represents either of the regression parameters  $R$  or  $i_0$ . In the linear case (Eq. (9)),  $\chi$  denotes the intercept and  $\chi'$  the slope, describing a constant rate of change over time. In the exponential case (Eq. (10)),  $\gamma$  corresponds to the prefactor and  $\gamma'$  to the exponential growth or decay rate, capturing accelerated or decelerated parameter evolution. The selection of a suitable statistical model is carried out iteratively by comparing regression results and choosing the one that yields the best values according to the performance metrics defined in Section 2.3.6.

### 2.3.5. Uncertainty propagation and RUL prediction

The least-squares regression of statistical trends for  $R$  and  $i_0$ , together with a fixed Tafel slope  $\bar{b}_{N_b}$ , provides parameter estimates and associated uncertainties. These uncertainties must be propagated to the RUL estimation to quantify the associated confidence with the prediction [7]. Uncertainty propagation is carried out using Monte Carlo sampling:  $M = 1000$  random draws are taken from a multivariate normal distribution defined by the mean values and covariance matrix of the fitted parameters  $R$  and  $i_0$ . The Tafel slope is treated as deterministic,  $b(t) \equiv \bar{b}_{N_b}$  without additional uncertainty contribution. A detailed derivation of the residual variance, covariance estimation, and calculation of standard errors is given in the Supplementary Information.

To illustrate this with an example, assuming a linear trend for the resistance  $R$  and a constant trend for the exchange current density  $i_0$ , we obtain Eq. (11) as a time- and current-density-dependent expression for the cell voltage:

$$E_{\text{cell}}(t, i) = E_{\text{rev}}(T, p) + i \cdot (R + R' \cdot t) + \bar{b}_{N_b} \cdot \log\left(\frac{i}{i_0}\right). \quad (11)$$

For each Monte Carlo draw, values for  $R$ ,  $R'$ , and  $i_0$  are obtained and propagated through Eq. (11). The RUL prediction is then calculated by solving Eq. (12) with the `brentq` function from the Python SciPy package for the specified  $i_{\text{EoL}}$ . This results in a distribution of  $t_{\text{EoL}}$  values, from which the mean is used as a point estimate and the 95% quantiles are reported as confidence intervals.

$$0 = E_{\text{cell}}(t_{\text{EoL}}, i_{\text{EoL}}) - E_{\text{cell, EoL}}. \quad (12)$$

The definition of the failure threshold for the cell voltage strongly affects the EoL estimate and is highly system-dependent. It is influenced by profitability considerations and equipment constraints, such as the maximum voltage of electrical hardware [7]. While there is no universal consensus in the literature, the U.S. Department of Energy has suggested a 10% increase relative to the begin-of-life (BoL) voltage as a practical failure threshold [50].

### 2.3.6. Performance metrics

PHM performance metrics must account for the continuous update of the predicted RUL as new data becomes available [7]. Consequently, early-stage predictions are based on less data than predictions closer to the system's EoL. The literature distinguishes between offline and online metrics [8]. Offline metrics are best suited for the development and validation of PHM approaches, as they rely on full datasets and knowledge of the ground truth EoL. Online metrics, by contrast, are used during operation when the true EoL is unknown. Since the datasets used in this work include the true EoL, the evaluation focuses on offline metrics such as the prognostic horizon (PH),  $\alpha$ - $\lambda$  accuracy, relative accuracy (RA), and convergence, following the definitions of Saxena et al. (2010) [8].

The PH is defined as the difference between the ground truth time of the EoL,  $t_{\text{EoL}}^*$ , and the earliest time  $t_{j_{\alpha\beta}}$  at which the predicted RUL satisfies the  $\beta$ -criterion based on the data available up to that point (see Eq. (13)).

$$\text{PH} = t_{\text{EoL}}^* - t_{j_{\alpha\beta}} \quad (13)$$

Here,  $t_{j_{\alpha\beta}} = \min\{t_n \in t_p \mid \pi[\text{RUL}(t_n)]_{\alpha}^{\pm} \geq \beta\}$  is the first time at which the predicted RUL distribution meets the  $\beta$ -criterion for a given  $\alpha_{\text{err}}$ .  $t_p$  is the set of all prediction time points, and  $\text{RUL}(t_n)$  is the predicted RUL distribution at time  $t_n$ . The probability mass  $\pi[\text{RUL}(t_n)]_{\alpha}^{\pm}$  is evaluated within the fixed bounds  $\alpha_{\text{PH}}^{\pm} = \text{RUL}^*(t_n) \pm \alpha_{\text{err}} \cdot t_{\text{EoL}}^*$ , with  $\text{RUL}^*(t_n)$  denoting the ground truth RUL. The parameter  $\alpha_{\text{err}}$  reflects the tolerance for prediction error, and  $\beta$  defines the minimum required confidence.

The  $\alpha$ - $\lambda$  accuracy is a binary metric that evaluates whether the predicted RUL at time  $t_{\lambda}$  lies within dynamic  $\alpha_{\text{err}}$ -bounds around the true RUL (see Eq. (14)). In contrast to PH, these bounds scale with the predicted RUL itself, resulting in a narrowing accuracy cone as the EoL approaches. The metric returns 1 if the probability mass within the bounds exceeds  $\beta$ , and 0 otherwise.

$$\alpha\text{-}\lambda \text{ accuracy} = \begin{cases} 1 & \text{if } \pi[\text{RUL}(t_{\lambda})]_{\alpha}^{\pm} \geq \beta \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

The evaluation time  $t_{\lambda}$  is defined as  $t_{\lambda} = t_p + \lambda \cdot (t_{\text{EoL}} - t_p)$ , where  $t_p$  is the time of the first prediction and  $\lambda \in [0, 1]$  indicates the relative position within the prognostic window. The probability mass  $\pi[\text{RUL}(t_{\lambda})]_{\alpha}^{\pm}$  is computed over bounds given by  $\alpha_{\lambda}^{\pm} = \text{RUL}^*(t_{\lambda}) \pm \alpha_{\text{err}} \cdot \text{RUL}^*(t_{\lambda})$ . This metric is stricter than the PH, as it requires predictions to stay within a cone of accuracy.

The RA quantifies the normalized prediction error with respect to the true RUL at time  $t_{\lambda}$  (Eq. (15)). In this work, the mean of the predicted RUL distribution is used as the point estimate.

$$\text{RA}(t_{\lambda}) = 1 - \frac{|\text{RUL}^*(t_{\lambda}) - \mathbb{E}[\text{RUL}(t_{\lambda})]|}{\text{RUL}^*(t_{\lambda})} \quad (15)$$

Here,  $\mathbb{E}[\text{RUL}(t_{\lambda})]$  denotes the expected value of the predicted RUL at  $t_{\lambda}$ . Unlike the binary  $\alpha$ - $\lambda$  accuracy, RA provides a continuous measure of prediction accuracy.

Convergence is used here as a meta-metric to quantify how quickly RA improves over time. Following Saxena et al. (2010) [8], it is computed as the Euclidean distance between the origin and the center of mass of the RA curve between  $t_p$  and  $t_{\text{EoL}}$  (see Eq. (16)).

$$C_{\text{RA}} = \sqrt{(x_c - t_p)^2 + y_c^2} \quad (16)$$

Here,  $(x_c, y_c)$  is the center of mass of the RA curve, defined as:

$$x_c = \frac{0.5 \cdot \sum_{n=p}^{N_{\text{EoL}}} (t_{n+1}^2 - t_n^2) \cdot \text{RA}(t_n)}{\sum_{n=1}^{N_{\text{EoL}}} (t_{n+1} - t_n) \cdot \text{RA}(t_n)} \quad (17)$$

$$y_c = \frac{0.5 \cdot \sum_{n=p}^{N_{\text{EoL}}} (t_{n+1} - t_n) \cdot \text{RA}(t_n)^2}{\sum_{n=1}^{N_{\text{EoL}}} (t_{n+1} - t_n) \cdot \text{RA}(t_n)}$$

**Table 1**

Relative errors (%) between fitted model parameters and their ground truth values, evaluated at increasing levels of relative measurement inaccuracy ( $f_{\text{rel,max}} = 0\%$ ,  $1\%$ , and  $3\%$ ).

Parameter	Ground truth	Relative error in %		
		$f_{\text{rel,max}} = 0\%$	$f_{\text{rel,max}} = 1\%$	$f_{\text{rel,max}} = 3\%$
$R$	80.6 mΩ cm <sup>2</sup>	0.27	-1.48	14.87
$R'$	$5.08 \times 10^{-3}$ mΩ cm <sup>2</sup> h <sup>-1</sup>	-0.04	-0.26	-2.64
$\bar{b}_{N_b}$	65 mV dec <sup>-1</sup>	-0.09	-0.077	-25.1
$i_0$	$1.01 \times 10^{-6}$ A cm <sup>-2</sup>	-1.05	-10.54	-97.97

where  $t_n$  denotes the discrete evaluation times, and  $\text{RA}(t_n)$  is the relative accuracy at  $t_n$ , as defined in Eq. (15).  $N_{\text{EoL}}$  is the index corresponding to the final evaluation time at the EoL. A lower value of  $C_{\text{RA}}$  indicates faster convergence, as it implies that high accuracy is achieved earlier in the prognostic window.

### 3. Results and discussion

#### 3.1. Synthetic data

First, data processing is applied to exclude data points where the PEMWE stacks are not in operation. Specifically, all data points with a cell voltage below  $E_{\text{cell}} < 1.25$  V are removed, which approximately corresponds to the Nernst potential under the assumed operating conditions. Data points with current density below  $i < 0.005$  A cm<sup>-2</sup> are also excluded to avoid convergence issues during least-squares fitting of the polarization curves.

The applied data sectioning criteria are  $\Delta i = 1.6$  A cm<sup>-2</sup> to cover most of the operating range of  $i_{\text{max}} = 2$  A cm<sup>-2</sup>, a minimum number of data points  $n_{t,\text{min}} = 300$  to ensure sufficient data density within each interval, and a maximum number  $n_{t,\text{max}} = 2000$  to ensure quasi-steady-state conditions. Using these criteria, the number of extracted polarization curves across the different levels of relative inaccuracy ranged between 199 and 208, with an average duration per extracted segment between 124.8 h and 130.5 h. Given a known cell voltage decay rate of  $\delta E_{\text{cell}} = 10$  μV/h, this corresponds to an absolute voltage degradation of approximately 1.3 mV at  $i_{\text{max}} = 2$  A cm<sup>-2</sup>, thereby satisfying the quasi-stationarity condition for the extracted polarization curves. Fig. 4a-c displays the extracted polarization curves at the BoT and the end of test (EoT), together with the fitted polarization curves utilizing Eq. (8) for the varied levels of relative inaccuracy.

Fig. 4d-f shows the time evolution of the extracted parameters  $R_n$  and  $i_{0,n}$ , along with the fitted parameter trends. To extract a stable value for the Tafel slope  $\bar{b}_{N_b}$ , the first  $N_b = 4$  polarization curves are used. A linear trend is applied to the resistance  $R$ , while a constant trend is used for the exchange current density  $i_0$ . These extracted trends align with the synthetic dataset. Specifically, the linear voltage decay rate ( $\delta E_{\text{cell}} = 10$  μV/h) mathematically translates into a linear increase in resistance and a constant  $i_0$ , as depicted in Eq. (18).

$$\begin{aligned}
 E_{\text{cell}} &= E_{\text{rev}} + i \cdot R + b \cdot \log\left(\frac{i}{i_0}\right) + \delta E_{\text{cell}} \cdot \frac{i}{i_{\text{max}}} \cdot t \\
 &= E_{\text{rev}} + i \cdot \left(R + \frac{\delta E_{\text{cell}}}{i_{\text{max}}} \cdot t\right) + b \cdot \log\left(\frac{i}{i_0}\right) \\
 &= E_{\text{rev}} + i \cdot (R + R' \cdot t) + b \cdot \log\left(\frac{i}{i_0}\right)
 \end{aligned} \quad (18)$$

Table 1 presents the relative error of the extracted model parameters compared to their ground truth values across increasing levels of relative measurement inaccuracy  $f_{\text{rel,max}}$ . As expected, the relative error of the fitted parameters generally increases with higher noise levels.

At  $f_{\text{rel,max}} = 3\%$ , the resistance  $R$ , the Tafel slope  $\bar{b}_{N_b}$ , and the exchange current density  $i_0$  exhibit the largest deviations from the ground truth. This originates from the fitting of the first  $N_b = 4$  polarization curves, where the root mean square error (RMSE) between

**Table 2**

Evaluation of  $\alpha$ - $\lambda$  accuracy, probability mass  $\pi[\text{RUL}(t_\lambda)]$ , and RA at different prediction times ( $\lambda = 0.25, 0.5$ , and  $0.75$ ), for varying levels of relative measurement inaccuracies ( $f_{\text{rel,max}} = 0\%$ ,  $1\%$ , and  $2\%$ ). Metrics were computed with accuracy margin  $\alpha_{\text{err},\lambda} = 0.2$  and confidence level  $\beta = 95\%$ .

$f_{\text{rel,max}}$	0%			1%			3%			
	$\lambda$	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75
$\alpha$ - $\lambda$ -accuracy( $t_\lambda$ )	1	1	1	0	0	1	0	0	0	0
$\pi[\text{RUL}(t_\lambda)]$	1	1	1	0.27	0.7	1.00	0.19	0.41	0.86	
$\text{RA}(t_\lambda)$	1.00	1.00	1.00	0.61	0.85	0.96	-1.11	0.73	1.00	

measured and modeled voltage is minimized with an unfixed Tafel slope. Because both the Tafel slope and the resistance enter linearly in the voltage model, a lower  $\bar{b}_{N_b}$  can be compensated by a higher  $R$  without affecting the RMSE. Once  $\bar{b}_{N_b}$  is fixed to its average value, this compensation effect propagates into all subsequent fits. The deviation of  $i_0$  can be explained similarly, the voltage model depends linearly on  $\bar{b}_{N_b}$  but logarithmically on  $i_0$ . A reduction in  $\bar{b}_{N_b}$  can therefore be balanced by a corresponding decrease in  $i_0$ , yielding an accurate voltage prediction but biased parameter values. These compensation mechanisms can persist in the regression procedure and distort the estimation of time-dependent degradation trends.

Because the extracted parameters serve as intermediate quantities to model voltage degradation, a comparison of the calculated cell voltage  $E_{\text{cell}}(t_n)$  with the ground truth  $E_{\text{cell}}^*(t_n)$  must also be considered to assess model quality. The resulting RMSE values of the predicted voltage are 0.1 mV at  $f_{\text{rel,max}} = 0\%$ , 19.9 mV at  $1\%$ , and 56.8 mV at  $3\%$ . These results demonstrate the robustness and numerical accuracy of the proposed method.

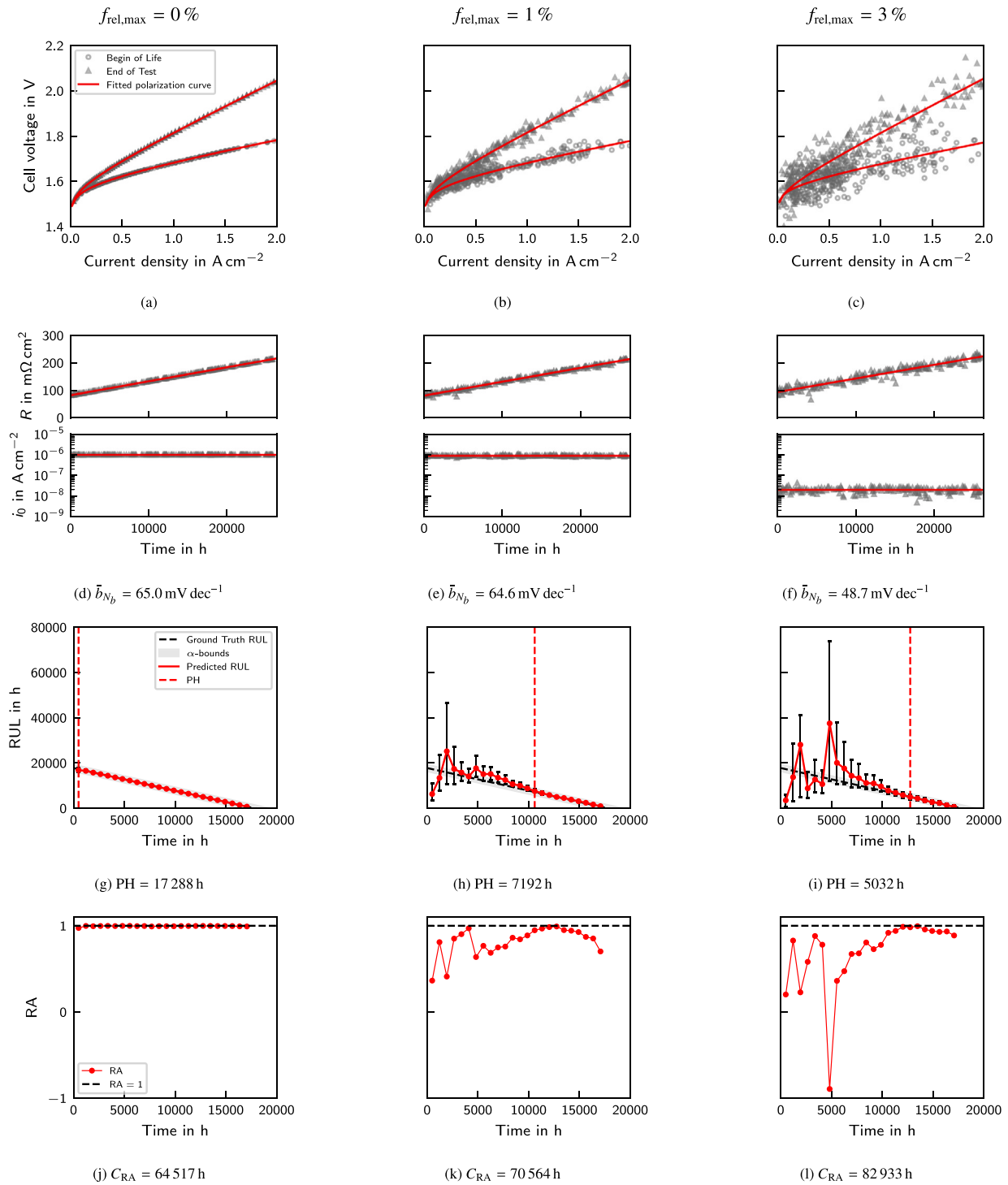
The aforementioned results are based on the full dataset. In the following passages, the PHM metrics described in Section 2.3.6 are explored in more detail. Figs. 4g-i show the predicted RUL values for a prediction start time of approximately  $t_p = 500$  h, compared to the ground truth RUL, thereby illustrating the PH. The ground truth EoL time,  $t_{\text{EoL}}^* = 17787$  h, is defined as the time at which the voltage increases by 10% compared to the BoL voltage at  $i_{\text{max}} = 2$  A cm<sup>-2</sup>, using the ground truth parameters  $R$ ,  $R'$ ,  $b$ , and  $i_0$  in Eq. (11).

The criteria for the PH are set to  $\alpha_{\text{err}} = 0.08$  and  $\beta = 95\%$ . This choice of  $\alpha_{\text{err}}$  corresponds to an absolute error margin around  $t_{\text{EoL}}^*$  of approximately two months, providing the plant operator sufficient time to schedule maintenance. As expected, the PH decreases with increasing measurement inaccuracy. For example, at a relative inaccuracy of  $3\%$ , the PH is reduced to  $\text{PH} = 5.172$  h. However, these PH values should be interpreted as comparative indicators for evaluating the approach rather than as absolute values applicable to a specific data quality.

Table 2 presents the  $\alpha$ - $\lambda$  accuracy, the probability mass  $\pi[\text{RUL}(t_\lambda)]$ , and the RA for  $\lambda = 0.25, 0.5$ , and  $0.75$  across the different levels of relative inaccuracy. The  $\alpha$ - $\lambda$  accuracy is evaluated using the criteria  $\alpha_{\text{err},\lambda} = 0.2$  and  $\beta = 95\%$ . These settings correspond to error margins of approximately four, two and a half, and one month around  $t_{\text{EoL}}^*$  for  $\lambda = 0.25, 0.5$ , and  $0.75$ , respectively. Therefore, presenting a stricter cut-off criterion compared to the PH due to the increasing accuracy requirement even with a higher  $\alpha_{\text{err}}$  value.

The  $\alpha$ - $\lambda$  accuracy criterion is fulfilled for all  $\lambda$  values at  $f_{\text{rel,max}} = 0\%$ . At  $f_{\text{rel,max}} = 1\%$ , it is only satisfied at  $\lambda = 0.75$ , and at  $f_{\text{rel,max}} = 3\%$ , the criterion is not met for any  $\lambda$ . This demonstrates that data quality strongly governs prediction accuracy. The increasing probability mass and the convergence of RA toward 1 with increasing  $\lambda$  further emphasize the need for larger datasets at higher levels of relative inaccuracy to achieve reliable voltage predictions.

This can be further illustrated by the visualization of the RA over time as the dataset grows until EoL in Figs. 4j-l. For higher levels of relative inaccuracy, the RA converges more slowly toward the ideal value of  $\text{RA} = 1$ . The slower convergence of RA can also be quantified



**Fig. 4.** Results from synthetic data at varying levels of relative measurement inaccuracies ( $f_{rel,max} = 0\%$ ,  $1\%$ , and  $3\%$ , left to right). Top row (a-c): Extracted polarization curves using data sectioning criteria ( $\Delta i = 1.6\ A\ cm^{-2}$ ,  $n_{i,min} = 300$ , and  $n_{i,max} = 2000$ ) with fitted polarization curves based on Eq. (8). Upper middle row (d-f): Temporal evolution of fitted parameters  $R_n$  and  $i_{0,n}$ , applying a linear trend for  $R$  (Eq. (9)) and a constant trend for  $i_0$ . Lower middle row (g-i): Point estimates of the RUL obtained from Eq. (11), including 95% confidence intervals and PH visualization with  $\alpha_{err} = 0.08$  and  $\beta = 95\%$ . Bottom row (j-l): Point estimates of the RA calculated according to Eq. (15).

by the increasing values of  $C_{RA}$ :  $C_{RA} = 64517\ h$  at  $f_{rel,max} = 0\%$ ,  $C_{RA} = 70564\ h$  at  $1\%$ , and  $C_{RA} = 82933\ h$  at  $3\%$ .

Nonetheless, even at early prediction stages, the RA values are close to 1, demonstrating good prediction quality. As with PH, the confidence in the RA estimates improves with a larger dataset, so some early-stage

outliers are to be expected. Negative RA values occur when the RUL is substantially overestimated, since in that case the numerator in Eq. (15) exceeds the denominator. Notably, for both  $f_{rel,max} = 1\%$  and  $f_{rel,max} = 3\%$ , a divergence from  $RA = 1$  towards negative values can be observed near  $t_{EoL}^*$ . This behavior is explained by the denominator in Eq. (15)

**Table 3**

Fitted regression parameters  $R$ ,  $R'$ ,  $\bar{b}_{N_b}$ , and  $i_0$  obtained using Eq. (11), along with the root-mean-square error  $\text{RMSE}_{E_{\text{cell}}(t_n)}$  between the modeled cell voltage  $E_{\text{cell}}(t_n)$  and the measured ground truth  $E_{\text{cell}}^*(t_n)$ , based on the full experimental dataset.

Parameter	Cell						
	Mean	1	2	3	4	5	6
$R$ in $\text{m}\Omega\text{cm}^2$	286	269	333	276	288	274	275
$R'$ in $\text{m}\Omega\text{cm}^2\text{h}^{-1}$	$8.03 \times 10^{-2}$	$2.90 \times 10^{-2}$	$1.71 \times 10^{-1}$	$3.72 \times 10^{-2}$	$1.08 \times 10^{-1}$	$2.96 \times 10^{-2}$	$1.07 \times 10^{-1}$
$\bar{b}_{N_b}$ in $\text{mVdec}^{-1}$	63.8	64.6	74.4	60.9	63.1	60.9	58.6
$i_0$ in $\text{Acm}^{-2}$	$1.57 \times 10^{-6}$	$1.84 \times 10^{-6}$	$5.51 \times 10^{-6}$	$1.09 \times 10^{-6}$	$1.48 \times 10^{-6}$	$1.04 \times 10^{-6}$	$6.95 \times 10^{-7}$
$\text{RMSE}_{E_{\text{cell}}(t_n)}$ in mV	13.6	14.7	31.2	15.1	12.7	13.2	9.1

approaching zero as EoL is reached which makes the equation more sensitive for overestimation of the RUL.

### 3.2. Experimental data

The experimental dataset is processed first according to the processing steps described in Section 2.3.1. Only data points with a cell voltage above 1.25 V and a current density above  $0.005\text{ A/cm}^2$  were considered. Additionally, the anode inlet temperature was restricted to the range of  $53^\circ\text{C}$  to  $63^\circ\text{C}$ , and the cathode pressure to 0.5 bar to 2 bar. This processing ensures quasi-steady-state operating conditions as required by the applied degradation model.

The data sectioning criteria were set to  $\Delta i = 1.05\text{ Acm}^{-2}$ ,  $n_{i_n,\text{min}} = 13$ , and  $n_{i_n,\text{max}} = 20$ . Based on these parameters, a total of eight polarization segments were extracted, corresponding exactly to the number of measured polarization curves. The average duration per segment was 0.7 h, which satisfies the quasi-stationarity requirement, given the stack voltage decay rate of approximately  $\delta E_{\text{cell}} = 75\text{ }\mu\text{V/h}$  per-cell (assuming a linear degradation behavior over the whole testing duration). The extracted polarization curves at BoT and EoT are displayed in Fig. 5a-f.

Table 3 summarizes the extracted model parameters for the mean cell and each of the six individual cells, assuming a linear time dependence of the resistance and a constant exchange current density. Most extracted parameter values lie within a similar range across the cells. However, the coated cells (cells 2, 4, and 6) exhibit a resistance slope  $R'$  that is approximately one order of magnitude higher than that of the uncoated reference cells (cells 1, 3, and 5), indicating a significantly higher voltage decay rate associated with the experimental coatings. Fig. 5g-l shows the temporal evolution of the extracted parameters. The observed increase in resistance combined with the nearly constant exchange current density is consistent with trends reported in the literature for constant-current testing [12,27]. For example, Rakousky et al. (2017) observed an increase in ohmic cell resistance due to passivation of titanium components without protective coating, while the exchange current density remained largely unchanged under constant-current operation [28]. These parallels strengthen the validity of the extracted parameter trends in this study. Notably, the increase in resistance of the coated cells is considerably steeper compared to the reference cells, suggesting that the experimental coating introduces additional ohmic losses.

The extracted BoT resistance values  $R$  are higher than typical literature values for the HFR in PEMWE single cells, which are reported to range from  $150\text{ m}\Omega\text{cm}^2$  to  $240\text{ m}\Omega\text{cm}^2$  under comparable operating conditions (temperature, pressure, membrane type, catalyst, and loading) [28,51]. This deviation can be attributed to the inherently higher internal resistance of the research stack compared to single-cell systems. Supporting this hypothesis, impedance measurements at BoT show real parts of the impedance ranging from  $288\text{ m}\Omega\text{cm}^2$  to  $392\text{ m}\Omega\text{cm}^2$ , which are in good agreement with the extracted  $R$  values.

The elevated impedance values may also stem from limitations of the measurement method. The applied frequency of 1 kHz may not coincide with the true high-frequency intercept of the x-axis in the Nyquist plot, which defines the HFR. At lower frequencies, charge-transfer and double-layer effects contribute, while at higher frequencies parasitic inductances can increase the apparent resistance [52,53].

At EoT, the real part of the impedance measurements ranged from  $258\text{ m}\Omega\text{cm}^2$  to  $1028\text{ m}\Omega\text{cm}^2$ , which corresponds well to the resistance values extracted with the model after 2200 h of operation. These extracted values lie between  $333\text{ m}\Omega\text{cm}^2$  to  $706\text{ m}\Omega\text{cm}^2$ . The model particularly captures the steeper resistance increase in the coated research cells. However, for the uncoated commercial reference cells, a decrease in the real part of the impedance between BoT and EoT was observed, which is not reflected in the extracted model parameters. Detailed impedance data can be found in Table S3 of the Supplementary Information.

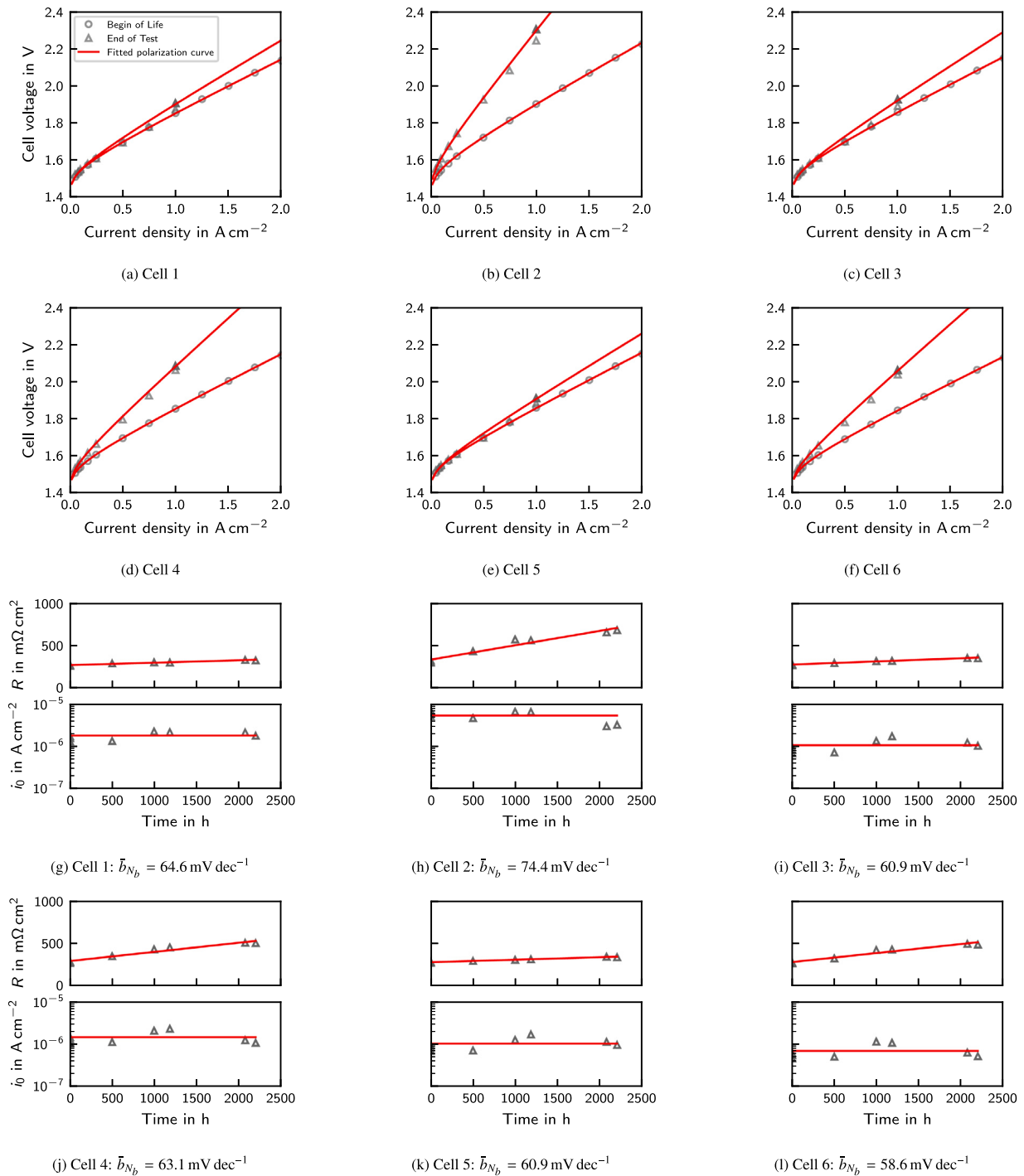
The extracted Tafel slopes range from  $58.6\text{ mVdec}^{-1}$  to  $74.4\text{ mVdec}^{-1}$ . Literature values are typically determined from linear fits of iR-corrected polarization data in the low-current region ( $10\text{ mAcm}^{-2}$  to  $100\text{ mAcm}^{-2}$ ), where kinetic control dominates and contributions from ohmic or mass transport limitations are negligible [33]. In contrast, the slopes reported here are obtained from full quasi-steady-state polarization curves. As a result, they inherently incorporate additional voltage losses beyond the pure kinetic overpotential. Consequently, the fitted average Tafel slope  $\bar{b}_{N_b}$  may slightly overestimate the intrinsic Tafel slope in the low-current region [43]. Nonetheless, the extracted values lie at the upper end of the literature range for Ir-based catalysts in PEMWE, typically reported between  $40\text{ mVdec}^{-1}$  to  $80\text{ mVdec}^{-1}$ , depending on catalyst type, loading, and operating conditions [54–56].

The extracted exchange current densities lie in the range of  $6.95 \times 10^{-7}\text{ Acm}^{-2}$  to  $5.51 \times 10^{-6}\text{ Acm}^{-2}$ , which is approximately one to two orders of magnitude higher than literature values reported for Ir-based catalysts at  $60^\circ\text{C}$ , typically ranging from  $1.5 \times 10^{-8}\text{ Acm}^{-2}$  to  $2.5 \times 10^{-7}\text{ Acm}^{-2}$  [46,57]. This discrepancy is consistent with the elevated Tafel slopes discussed above. In the literature, apparent exchange current densities are usually extracted by extrapolating the Tafel slope, which is determined in the kinetically controlled low-current region, to its intercept at  $\eta_{\text{OER}} = 0\text{ V}$ . Rogler et al. (2023) demonstrated that even small deviations in the extracted Tafel slope can lead to substantial changes in the estimated exchange current density, often by several orders of magnitude [43]. As discussed, the slopes extracted here likely include additional losses, leading to systematically higher  $b$  values and, consequently, an overestimation of the apparent exchange current density.

Despite these differences, most extracted parameters remain within physically reasonable bounds and in good agreement with literature ranges. This supports the physical validity and robustness of the parameterization. However, due to the modeling simplifications described in Section 2.3, the extracted parameters should not be interpreted as exact physical quantities.

Physical accuracy of the parameters is not a requirement for the underlying PHM objective, which is to accurately model voltage degradation over time. Instead, model quality is quantitatively assessed by the root-mean-square error  $\text{RMSE}_{E_{\text{cell}}(t_n)}$  between the modeled cell voltage  $E_{\text{cell}}(t_n)$  and the measured ground truth  $E_{\text{cell}}^*(t_n)$  across the entire experimental dataset. The resulting RMSE values range from 9.1 mV to 31.2 mV, indicating high overall model accuracy (see Table 3).

The PHM suitability of the approach for the experimental dataset can be further evaluated using the PHM metrics outlined in Section 2.3.6. As described earlier, the coated cells exhibit higher voltage decay rates compared to the uncoated cells. However, a differentiation



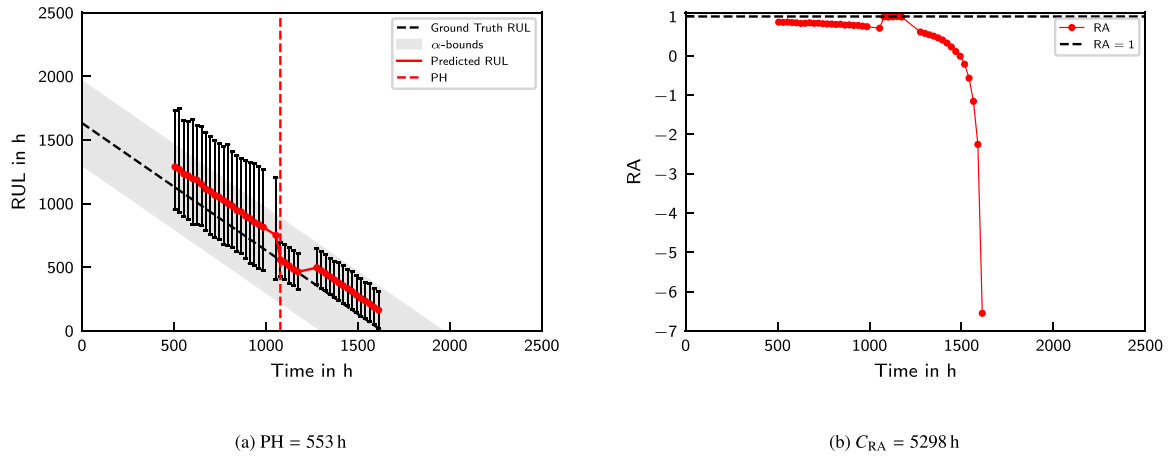
**Fig. 5.** Results from the experimental dataset. Top rows (a-f): Extracted polarization curves using data sectioning criteria ( $\Delta t = 1.05 \text{ A cm}^{-2}$ ,  $n_{t,\min} = 13$ , and  $n_{t,\max} = 20$ ) with fitted polarization curves based on Eq. (8). Bottom rows (g-l): Temporal evolution of fitted parameters  $R_n$  and  $i_0$ , applying a linear trend for  $R$  (Eq. (9)) and a constant trend for  $i_0$ .

between individual cells for estimating the EoL of the stack based on voltage degradation is not necessary. The voltage failure threshold of a PEMWE stack is ultimately determined by the overall stack voltage or the mean cell voltage, since the limiting factors are typically equipment-related constraints, such as the maximum allowable voltage of the power converters [7]. Therefore, the following results were obtained with the mean cell voltage of the stack.

The EoL voltage failure threshold is defined as a 10% increase relative to the BoT voltage at  $i = 1 \text{ A cm}^{-2}$ , which corresponds to  $E_{\text{cell,EoL}} = 2 \text{ V}$ . The ground truth EoL time  $t_{\text{EoL}}^* = 1632 \text{ h}$  is determined as the

intersection point between the mean cell voltage, which is smoothed with a moving average filter over 24 h and the threshold voltage  $E_{\text{cell,EoL}}$  (see Supplementary Information, Figure S4). It is important to note that, although this EoL threshold is consistent with the definition used by the U.S. Department of Energy [50], it must also fall within the duration of the available dataset in order to be used as a ground truth reference. Only then can the predicted voltage trajectory be validated using offline PHM performance metrics.

Fig. 6a shows the predicted RUL of the mean cell voltage for a prediction start time of approximately  $t_p = 500 \text{ h}$ , compared to the



**Fig. 6.** Results from the experimental data. Upper graph (a): Point estimates of the RUL obtained from Eq. (11), including 95 % confidence intervals and PH visualization with  $\alpha_{err} = 0.2$  and  $\beta = 95\%$ . Lower Graph (b): Point estimates of the RA calculated according to Eq. (15).

**Table 4**

Evaluation of  $\alpha$ - $\lambda$  accuracy, probability mass  $\pi[RUL(t_\lambda)]$ , and RA at different prediction times ( $\lambda = 0.25, 0.5, \text{ and } 0.75$ ), for the experimental dataset. Metrics were computed with accuracy margin  $\alpha_{err,\lambda} = 0.3$  and confidence level  $\beta = 95\%$ .

$\lambda$	0.25	0.5	0.75
$\alpha$ - $\lambda$ -accuracy( $t_\lambda$ )	0	0	0
$\pi[RUL(t_\lambda)]$	0.01	0.66	0.36
RA( $t_\lambda$ )	-0.27	0.80	0.60

ground truth RUL, illustrating the PH. The PH criteria are set to  $\alpha_{err} = 0.2$  and  $\beta = 95\%$ , where the chosen  $\alpha_{err}$  corresponds to an absolute error margin of approximately two weeks around  $t_{EoL}^*$ . The extracted PH is approximately three weeks (PH = 553 h).

Table 4 presents the  $\alpha$ - $\lambda$  accuracy, the probability mass  $\pi[RUL(t_\lambda)]$ , and the RA for  $\lambda = 0.25, 0.5, \text{ and } 0.75$  across different levels of relative inaccuracy. The  $\alpha$ - $\lambda$  accuracy is evaluated using the criteria  $\alpha_{err,\lambda} = 0.3$  and  $\beta = 95\%$ . These settings correspond to error margins of approximately 15, 10, and 4 days around  $t_{EoL}^*$  for  $\lambda = 0.25, 0.5, \text{ and } 0.75$ , respectively. The  $\alpha$ - $\lambda$  accuracy criterion is not fulfilled at any of the evaluated times, indicating that the approach cannot reliably achieve day-level accuracy for this specific dataset.

These results must be interpreted in the context of the dataset and method: Due to the nature of the data sectioning method described in Section 2.3.2, only a limited number of polarization curves are available for parameter fitting (see, Fig. 5g-l). This results in higher predictive uncertainty compared to datasets with greater temporal resolution and variability, such as the synthetic dataset, which allow more frequent curve extraction. Nevertheless, achieving the observed accuracy with only eight polarization curves demonstrates the method's feasibility for constant operated plants when occasional polarization curves can be obtained.

The point estimates of the RA displayed in Fig. 6b remain close to RA = 1 even at early prediction times, indicating overall good prediction quality. As observed in the synthetic dataset, a divergence toward negative RA values occurs near  $t_{EoL}^*$ . This behavior, as previously discussed, results from the denominator in Eq. (15) approaching zero near EoL, making the RA metric increasingly sensitive to even slight overestimations of the RUL.

### 3.3. Comparative analysis of PHM performance metrics across datasets

To enable comparison across the long synthetic dataset and the shorter experimental dataset, the PH and  $C_{RA}$  must be normalized. The normalized PH is defined as the ratio of the absolute PH to the total runtime,  $\hat{PH} = PH/t_{EoL}^*$ . While the absolute PH of the experimental

dataset appears small compared to the synthetic dataset, where PH values range from 5032 h to 17 288 h depending on the applied maximum relative inaccuracy  $f_{rel,max}$  (0% to 3%), the normalized PH of the experimental data ( $\hat{PH} = 0.34$ ) lies within the same order of magnitude. In the synthetic dataset, normalized PH values range from 0.28 ( $f_{rel,max} = 3\%$ ) to 0.97 ( $f_{rel,max} = 0\%$ ), indicating that the experimental data offers a comparable prognostic reach to the synthetic data under the investigated relative accuracy criteria.

Similarly, the value of  $C_{RA} = 5298$  h appears relatively small compared to the values obtained from the synthetic dataset. However, due to the strong dependency of  $C_{RA}$  on the temporal extent of the dataset,  $C_{RA}$  is normalized by the corresponding  $t_{EoL}^*$  according to  $\hat{C}_{RA} = C_{RA}/t_{EoL}^*$ . This yields a normalized value of  $\hat{C}_{RA} = 3.2$  for the experimental dataset and values in the range of  $\hat{C}_{RA} = 3.6$  to 4.7 for the synthetic dataset, indicating a slightly faster convergence of the RA in the experimental case.

Overall, the similarity of normalized PHM performance metrics between the experimental and synthetic datasets indicates that a combined validation strategy is viable. Moreover, the absolute RMSE  $E_{cell}(t_n)$  being overall in the two digit mV range strengthens the comparability of results across the datasets. Therefore, we argue that the long-duration synthetic dataset establishes stability and uncertainty trends over industry relevant lifetimes, while the shorter experimental dataset confirms applicability under real operating conditions, despite its limited duration.

To place these results in a broader context, Table 5 shows that the obtained voltage errors over full datasets are of the same order of magnitude as recent physical, empirical, and data-driven approaches. A direct comparison of PHM-specific performance metrics is not possible because these metrics are not reported in those studies. Notably, the proposed method also maintains high accuracy in the low-current-density region, where the logarithmic activation contribution dominates. Its low computational effort — relying solely on least-squares polarization-curve fits and low-order regression — further highlights its practical advantages.

Beyond the presented datasets, the method is generally applicable to any dataset that contains quasi-steady-state polarization-curve segments. Thus, it remains valid for both steady-state and dynamic operation provided that (i) operating conditions stay within defined stability bounds and (ii) sufficient current-density variation is present. Under strictly constant operation, fewer usable segments are obtained, which increases uncertainty and reduces prognostic horizons.

Finally, the combined use of long term synthetic and short-term experimental data allows researchers and operators to assess prognostic capability without multi-year field trials, thereby accelerating model deployment and reducing risks associated with unvalidated lifetime

**Table 5**

Accuracy comparison of the model against literature. Ranges denote the spread reported by each source (noise levels or operating regions). Direct numerical comparison is limited by differing datasets and operating profiles.

Source	Approach	RMSE in mV	MSE in V <sup>2</sup>	Absolute error in %
Synthetic data (low to high noise)	physical	0.1 to 56.8	$7.17 \times 10^{-9}$ to $3.22 \times 10^{-3}$	$3.3 \times 10^{-3}$ to 2.8 (MAE)
Experimental data (mean cell voltage)	physical	13.6	$1.85 \times 10^{-4}$	0.67 (MAE)
Pape et al. (2025) [20]	physical	29.7	–	–
Woelke et al. (2025) [19] (high to low current densities)	data-driven	–	–	2 to 13
Xu et al. (2024) [18] (current profile/ model dependent)	data-driven	–	$1.18 \times 10^{-8}$ to $2.08 \times 10^{-4}$	–
Yan et al. (2024) [14] (current region/ model dependent)	empirical	–	$5.4 \times 10^{-5}$ to $3.9 \times 10^{-3}$	–

predictions. From an industrial perspective, the resulting RUL forecasts enable predictive maintenance and degradation-aware operation [7]. This supports cost-optimal scheduling of stack replacement and reduces unplanned downtime, suggesting operating-cost benefits similar to those reported for PHM in related electrolysis technologies [58]. Furthermore, degradation-aware operation has been indicated to reduce operating costs in techno-economic case studies for PEMWE plants [59, 60].

#### 4. Conclusion

This study presented a layer–unspecific physical model to predict voltage degradation and RUL of PEMWE suited for plant-level operation data. Quasi-steady-state polarization curves are extracted in situ, without diagnostic downtime, and fitted with a simplified electrochemical expression. As a central methodological feature, the model applies a fixed Tafel slope to avoid parameter collinearity and enable stable tracking of the temporal evolution of resistance and exchange current density.

Monte Carlo sampling propagates parameter regression uncertainties into RUL predictions, enabling confidence levels to be quantified. Validation against a 26 200 h synthetic dataset and a 2200 h experimental dataset demonstrated millivolt-scale voltage errors and prognostic horizons sufficient for practical maintenance planning, while relying only on plant-typical voltage and current signals.

A key constraint is the required structure of the data. Extracting polarization curves demands stable operation and sufficient current-density variability within short windows; under strictly constant operation, only a few usable segments are obtained, which increases predictive uncertainty and shortens prognostic horizons. This limitation was evident in the experimental dataset, yet the method still achieved sufficient accuracy when applied to stack mean voltage, underlining feasibility in data-sparse settings.

The electrochemical description makes simplifying assumptions: a time-invariant Tafel slope, omission of explicit mass-transport effects to avoid overfitting, and reliance on predefined statistical regression trends. As a result, parameters should be interpreted as apparent values for prediction rather than exact descriptors of electrode processes. Nevertheless, extracted values aligned with literature ranges, and the regression proved robust under steady conditions, capturing trends consistent with experimental observations.

These limitations open opportunities for further work, including explicit treatment of temperature and pressure effects, more adaptive handling of changes in operation behavior through filtering or machine learning for parameter trend prediction, and validation across broader long-term datasets and operating regimes.

In conclusion, the proposed model provides a robust, plant-level–suited method to predict PEMWE voltage decay and deliver RUL estimates with quantified uncertainty. The integration of synthetic data further extends validation to industry-relevant lifetimes without the need for multi-year experiments, thereby providing a solid basis for efficiency management and predictive maintenance of PEMWE plants.

#### Glossary

Symbols and corresponding SI units		
$b$	Tafel slope	mV dec <sup>-1</sup>
$\bar{b}_{N_b}$	Mean Tafel slope fitted over $N_b$ polarization curves	mV dec <sup>-1</sup>
$C_{RA}$	Convergence of the RA	h
$\hat{C}_{RA}$	Normalized convergence of the RA	–
$E$	Voltage	V
$\delta E$	Voltage decay rate	$\mu\text{V h}^{-1}$
$F$	Faraday constant	C mol <sup>-1</sup>
$f_{rel,max}$	Relative maximum inaccuracy	%
$i$	Current density	A cm <sup>-2</sup>
$i_0$	Exchange current density	A cm <sup>-2</sup>
$\Delta i$	Current density interval	A cm <sup>-2</sup>
$j$	Index parameter	–
$N_{EoL}$	Index of final time instance at EoL	–
$n$	Index parameter	–
$n_{t_n,max}$	Maximum number of data points of extracted polarization curve	–
$n_{t_n,min}$	Maximum number of data points of extracted polarization curve	–
$R$	Resistance	m $\Omega$ cm <sup>2</sup>
$\bar{R}$	Ideal gas constant	J mol <sup>-1</sup> K <sup>-1</sup>
$P$	Power	W
PH	Prognostic horizon	h
$\bar{P}\bar{H}$	Normalized prognostic horizon	–
$p$	Pressure	Pa
$p_{H_2O}$	Water partial pressure	Pa
$p_{s,H_2O}$	Saturated water pressure	Pa
$p^*$	Reference pressure	Pa
RA	Relative accuracy	–
RMSE	Root mean square error	mV
RUL	Remaining useful life	h
$T$	Temperature	K
$t_{EoL}$	Time instant at EoL	s
$t_n$	Time instant	s
$t_p$	Time instant of first prediction	s
$t_P$	Set of all prediction time instances	s
$t_\lambda$	Time instant at $\lambda$	s
$x_c$	X-coordinate of the center of mass of the RA curve	s

$y_c$	Y-coordinate of the center of mass of the RA curve	s
$\alpha$	Charge transfer coefficient	–
$\alpha_{err}$	Specified error margin	–
$\beta$	Probability mass bounds	–
$\gamma$	Exponential coefficient	–
$\eta$	Overpotential	V
$\vartheta$	Temperature	°C
$\lambda$	Relative time instant	–
$\pi$	Probability mass	–
$\chi$	Linear coefficient	–

Sub- and superscripts	
a	Anode
c	Cathode
cell	PEMWE cell
cross	Crossover
eff	Effective
el	Electronic
F	Faraday
H <sup>+</sup>	Proton
mem	Membrane
mt	Mass transport
$N_b$	Number of fitted polarization curves for mean $b$ value
ohm	Ohmic
plant	PEMWE plant
rev	Reversible
stack	PEMWE stack
0	Standard
*	Ground truth

Abbreviations	
BoL	Begin of life
BoT	Begin of test
BPP	Bipolar plate
CCM	Catalyst coated membrane
ECSA	Electrochemically active surface area
EOl	End of life
EOt	End of test
HER	Hydrogen evolution reaction
HFR	High-frequency resistance
HI	Health indicator
OER	Oxygen evolution reaction
PEM	Proton exchange membrane
PEMWE	Proton exchange membrane water electrolysis
PHM	Prognostics and health management
PTL	Porous transport layer

### CRedit authorship contribution statement

**Felix Dittmar:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Thomas Lickert:** Writing – review & editing, Data curation. **Tom Smolinka:** Writing – review & editing. **Karsten Pinkwart:** Writing – review & editing. **Jens Tübke:** Writing – review & editing, Supervision.

### Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT by OpenAI in order to improve the spelling and general readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Felix Dittmar reports financial support was provided by German Federal Ministry of Education and Research (BMBF). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

The authors gratefully acknowledge the financial support from the German Federal Ministry of Education and Research (BMBF) through the project “hyBit” under the grant 03SF0687E.

### Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jpowsour.2025.239122>.

### Data availability

Experimental data are available from the corresponding author on reasonable request. The code for synthetic data generation and the degradation model are publicly available on GitHub and archived on Zenodo (<https://doi.org/10.5281/zenodo.18268174>).

### References

- [1] International Energy Agency, Hydrogen production and infrastructure projects database, 2025, URL <https://www.iea.org/data-and-statistics/data-product/hydrogen-production-and-infrastructure-projects-database>. (Accessed 30 September 2025).
- [2] International Energy Agency, Global hydrogen review 2025, 2025, URL <https://www.iea.org/reports/global-hydrogen-review-2025>. (Accessed 30 September 2025).
- [3] S.A. Grigoriev, V.N. Fateev, D.G. Bessarabov, P. Millet, Current status, research trends, and challenges in water electrolysis science and technology, *Int. J. Hydrog. Energy* 45 (49) (2020) 26036–26058, <http://dx.doi.org/10.1016/j.ijhydene.2020.03.109>.
- [4] A. Javed, N.L. Wolf, F. Meyer, L. Treutlein, H. Kungl, A. Karl, E. Jodan, R.-A. Eichel, Exploring the state-of-operation of proton exchange membrane electrolyzers, *Int. J. Hydrog. Energy* 98 (2025) 280–294, <http://dx.doi.org/10.1016/j.ijhydene.2024.12.055>.
- [5] N.-H. Kim, D. An, J.-H. Choi, Prognostics and Health Management of Engineering Systems, Springer International Publishing, Cham, 2017, <http://dx.doi.org/10.1007/978-3-319-44742-1>.
- [6] D. Thomas, B. Weiss, Maintenance costs and advanced maintenance techniques in manufacturing machinery: Survey and analysis, *Int. J. Progn. Health Manag.* 12 (1) (2021) <http://dx.doi.org/10.36001/ijphm.2021.v12i1.2883>.
- [7] F. Dittmar, H. Agarwal, J. Tübke, Prognostics and health management (PHM) of proton exchange membrane water electrolyzers: A review-based guideline, *Int. J. Hydrog. Energy* 106 (2025) 806–824, <http://dx.doi.org/10.1016/j.ijhydene.2025.01.443>.
- [8] A. Saxena, J. Celaya, B. Saha, S. Saha, K. Goebel, Metrics for offline evaluation of prognostic performance, *Int. J. Progn. Health Manag.* 1 (1) (2010) <http://dx.doi.org/10.36001/ijphm.2010.v1i1.1336>.
- [9] D. Bessarabov, P. Millet, PEM water electrolysis volume 2, in: *Hydrogen and Fuel Cells Primer Series, vol. 2*, Academic Press, London, 2018.
- [10] E. Wallnöfer-Ogris, I. Grimmer, M. Ranz, M. Höglinger, S. Kartusch, J. Rauh, M.-G. Macherhammer, B. Grabner, A. Trattner, A review on understanding and identifying degradation mechanisms in pem water electrolysis cells: Insights for stack application, development, and research, *Int. J. Hydrog. Energy* 65 (2024) 381–397, <http://dx.doi.org/10.1016/j.ijhydene.2024.04.017>.
- [11] International Organization for Standardization, ISO 13381-1:2015 condition monitoring and diagnostics of machines — prognostics: Part 1: General guidelines, (ISO 13381-1:2015) 2015.
- [12] A.Z. Tomić, I. Pivac, F. Barbir, A review of testing procedures for proton exchange membrane electrolyzer degradation, *J. Power Sources* 557 (2023) 232569, <http://dx.doi.org/10.1016/j.jpowsour.2022.232569>.

- [13] Q. Wen, L. Sun, F. Yang, X. Song, J. Gao, X. Wang, H. Xu, Time series data augmentation for deep learning: A survey, in: Z.-H. Zhou (Ed.), Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence // Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21, International Joint Conferences on Artificial Intelligence Organization and International Joint Conferences on Artificial Intelligence, 2021, pp. 4653–4660, <http://dx.doi.org/10.24963/ijcai.2021/631>.
- [14] X. Yan, C. Locci, F. Hiss, A. Niefse, State-of-health estimation for industrial H<sub>2</sub> electrolyzers with transfer linear regression, *Energies* 17 (6) (2024) 1374, <http://dx.doi.org/10.3390/en17061374>.
- [15] Z. Hua, Z. Zheng, E. Pahon, M.-C. Péra, F. Gao, A review on lifetime prediction of proton exchange membrane fuel cells system, *J. Power Sources* 529 (2022) 231256, <http://dx.doi.org/10.1016/j.jpowsour.2022.231256>.
- [16] B. Benschmann, L.V. Buehre, R. Hanke-Rauschenbach, Reference electrodes in PEM water electrolysis – a review and experimental investigation of oxygen and hydrogen evolution reaction kinetics, *ECS Meet. Abstr.* MA2022-01 (2022) 1372, <http://dx.doi.org/10.1149/MA2022-01341372mtgabs>.
- [17] M. Bahr, A. Gusak, S. Stypka, B. Oberschachtsiek, Artificial neural networks for aging simulation of electrolysis stacks, *Chem. Ing. Tech.* 92 (10) (2020) 1610–1617, <http://dx.doi.org/10.1002/cite.202000089>.
- [18] B. Xu, W. Ma, W. Wu, Y. Wang, Y. Yang, J. Li, X. Zhu, Q. Liao, Degradation prediction of PEM water electrolyzer under constant and start-stop loads based on CNN-LSTM, *Energy AI* 18 (2024) 100420, <http://dx.doi.org/10.1016/j.egyai.2024.100420>.
- [19] J. Woelke, A. Rex, C. Eckert, B. Benschmann, R. Hanke-Rauschenbach, Predicting future polarization curves from operating data: Machine learning-based investigation of degradation modeling concepts for PEM water electrolysis, *Energy AI* 21 (2025) 100547, <http://dx.doi.org/10.1016/j.egyai.2025.100547>.
- [20] S.-V. Pape, S. Zerressen, M.F. Seidler, R. Keller, F. Lohmann-Richters, M. Müller, U.-P. Apfel, A.K. Mechler, A. Glösen, Performance data extraction from dynamic long-term operation of proton exchange membrane and alkaline water electrolysis cells, *Int. J. Hydrog. Energy* 127 (2025) 51–63, <http://dx.doi.org/10.1016/j.ijhydene.2025.03.387>.
- [21] R. Heyne, *Lebensdauerprognose von elektrochemischen Systemen unter besonderer Berücksichtigung von Brennstoffzellen* (Ph.D. thesis), TU Clausthal, Clausthal, 2015.
- [22] Climate Data Center, Historische 10-minütige stationsmessung der windgeschwindigkeit in deutschland, version v24.03, 2025, URL [https://opendata.dwd.de/climate\\_environment/CDC/observations\\_germany/climate/10\\_minutes/wind/historical/](https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/wind/historical/). (Accessed 11 July 2025).
- [23] V112 3.0 MW onshore, 2011, URL [https://www.wind-still.ch/files/windstill/files/pdf/Vestas\\_V112\\_web\\_DE.pdf](https://www.wind-still.ch/files/windstill/files/pdf/Vestas_V112_web_DE.pdf). (Accessed 11 July 2025).
- [24] J. Hemaier, S. Rehfeldt, H. Klein, A. Peschel, Performance and cost modelling taking into account the uncertainties and sensitivities of current and next-generation PEM water electrolysis technology, *Int. J. Hydrog. Energy* 48 (66) (2023) 25619–25634, <http://dx.doi.org/10.1016/j.ijhydene.2023.03.050>.
- [25] H. Groenemans, G. Saur, C. Mittelstaedt, J. Lattimer, H. Xu, Techno-economic analysis of offshore wind PEM water electrolysis for H<sub>2</sub> production, *Curr. Opin. Chem. Eng.* 37 (2022) 100828, <http://dx.doi.org/10.1016/j.coche.2022.100828>.
- [26] J. Settino, R.N. Farrugia, D. Buhagiar, T. Sant, Offshore wind-to-hydrogen production plant integrated with an innovative hydro-pneumatic energy storage device, *J. Phys.: Conf. Ser.* 2151 (1) (2022) 012013, <http://dx.doi.org/10.1088/1742-6596/2151/1/012013>.
- [27] E. Kuhnert, V. Hacker, M. Bodner, A review of accelerated stress tests for enhancing MEA durability in PEM water electrolysis cells, *Int. J. Energy Res.* 2023 (2023) 3183108, <http://dx.doi.org/10.1155/2023/3183108>.
- [28] C. Rakousky, U. Reimer, K. Wippermann, S. Kuhri, M. Carmo, W. Lueke, D. Stolten, Polymer electrolyte membrane water electrolysis: Restraining degradation in the presence of fluctuating power, *J. Power Sources* 342 (2017) 38–47, <http://dx.doi.org/10.1016/j.jpowsour.2016.11.118>.
- [29] S.H. Frensch, F. Fouda-Onana, G. Serre, D. Thoby, S.S. Araya, S.K. Kær, Influence of the operation mode on PEM water electrolysis degradation, *Int. J. Hydrog. Energy* 44 (57) (2019) 29889–29898, <http://dx.doi.org/10.1016/j.ijhydene.2019.09.169>.
- [30] R. Keller, E. Rauls, M. Hehemann, M. Müller, M. Carmo, An adaptive model-based feedforward temperature control of a 100 kW PEM electrolyzer, *Control Eng. Pract.* 120 (2022) 104992, <http://dx.doi.org/10.1016/j.conengprac.2021.104992>.
- [31] T. Krenz, A. Rex, L. Helmers, P. Trinke, B. Benschmann, R. Hanke-Rauschenbach, Reversible degradation phenomenon in PEMWE cells: An experimental and modeling study, *J. Electrochem. Soc.* 171 (12) (2024) 124501, <http://dx.doi.org/10.1149/1945-7111/ad96e4>.
- [32] P. Olivier, C. Bourasseau, B. Bouamama, Dynamic and multiphysic PEM electrolysis system modelling: A bond graph approach, *Int. J. Hydrog. Energy* 42 (22) (2017) 14872–14904, <http://dx.doi.org/10.1016/j.ijhydene.2017.03.002>.
- [33] M. Bernt, H.A. Gasteiger, Influence of ionomer content in IrO<sub>2</sub>/TiO<sub>2</sub> electrodes on PEM water electrolyzer performance, *J. Electrochem. Soc.* 163 (11) (2016) F3179–F3189, <http://dx.doi.org/10.1149/2.0231611jes>.
- [34] P. Trinke, G.P. Keeley, M. Carmo, B. Benschmann, R. Hanke-Rauschenbach, Elucidating the effect of mass transport resistances on hydrogen crossover and cell performance in PEM water electrolyzers by varying the cathode ionomer content, *J. Electrochem. Soc.* 166 (8) (2019) F465–F471, <http://dx.doi.org/10.1149/2.0171908jes>.
- [35] S. Fahr, F.K. Engel, S. Rehfeldt, A. Peschel, H. Klein, Overview and evaluation of crossover phenomena and mitigation measures in proton exchange membrane (PEM) electrolysis, *Int. J. Hydrog. Energy* 68 (2024) 705–721, <http://dx.doi.org/10.1016/j.ijhydene.2024.04.248>.
- [36] S.G. Bratsch, Standard electrode potentials and temperature coefficients in water at 298.15 K, *J. Phys. Chem. Ref. Data* 18 (1) (1989) 1–21, <http://dx.doi.org/10.1063/1.555839>.
- [37] K. Sattler, *Thermische Trennverfahren: Grundlagen, Auslegung, Apparate*, 3. Auflage, VCH, Weinheim and Chichester, 2001.
- [38] A. Martin, P. Trinke, B. Benschmann, R. Hanke-Rauschenbach, Hydrogen crossover in PEM water electrolysis at current densities up to 10 A cm<sup>-2</sup>, *J. Electrochem. Soc.* 169 (9) (2022) 094507, <http://dx.doi.org/10.1149/1945-7111/ac908c>.
- [39] M. Bernt, J. Schröter, M. Möckl, H.A. Gasteiger, Analysis of gas permeation phenomena in a PEM water electrolyzer operated at high pressure and high current density, *J. Electrochem. Soc.* 167 (12) (2020) 124502, <http://dx.doi.org/10.1149/1945-7111/abaa68>.
- [40] C.H. Hamann, W. Vielstich, *Elektrochemie*, 4. Auflage, Wiley-VCH, Weinheim, 2005.
- [41] T. Krenz, O. Weiland, P. Trinke, L. Helmers, C. Eckert, B. Benschmann, R. Hanke-Rauschenbach, Temperature and performance inhomogeneities in PEM electrolysis stacks with industrial scale cells, *J. Electrochem. Soc.* 170 (4) (2023) 044508, <http://dx.doi.org/10.1149/1945-7111/accb68>.
- [42] P. Olivier, C. Bourasseau, P.B. Bouamama, Low-temperature electrolysis system modelling: A review, *Renew. Sustain. Energy Rev.* 78 (2017) 280–300, <http://dx.doi.org/10.1016/j.rser.2017.03.099>.
- [43] M. Rogler, M. Suermann, R. Wagner, S. Thiele, J. Straub, Advanced method for voltage breakdown analysis of PEM water electrolysis cells with low iridium loadings, *J. Electrochem. Soc.* 170 (11) (2023) 114521, <http://dx.doi.org/10.1149/1945-7111/ad0b74>.
- [44] A. Hartig-Weiß, A. Siebel, M. Bernt, T.-H. Shen, V. Tileli, H.A. Gasteiger, Impact of intermittent operation on lifetime and performance of a PEM water electrolyzer, *J. Electrochem. Soc.* 166 (8) (2019) F487–F497, <http://dx.doi.org/10.1149/2.0421908jes>.
- [45] E. Padgett, H. Yu, S.J. Blair, D.A. Cullen, R.K. Ahluwalia, D.J. Myers, B. Pivovar, S.M. Alia, Quantifying sources of voltage decay in long-term durability testing for PEM water electrolysis, *J. Electrochem. Soc.* 172 (5) (2025) 054508, <http://dx.doi.org/10.1149/1945-7111/add184>.
- [46] M. Suermann, B. Benschmann, R. Hanke-Rauschenbach, Degradation of proton exchange membrane (PEM) water electrolysis cells: Looking beyond the cell voltage increase, *J. Electrochem. Soc.* 166 (10) (2019) F645–F652, <http://dx.doi.org/10.1149/2.1451910jes>.
- [47] M. Milosevic, T. Böhm, A. Körner, M. Bierling, L. Winkelmann, K. Ehelebe, A. Hutzler, M. Suermann, S. Thiele, S. Cherevko, In search of lost iridium: Quantification of anode catalyst layer dissolution in proton exchange membrane water electrolyzers, *ACS Energy Lett.* 8 (6) (2023) 2682–2688, <http://dx.doi.org/10.1021/acseenergylett.3c00193>.
- [48] A. Luxa, N. Jöres, C. Yáñez, M. Souza, G. Pangalos, L. Schnelle, G. Lichtenberg, Multilinear modeling and simulation of a multi-stack PEM electrolyzer with degradation for control concept comparison, in: Proceedings of the 12th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, Scite Press, 2022, pp. 52–62, <http://dx.doi.org/10.5220/0011263300003274>.
- [49] S. Stucki, G.G. Scherer, S. Schlagowski, E. Fischer, PEM water electrolyzers: evidence for membrane failure in 100kW demonstration plants, *J. Appl. Electrochem.* 28 (10) (1998) 1041–1049, <http://dx.doi.org/10.1023/A:1003477305336>.
- [50] Technical targets for proton exchange membrane electrolysis, 2024, URL <https://www.energy.gov/eere/fuelcells/technical-targets-proton-exchange-membrane-electrolysis>. (Accessed 09 December 2024).
- [51] M. Suermann, T.J. Schmidt, F.N. Büchi, Cell performance determining parameters in high pressure water electrolysis, *Electrochim. Acta* 211 (2016) 989–997, <http://dx.doi.org/10.1016/j.electacta.2016.06.120>.
- [52] S. Siracusanò, S. Trocino, N. Briggoglio, V. Baglio, A.S. Aricò, Electrochemical impedance spectroscopy as a diagnostic tool in polymer electrolyte membrane electrolysis, *Mater. (Basel, Switzerland)* 11 (8) (2018) <http://dx.doi.org/10.3390/ma11081368>.
- [53] T. Malkow, A. Pilega, G. Tsotridis, EU harmonised test procedure: Electrochemical impedance spectroscopy for water electrolysis cells, in: EUR, vol. 29267, Publications Office of the European Union, Luxembourg, 2018.
- [54] M. Suermann, T.J. Schmidt, F.N. Büchi, Comparing the kinetic activation energy of the oxygen evolution and reduction reactions, *Electrochim. Acta* 281 (2018) 466–471, <http://dx.doi.org/10.1016/j.electacta.2018.05.150>.
- [55] S. Cherevko, S. Geiger, O. Kasian, A. Mingers, K.J. Mayrhofer, Oxygen evolution activity and stability of iridium in acidic media. Part 1. – metallic iridium, *J. Electroanal. Chem.* 773 (2016) 69–78, <http://dx.doi.org/10.1016/j.jelechem.2016.04.033>.

- [56] S. Cherevko, S. Geiger, O. Kasian, A. Mingers, K.J. Mayrhofer, Oxygen evolution activity and stability of iridium in acidic media. Part 2. – electrochemically grown hydrous iridium oxide, *J. Electroanal. Chem.* 774 (2016) 102–110, <http://dx.doi.org/10.1016/j.jelechem.2016.05.015>.
- [57] M. Rogler, R. Wagner, S. Thiele, M. Suermann, Guidance for targeted degradation analysis of OER kinetics of low-loading iridium-based catalysts in PEM water electrolysis cells, *Electrochim. Acta* 510 (2025) 145360, <http://dx.doi.org/10.1016/j.electacta.2024.145360>.
- [58] H. Lee, J. Gu, B. Lee, H.-S. Cho, H. Lim, Prognostics and health management of alkaline water electrolyzer: Techno-economic analysis considering replacement moment, *Energy AI* 13 (2023) 100251, <http://dx.doi.org/10.1016/j.egyai.2023.100251>.
- [59] H. Zhang, T. Yuan, Optimization and economic evaluation of a PEM electrolysis system considering its degradation in variable-power operations, *Appl. Energy* 324 (2022) 119760, <http://dx.doi.org/10.1016/j.apenergy.2022.119760>.
- [60] W. Zheng, B. Lv, Z. Shao, B. Zhang, Z. Liu, J. Sun, J. Yuan, C. Jiang, Optimization of power allocation for the multi-stack PEMEC system considering energy efficiency and degradation, *Int. J. Hydrog. Energy* 53 (2024) 1210–1225, <http://dx.doi.org/10.1016/j.ijhydene.2023.11.241>.