

Article

Physics-Anchored Dual Network for Conditional SOFC Health-Indicator Trajectory Forecasting Under Variable-Load Operation

Tao Zhang ¹, Zhixiang Liu ²  and Jingxiang Xu ^{1,*} 

¹ College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China; 26792logan@gmail.com

² College of Information Technology, Shanghai Ocean University, Shanghai 201306, China; zxliu@shou.edu.cn

* Correspondence: jxxu@shou.edu.cn

Abstract

Reliable health-indicator forecasting under variable-load operation is important for SOFC durability management, maintenance scheduling, and system control. This study investigates offline conditional forecasting using an equivalent ohmic-resistance health indicator. The indicator is extracted from experimental data through reduced-order model inversion and subsequently processed using causal smoothing. In the test horizon, the future operating-condition profile (t, U, T) is supplied as the conditioning input, whereas future health-indicator labels are hidden during inference and used only for post-test evaluation. The measured current $I(t)$ is used only for label extraction and is not supplied as a future covariate. We propose a Physics-Anchored Dual-Network (PADN) framework that couples a surrogate network and a dynamic network through an anchored dynamic relation built around a calibrated empirical degradation skeleton. The prediction target serves as a macroscopic health indicator rather than as a direct microscopic degradation measurement, so the task is not formulated as RUL prediction or unknown-load forecasting. PADN is evaluated on a public variable-load aging dataset containing two SOFC single cells of about 1700 h and is compared with EXP-only, Direct-MLP, CNN-LSTM, and Transformer baselines using chronological split points at 50%, 63%, and 78% of the recorded timeline; data before each split point are used for model development, including parameter updates and validation, whereas data after it are held out for testing. Across the six cell-split-point combinations, PADN achieves the lowest MAPE in five combinations and the lowest average errors at each split point among the compared methods. In the most data-limited scenario, with the chronological split point at 50% of the recorded timeline, PADN maintains MAPE values of 2.645% and 2.934% for Cell-1 and Cell-2, respectively. Ablation results show that the empirical degradation skeleton, physics-consistency loss, and residual regularization each contribute to improved extrapolation stability under the present two-cell conditional forecasting benchmark. Overall, PADN provides a practical gray-box approach for conditional forecasting of extracted and smoothed SOFC health-indicator trajectories under limited-data variable-load settings, although the current evidence remains limited to two cells from one experimental platform.



Academic Editors: Antonino S. Arico and Young-Wan Ju

Received: 13 May 2026

Revised: 6 July 2026

Accepted: 30 July 2026

Published: 3 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: solid oxide fuel cell; health-indicator forecasting; conditional trajectory forecasting; physics-anchored dual network; gray-box modeling

1. Introduction

Amid the transition to low-carbon energy systems, high-efficiency electrochemical conversion technologies have attracted growing attention. Solid oxide fuel cells (SOFCs) are particularly promising because they combine high energy-conversion efficiency with broad fuel flexibility and compatibility with combined heat and power systems [1,2]. Their ability to operate on hydrogen, natural gas, and syngas makes them attractive for distributed generation and auxiliary power applications. However, long-term SOFC operation at high temperatures and under harsh electrochemical conditions inevitably leads to microstructural degradation, including Ni coarsening, loss of triple-phase boundaries, secondary-phase formation, and chromium poisoning [3–6]. These processes increase ohmic and polarization losses, reduce output performance, and ultimately limit commercial durability. Accurate degradation forecasting is therefore important for health management, maintenance scheduling, and life-cycle cost reduction in SOFC systems [7,8].

Mechanism-based degradation modeling begins with electrochemical kinetics, transport, and material evolution. Khan et al. linked anode performance loss to Ni-YSZ coarsening through image-informed coarsening theory [9]. Mason et al. developed a multidimensional framework for analyzing performance decay caused by microstructural coarsening in SOFC electrodes [10]. Zaccaria et al. proposed a distributed real-time degradation model for SOFC operation under changing conditions [11]. These studies offer strong interpretability and can connect measurable degradation behavior to physical causes. However, multi-physics coupling, parameter identification, and cell-to-cell variation make such models difficult to calibrate and generalize under variable-load operation [12].

Data-driven methods bypass explicit full-order physics and instead learn degradation patterns from monitored signals [13]. Rao et al. used deep learning for SOFC state prediction under complex operating conditions [14]. Yang et al. applied machine learning to predict long-term cathode degradation induced by chromium poisoning [15]. Li et al. developed a deep-learning voltage prognostic model for SOFC systems [16]. More broadly, recurrent and attention-based architectures such as LSTM and Transformer are standard tools for nonlinear sequence modeling [17,18]. These approaches are flexible and effective for high-dimensional inputs, but they depend strongly on the quality of the training data [19]. Because SOFC durability experiments are expensive and time-consuming, available full-life datasets are often limited. Under small-sample and variable-load settings, purely data-driven models can overfit transient fluctuations and produce unstable or non-physical long-horizon extrapolations.

Hybrid modeling and physics-informed approaches seek to balance physical structure with data adaptability [20]. In SOFC research, Chi et al. introduced the link function degradation model, which explicitly represents the degradation rate as a function of operating conditions and enables online identification under variable-load operation [21]. Subsequent studies further incorporated degradation-aware prognostic models into SOFC operation optimization and control frameworks [22–24]. Shi et al. combined a neural network with a dynamic decay model to improve life prediction for SOFC combined heat and power systems [25].

More broadly, physics-informed neural networks (PINNs) typically exploit prior physical knowledge by incorporating differential equations or physical residuals into the learning process [26,27]. Studies by Wang et al., Ko et al., and Xiong et al. have shown that such methods can improve the stability of degradation prognosis in lithium-ion batteries and proton-exchange membrane fuel cells [28–30]. However, many PINN-style methods rely on a basic assumption that sufficiently complete governing equations are already available. This condition is often difficult to satisfy for SOFC degradation. The degradation process involves strongly coupled multiscale and multiphysics mechanisms, which are

difficult to describe fully using a single unified equation set. This does not mean that PINN-style methods are inherently unsuitable. Rather, it suggests that gray-box modeling is more appropriate for the SOFC degradation setting considered in this study, because only an empirical and partially interpretable degradation skeleton is currently available.

PADN is developed on the basis of this gray-box formulation. It does not require a complete governing equation for SOFC degradation to be specified in advance. Instead, it uses a calibrated empirical degradation skeleton as a physics anchor and employs a dual-network formulation to learn the residual dynamics not captured by the skeleton. During training, the anchored dynamic relation is constrained through a physics-consistency loss. During inference, the forecast is obtained through direct forward evaluation of the surrogate network, without explicit stepwise rollout or test-time enforcement of physical residuals. PADN is therefore better understood as a gray-box surrogate forecaster trained under physical anchoring, rather than as a standard PINN formulation built on a fully specified governing equation.

In this work, we study offline conditional forecasting of an experiment-extracted and smoothed equivalent ohmic-resistance health indicator under variable-load operation, limited training data, and long-horizon extrapolation. The future operating-condition sequence is assumed to be known from the recorded load profile, so the task is not unknown-load forecasting or direct RUL prediction. The future $R_0(t)$ labels are hidden during inference and are used only for post-test evaluation. The reduced-order inversion is used before model training to construct evaluation labels, not as a test-horizon forecasting model. For this setting, we propose a Physics-Anchored Dual-Network (PADN) framework, a gray-box forecasting formulation in which a calibrated empirical degradation skeleton provides the physical trend and neural networks learn the residual dynamics that the skeleton cannot represent. We evaluate PADN on about 1700 h of variable-load aging data from two SOFC single cells against EXP-only, Direct-MLP, CNN-LSTM, and Transformer baselines, and we examine the empirical anchor, physics-consistency loss, and residual regularization through ablation experiments.

The main contributions of this experimental-data study are threefold: First, we formulate PADN as a gray-box conditional forecasting framework that couples a surrogate network and a dynamic network through an anchored dynamic relation, allowing residual trajectory behavior to be learned without requiring a complete governing equation. Second, we construct a forecasting-oriented health-indicator pipeline for the real variable-load dataset by extracting an equivalent ohmic-resistance trajectory from the released $D_1/D_2/D_3$ records and validating it against periodic I–V references before model training. Third, we provide a controlled two-cell benchmark on the same public single-cell dataset through comparative and ablation experiments using the same chronological split points at 50%, 63%, and 78% of the recorded timeline across all methods, which allows PADN and the baselines to be examined under a fixed label definition and task setting.

2. Overall Framework of PADN

The proposed Physics-Anchored Dual-Network (PADN) framework is designed for conditional forecasting of SOFC health-indicator trajectories under variable-load operation, limited training data, and long-horizon extrapolation. Rather than learning a complete degradation law from data alone, PADN combines a calibrated empirical degradation skeleton with two neural components. The empirical degradation skeleton provides a physically meaningful trend direction, while the neural networks compensate for nonlinear residual dynamics that the empirical form cannot capture. In this study, the predicted quantity is the extracted and smoothed equivalent ohmic-resistance trajectory, which is

used as a macroscopic health indicator rather than as a direct measurement of a single degradation mechanism; the extraction procedure is described in Section 3.

2.1. Problem Formulation

Let $\mathbf{x}(t) = [U(t), T(t)]^\top$ denote the operating-condition vector at time t , containing the recorded cell-voltage channel and temperature used by the empirical degradation skeleton. In the present offline benchmark, these channels are treated as part of the recorded conditioning profile; they are not autonomously predicted by PADN. Let $R_0(t)$ denote the extracted and smoothed equivalent ohmic-resistance health indicator to be predicted. Given a set of observations from an early-life interval, defined by (1),

$$\mathcal{D} = \{(t_i, \mathbf{x}_i, R_{0,i})\}_{i=1}^N, \quad (1)$$

this study aims to develop a conditional forecasting model that fits the observed health-indicator trajectory and predicts its evolution over a future horizon. The operating-condition sequence $\mathbf{x}(t)$ over this horizon is obtained from the recorded load profile. The future R_0 trajectory is not uniquely determined by its historical trajectory alone, because its evolution also depends on the subsequent operating conditions. The same health state at the forecast origin may lead to different future R_0 trajectories under different voltage and temperature profiles. Forecasting a specific R_0 trajectory therefore requires the corresponding future operating scenario to be specified. Forecasting under unknown future operating conditions constitutes a different task. Under this specified conditional setting, only (t, U, T) are used as model inputs during testing. During inference, PADN does not use future $R_0(t)$ values, pointwise inverted $R_0^{\text{raw}}(t)$ values, or measured current $I(t)$. The task is therefore offline conditional trajectory forecasting rather than autonomous operating-profile prediction or direct RUL prediction. During inference, no future R_0 labels or test-horizon parameter updates are used. Unlike standard black-box sequence prediction, PADN formulates the problem as continuous-state learning with an explicit dynamic constraint. The objective is not only low in-domain fitting error but also stable and physically plausible out-of-domain evolution under the specified conditional setting. Accordingly, this study does not claim autonomous prediction of unknown future operating conditions, direct remaining-useful-life estimation, or direct identification of the true microscopic degradation state. The setting corresponds to post-test conditional forecasting or planned-profile prognosis; deployment under unknown future load would require an additional operating-profile prediction module. A consolidated summary of the principal symbols used throughout the manuscript is provided in Appendix A.

2.2. Empirical Degradation Skeleton

Based on the link-function degradation formulation proposed by Chi et al. [21], the empirical degradation skeleton explicitly relates operating conditions to the degradation rate. For the experimental dataset considered here, the reproduced exponential link-function form is given in (2):

$$g_{\text{emp}}(\mathbf{x}(t); \hat{\lambda}) = \hat{\lambda}_1 (1.2 - U(t)) \exp\left(-\frac{\hat{\lambda}_2}{T(t)}\right) \quad (2)$$

where $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2)$ is the set of identified parameters. The factor $(1.2 - U)$ approximates the effect of operating overpotential in the present experiment, where the open-circuit voltage is about 1.2 V, and the exponential term introduces an Arrhenius-like temperature dependence. The variables U and T in (2) are evaluated in their physical units, namely, V and K. Network inputs may be normalized for optimization, but the empirical skeleton

itself is evaluated using the corresponding physical-unit values, or equivalently using values transformed back from the normalized representation before (2) is applied. Only U and T are included in the link function so that the empirical degradation skeleton remains an explicit operating-condition-to-degradation-rate relation rather than a response-dependent mapping. This structure preserves interpretability and load sensitivity, but its fixed functional form is usually insufficient to represent cell-specific and nonlinear residual dynamics. PADN, therefore, does not replace the empirical degradation skeleton; instead, it learns only the unexplained component.

The exponential variant is retained here to define the empirical degradation skeleton for the experimental dataset because it provides an interpretable operating-condition-dependent anchor that captures the dominant degradation direction while remaining consistent with the temperature-sensitive trend in the data, even though it cannot absorb all nonlinear residual effects.

2.3. Physics-Anchored Dual-Network Architecture

PADN contains a surrogate network and a dynamic network. Both are implemented as multilayer perceptrons so that the contribution of physical anchoring can be evaluated independently of more elaborate sequence-model backbones. The surrogate network maps time and operating conditions to the health indicator as defined in (3):

$$\hat{R}_0(t) = f_s(t, \mathbf{x}(t); \theta_s) \quad (3)$$

where θ_s denotes the parameters of the surrogate network. Its role is to generate the full health-indicator trajectory.

The dynamic network predicts only the residual degradation rate that is not captured by the empirical degradation skeleton, as defined in (4):

$$\delta_d(t) = f_d(t, \mathbf{x}(t), \hat{R}_0(t); \theta_d) \quad (4)$$

where θ_d denotes the parameters of the dynamic network. Including $\hat{R}_0(t)$ in the input allows the residual term to depend on the current degradation stage rather than on operating conditions alone. This design reflects the fact that the same load condition can induce different residual dynamics at different health states.

The two networks are coupled through a single anchored dynamic relation given by (5):

$$\frac{d\hat{R}_0(t)}{dt} = g_{\text{emp}}(\mathbf{x}(t); \hat{\lambda}) + \delta_d(t). \quad (5)$$

The derivative on the left-hand side of (5) is taken with respect to physical time. In implementation, automatic differentiation is taken with respect to the normalized time coordinate and then converted back to the physical scale before the physics residual is evaluated. If $t' = (t - \mu_t)/\sigma_t$ and $R' = (R - \mu_R)/\sigma_R$, the conversion is

$$\frac{dR}{dt} = \frac{\sigma_R}{\sigma_t} \frac{dR'}{dt'}. \quad (6)$$

The first term on the right-hand side represents the physical trend provided by the empirical degradation skeleton, and the second term is the learned residual correction. In PADN, this anchored dynamic relation is enforced through the physics-consistency loss during training rather than solved through explicit recurrent rollout during inference. The formulation therefore turns unconstrained black-box trajectory fitting into residual learning around a physically meaningful anchor. The overall architecture of PADN is shown in Figure 1.

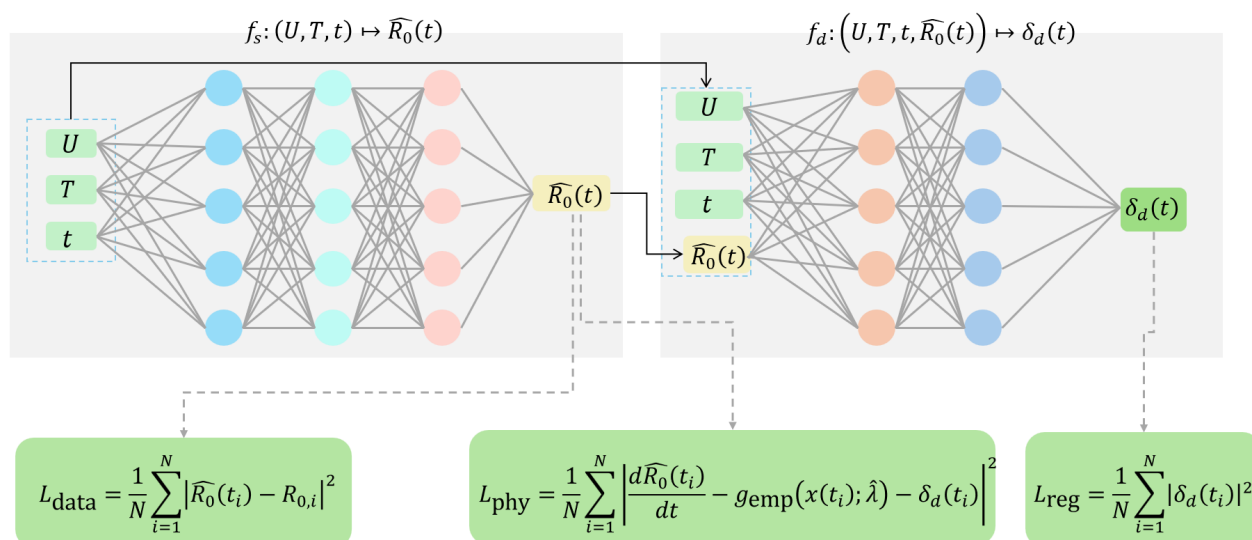


Figure 1. Schematic diagram of the Physics-Anchored Dual-Network (PADN) framework, showing the surrogate branch, dynamic branch, empirical degradation skeleton, and training losses.

2.4. Physics Constraints and Composite Loss

Training is driven by a composite loss with three components, defined in (7):

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \alpha \mathcal{L}_{\text{phy}} + \beta \mathcal{L}_{\text{reg}}, \quad (7)$$

where α and β are weighting coefficients. The data-fitting term in (8) enforces agreement with observed health-indicator values:

$$\mathcal{L}_{\text{data}} = \frac{1}{N} \sum_{i=1}^N |\hat{R}_0(t_i) - R_{0,i}|^2. \quad (8)$$

The physics-consistency loss in (9) penalizes disagreement between the surrogate trajectory and the anchored dynamic relation:

$$\mathcal{L}_{\text{phy}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{d\hat{R}_0(t_i)}{dt} - g_{\text{emp}}(\mathbf{x}(t_i); \hat{\lambda}) - \delta_d(t_i) \right|^2. \quad (9)$$

The regularization term in (10) limits the magnitude of the residual dynamics:

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^N |\delta_d(t_i)|^2 \quad (10)$$

These three terms address different failure modes. $\mathcal{L}_{\text{data}}$ anchors the surrogate trajectory to the observed health-indicator points. $\mathcal{L}_{\text{phy}}$ keeps the learned evolution aligned with the empirical degradation skeleton through the anchored dynamic relation. $\mathcal{L}_{\text{reg}}$ bounds the residual correction so that the dynamic network does not overwhelm the baseline physical trend. Because the data-fitting and rate-based loss terms have different numerical scales, the reported α and β values should be interpreted as validation-selected scale-balancing hyperparameters for the normalized implementation rather than as universal physical constants. Figure 2 summarizes the composite objective $\mathcal{L} = \mathcal{L}_{\text{data}} + \alpha \mathcal{L}_{\text{phy}} + \beta \mathcal{L}_{\text{reg}}$ and highlights how these three terms jointly promote accurate fitting, physically plausible evolution, and stable extrapolation.

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \alpha \mathcal{L}_{\text{phy}} + \beta \mathcal{L}_{\text{reg}}$$

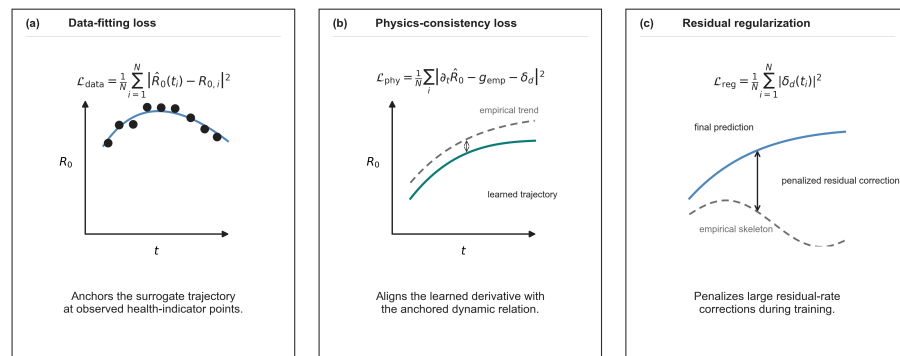


Figure 2. Conceptual decomposition of the PADN composite loss. The diagram shows how the data-fitting, physics-consistency, and residual-regularization terms jointly support accurate fitting, physics-consistent evolution, and stable extrapolation.

2.5. Training and Forecasting Procedure

PADN uses a two-stage training strategy. For each cell and normalized chronological split point ρ , the development segment before the split point is divided into a parameter-update subset and a validation subset, as described in Section 3.2. In Stage 1, the empirical degradation skeleton is fitted only on the parameter-update subset, and its identified parameters $\hat{\lambda}$ are then fixed. In Stage 2, the surrogate network and the dynamic network are jointly optimized under the composite loss using the same parameter-update subset, while the validation subset is used for early stopping and hyperparameter selection. The held-out test segment is not used for fitting the empirical skeleton, selecting hyperparameters, or updating neural-network parameters. This separation is important because it prevents drift of the empirical degradation skeleton during neural-network training and preserves a clear division between the interpretable component and the data-driven residual correction.

During inference, the health indicator is predicted by a forward pass through the surrogate network using the held-out (t, U, T) profile. The true held-out $R_0(t)$ trajectory is withheld until evaluation metrics and plots are computed after prediction. The dynamic network and the empirical degradation skeleton are not iteratively rolled out at each test step because their information has already been embedded in the surrogate trajectory through the physics-consistency loss during training. PADN should therefore be interpreted at test time as a surrogate forecaster trained under the anchored dynamic relation, not as an explicit stepwise ODE solver. If residual-rate decomposition is required for analysis, the dynamic network can still be evaluated separately.

The extrapolation behavior of PADN is influenced by the residual regularization. When the model moves away from regions well supported by training data, the composite objective penalizes large residual corrections and, therefore, biases the learned dynamics toward smaller $|\delta_d(t)|$. In the limiting case where the residual correction becomes small, the learned dynamics approach the empirical degradation skeleton according to (11):

$$\delta_d(t) \rightarrow 0 \quad \Rightarrow \quad \frac{d\hat{R}_0(t)}{dt} \approx g_{\text{emp}}(\mathbf{x}(t); \hat{\lambda}). \quad (11)$$

This limiting-case relation in (11) should be interpreted as an inductive bias rather than as a guaranteed automatic fallback at every unseen point. In practice, PADN is encouraged to remain closer to the empirical trend as extrapolation uncertainty grows, which is consistent with the observed extrapolation behavior. This physics-anchoring mechanism constrains the learned residual correction without forcing PADN to become a purely

empirical model: the surrogate network can still learn cell-specific nonlinear trajectory behavior, while the composite loss discourages extrapolations that require excessively large residual rates.

3. Experimental Data and Evaluation Setup

3.1. Variable-Load Aging Dataset

The study uses the publicly available variable-load SOFC aging dataset reported by Chi et al. [21]. The experimental objects are two planar single cells, denoted Cell-1 and Cell-2. Hydrogen and air flow rates were maintained at 40 sccm and 300 sccm, respectively, while voltage, current, and temperature were continuously recorded. The experiment included an initial stabilization stage near 903 K, an initial I–V characterization stage at multiple temperatures, and a subsequent variable-load aging stage of about 1700 h. Across 14 operating segments, the temperature varied from 873 K to 923 K, and the voltage setpoint varied from 0.5 V to 0.9 V. Following the original data organization, we distinguished three subsets: an initial I–V dataset D_1 , a periodic I–V dataset D_2 , and a continuous variable-load operation dataset D_3 . The first subset was used to identify the fixed cell-model parameters, the second was used only for a separately refitted cross-check of the extracted health indicator, and the third provided the chronological degradation sequence used for model development and testing. The detailed experimental control ranges are summarized in Table 1. These repeated switches in operating conditions make the resulting trajectories strongly non-stationary and, therefore, suitable for evaluating early-life forecasting under realistic load changes. The operating profiles of the two cells are shown in Figure 3.

Table 1. Experimental control ranges for the variable-load aging dataset.

Segment	Cell-1			Cell-2		
	Duration	Temperature	Voltage	Duration	Temperature	Voltage
1	0–104 h	903 K	0.9 V	0–105 h	903 K	0.9 V
2	104–188 h	873 K	0.9 V	105–189 h	873 K	0.9 V
3	188–282 h	873 K	0.7 V	189–283 h	873 K	0.7 V
4	282–378 h	873 K	0.5 V	283–380 h	873 K	0.5 V
5	378–471 h	923 K	0.9 V	380–470 h	923 K	0.9 V
6	471–573 h	923 K	0.7 V	470–564 h	923 K	0.7 V
7	573–594 h	923 K	0.5 V	564–658 h	923 K	0.6 V
8	594–692 h	923 K	0.6 V	658–797 h	903 K	0.7 V
9	692–861 h	903 K	0.7 V	797–904 h	903 K	0.5 V
10	861–1025 h	903 K	0.5 V	904–998 h	903 K	0.5 V
11	1025–1119 h	903 K	0.5 V	998–1204 h	903 K	0.7 V
12	1119–1326 h	903 K	0.7 V	1204–1417 h	923 K	0.7 V
13	1326–1539 h	923 K	0.7 V	1417–1631 h	923 K	0.5 V
14	1539–1753 h	923 K	0.5 V	1631–1669 h	923 K	0.9 V

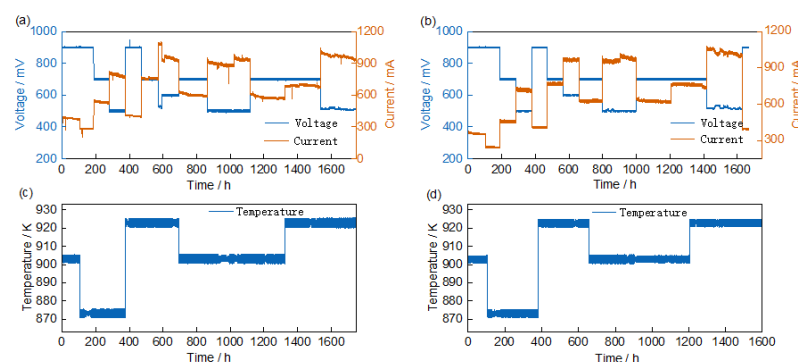


Figure 3. Variable-load operating profiles plotted against operating time for the two SOFC single cells used for evaluation. (a,b) Voltage and current versus time for Cell-1 and Cell-2, respectively; (c,d) the corresponding temperature profiles versus time.

3.2. Data Preprocessing

The analysis started from the released continuous-operation sequence in D_3 , from which abnormal-value samples had already been removed in the public data source, and to which no smoothing had been applied. We did not perform additional response-derived outlier removal, interpolation, or gap filling before constructing the forecasting labels. The retained D_3 records were treated as the chronological measurement grid, with a nominal sampling interval of about 0.1 h. Because degradation prediction is a causal task, all data splits preserve temporal order. Let ρ denote the normalized chronological split point, expressed as a fraction of each recorded D_3 timeline. The data before each split point form the development segment used for training and validation, whereas the data after it form the held-out test segment. The health-indicator labels in the test segment remain fully unseen by the forecasting models. Only the time grid and the voltage and temperature sequences defining the future operating scenario are retained as conditioning inputs, whereas the future measured current, raw inverted resistance, and smoothed R_0 values are withheld during inference. Inversion and smoothing on the held-out segment are used only to define reference labels for evaluation after predictions have been generated. Within the development segment, the last 10% is held out as a validation subset for early stopping and hyperparameter selection, and the preceding 90% is used for parameter updates. No random shuffling is applied across time. We use three chronological split points along the recorded timeline, $\rho \in \{50\%, 63\%, 78\%\}$, to assess sensitivity to the split location. The task considered here remains offline conditional trajectory forecasting: the future operating-condition sequence is known, whereas future health-indicator labels are not.

3.3. Health-Indicator Definition and Extraction

Under variable-load operation, terminal voltage and current fluctuate rapidly in response to reversible changes in load and temperature, making them unsuitable for direct use as health indicators. In this study, the equivalent ohmic-resistance trajectory $R_0(t)$ is used to characterize cell health. It is derived from experimental data through reduced-order model inversion and causal smoothing. Here, $R_{\text{ohm}}(T, t)$ refers to the temperature-dependent ohmic term in the reduced-order model, $R_0^{\text{raw}}(t)$ is the raw equivalent ohmic-resistance sequence obtained by pointwise inversion, and $R_0(t)$ is the smoothed trajectory used as the forecasting target. This choice offers two advantages: First, the health indicator partially reduces the direct influence of reversible operating-condition variations and serves as a more degradation-oriented indicator than raw voltage or current. Second, it can be continuously inferred from routine operating signals and cross-checked against separately refitted periodic I–V references. In this study, $R_0(t)$ is treated as a macroscopic equivalent representation of multiple degradation mechanisms rather than as a one-to-one measurement of a single microscopic process or as a direct observation of the true degradation state.

Following the reduced-order SOFC cell model adopted by Chi et al. [21], the voltage–current relation used for health-indicator extraction is written as (12):

$$U = E_{\text{OCV}}(T) - IR_{\text{ohm}}(T, t) - c_{\text{act}}(T) \ln \left(\frac{I + \sqrt{I^2 + 4i_0(T)^2}}{2i_0(T)} \right), \quad (12)$$

where the ohmic-resistance term follows Chi et al. [21] and is written as (13):

$$R_{\text{ohm}}(T, t) = R_0^{\text{raw}}(t) \exp \left(\frac{b_{\text{ohm}}}{T} \right), \quad (13)$$

This formulation separates the ohmic contribution into a temperature-independent raw equivalent ohmic-resistance sequence, $R_0^{\text{raw}}(t)$, and an Arrhenius-type temperature factor. Equivalently, after calibration of the fixed reduced-order model from D_1 , the raw equivalent ohmic-resistance value at each processed operating point is obtained by (14):

$$R_0^{\text{raw}}(t_k) = \frac{E_{\text{OCV}}(T(t_k)) - U(t_k) - c_{\text{act}}(T(t_k)) \ln \left(\frac{I(t_k) + \sqrt{I(t_k)^2 + 4i_0(T(t_k))^2}}{2i_0(T(t_k))} \right)}{I(t_k) \exp \left(\frac{b_{\text{ohm}}}{T(t_k)} \right)}. \quad (14)$$

The fixed reduced-order model parameters are first identified from D_1 . After that calibration step, each sample in the continuous-operation dataset D_3 is inverted pointwise from the measured triplet $(U(t_k), I(t_k), T(t_k))$ to obtain the discrete raw health-indicator sequence $R_0^{\text{raw}}(t_k)$. The current $I(t)$ is used only in this label-extraction step and is not used as a future covariate by PADN or by the neural baselines because it is response-dependent under the operating profile. Therefore, the voltage and temperature sequences supplied to the forecasting model are insufficient to reconstruct the future R_0 labels, because the future measured current required by (14) is unavailable during inference. Because it remains adequate over the observed operating range, the reduced-order model identified from D_1 is kept fixed throughout the approximately 1700 h horizon when processing the continuous-operation sequence. Because the released D_3 sequence already excludes abnormal-value samples, the inversion is applied only to valid operating points within the calibrated experimental range, without additional response-derived filtering or mid-course recalibration. The full extraction workflow is summarized in Figure 4.

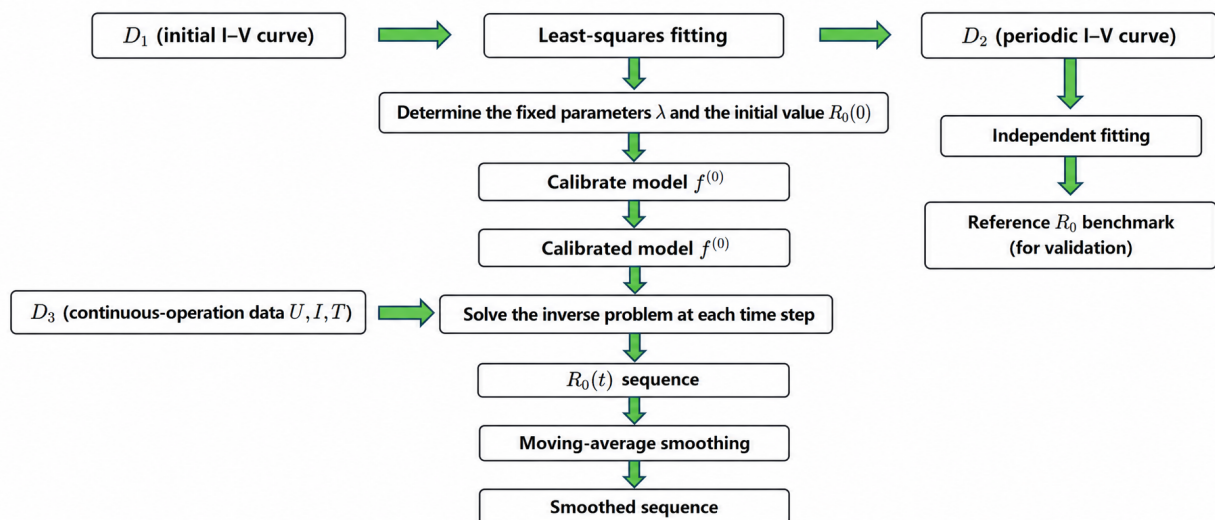


Figure 4. Flowchart of the equivalent ohmic-resistance health-indicator extraction workflow from the experimental records. The procedure maps initial I–V data, periodic I–V references, and continuous-operation measurements to the raw and smoothed $R_0(t)$ sequences used for validation and forecasting.

The extracted $R_0^{\text{raw}}(t)$ trajectory is cross-checked against the periodic I–V dataset D_2 . For this validation step only, the cell model is separately refitted at each I–V test time to obtain a reference ohmic-resistance value, and this reference trajectory is compared with the continuously inverted trajectory, as shown in Figure 5. Chi et al. [21] used the same comparison logic and noted that estimating $R_0(t_k)$ from D_3 uses only the sample at t_k , which makes the pointwise estimate sensitive to noise. Consistently, the continuously inverted trajectory follows the periodic I–V-based reference at the trend level but contains stronger short-term fluctuations.

Therefore, the fluctuations and occasional spikes in $R_0^{\text{raw}}(t)$ should not be interpreted as individual microscopic degradation events. Because this sequence is obtained by point-wise inversion of the measured U , I , and T , the process can amplify measurement noise, transients caused by load switching, thermal relaxation effects, and mismatch in the reduced-order model. The D_2 -based refits are used only to construct a reference trajectory and do not update the fixed reduced-order model used for continuous inversion on D_3 , nor do they convert $R_0(t)$ into a direct ground-truth measurement of a single physical mechanism. At the same time, this cross-check does not fully remove prior coupling because the target remains obtained by reduced-order-model inversion, while the PADN anchor is also model-informed. The present study therefore interprets the task as forecasting a cross-checked, reduced-order-model-inverted and smoothed health indicator, rather than as direct prediction of an independently observed degradation ground truth. To reduce short-term fluctuations before downstream forecasting, the inferred sequence is smoothed by a trailing causal moving average defined in (15):

$$R_0(t_k) = \frac{1}{\min(k+1, N_{\text{MA}})} \sum_{j=0}^{\min(k, N_{\text{MA}}-1)} R_0^{\text{raw}}(t_{k-j}). \quad (15)$$

Consistent with the moving-average noise-filtering role and the $N_{\text{MA}} = 450$ setting used by Chi et al. [21], we use $N_{\text{MA}} = 450$, which corresponds to a smoothing window of about 45 h under the nominal 0.1 h sampling interval. The moving-average definition in (15) is applied here in a strictly causal form with a shorter available history at the beginning of each sequence, so each $R_0(t_k)$ depends only on the current and past raw values, and no future samples beyond time t_k are used in label construction. This causal construction is important for the chronological split protocol because it prevents leakage of future information from the test horizon into the supervised target. The smoothed trajectory $R_0(t)$ is used as the target label for PADN and the neural baselines. This choice should be understood as part of the benchmark definition in the present study rather than as a universally optimal level of smoothing. Because no smoothing-sensitivity analysis is included here, the conclusions of this study are limited to the extracted-and-smoothed target defined above. The validation curves and the effect of smoothing are displayed in Figures 5 and 6, respectively.

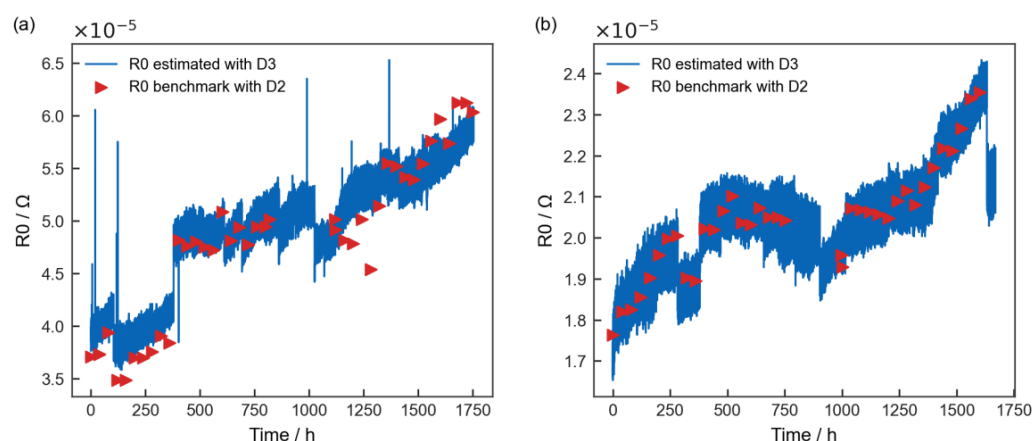


Figure 5. Validation of the extracted equivalent ohmic-resistance health indicator $R_0(t)$ plotted against operating time. Panels (a,b) correspond to Cell-1 and Cell-2, respectively. The blue curves show the D_3 -extracted $R_0(t)$ resistance trajectories, and the red markers show periodic I–V-based D_2 reference resistance values.

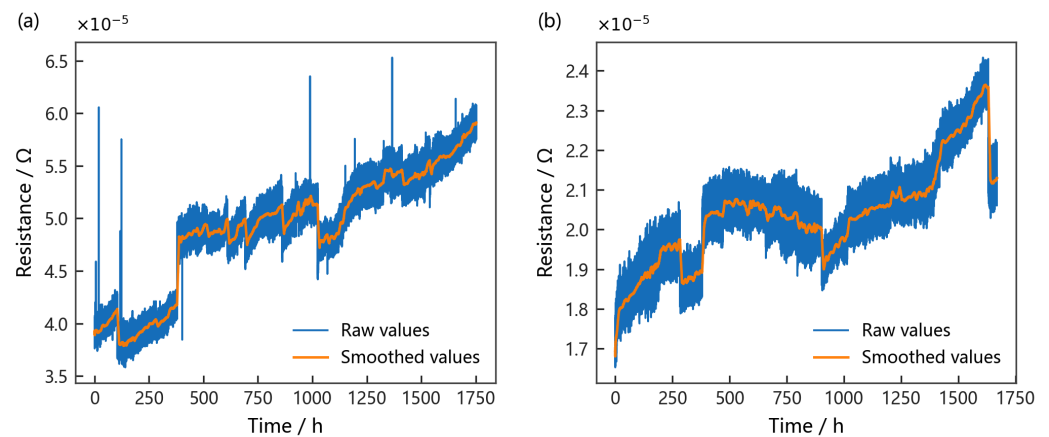


Figure 6. Moving-average smoothing of the extracted equivalent ohmic-resistance health indicator $R_0(t)$ plotted against operating time. Panels (a,b) correspond to Cell-1 and Cell-2, respectively. The blue curves show the raw resistance values, and the orange curves show the smoothed resistance trajectories used as forecasting targets.

3.4. Baselines and Evaluation Metrics

Five forecasting methods are evaluated using the same chronological split points and extracted health-indicator labels: PADN, the proposed gray-box model; EXP-only, a standalone empirical-anchor ODE baseline; Direct-MLP, a data-only direct surrogate baseline; CNN-LSTM, which represents convolution-recurrent sequence modeling; and Transformer, which represents attention-based long-sequence forecasting [18]. EXP-only fits only the empirical-link parameters $\hat{\lambda}_1$ and $\hat{\lambda}_2$ on the development segment, initializes the test prediction from the last observed development-segment R_0 value, and then integrates the empirical degradation relation in (2) over the held-out horizon using the known future (U, T, t) profile. It therefore tests whether the calibrated empirical skeleton alone is sufficient without a neural residual correction. Direct-MLP uses the same test-horizon inputs as the PADN surrogate, namely, (t, U, T) , and directly maps them to $\hat{R}_0(t)$ without the empirical degradation skeleton, physics-consistency loss, dynamic branch, or residual regularization. It therefore provides a mechanism-closer data-only baseline for separating direct surrogate inference from physics-anchored training. The CNN-LSTM and Transformer baselines are intentionally configured with parameter counts no lower than that of PADN, so that the comparison is not biased in favor of PADN by neural capacity alone. This design controls only one aspect of model capacity; architecture, input handling, and rollout strategy remain possible confounding factors. The *Without EXP* ablation still complements the standalone EXP-only baseline by removing the empirical skeleton from the anchored dynamic relation inside PADN. All methods are trained or fitted and evaluated on the same extracted and smoothed $R_0(t)$ target sequence and at the same chronological split points. Because the available experiment spans about 1700 h and does not reach a clearly defined end-of-life threshold for either cell, we do not formulate the task as remaining useful life (RUL) prediction. Instead, we evaluate how accurately each method reconstructs and extrapolates the extracted and smoothed health-indicator trajectory itself. The benchmark equalizes data source, label, and chronological split protocol, but the forecast-generation mechanisms are only partly identical across methods: PADN and Direct-MLP perform direct surrogate evaluation, and EXP-only performs empirical ODE integration, whereas CNN-LSTM and Transformer use a rolling one-step sequence-forecasting protocol. The resulting comparison should therefore be interpreted as controlled benchmark evidence under a shared data source, label definition, chronological split protocol, and added empirical-anchor and direct-surrogate baselines, rather than as a complete causal

attribution study. Prediction accuracy is quantified by root-mean-square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), defined in (16), (17), and (18), respectively:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{R}_{0,k} - R_{0,k})^2} \quad (16)$$

$$\text{MAE} = \frac{1}{N} \sum_{k=1}^N |\hat{R}_{0,k} - R_{0,k}| \quad (17)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{k=1}^N \left| \frac{\hat{R}_{0,k} - R_{0,k}}{R_{0,k}} \right|. \quad (18)$$

These three metrics jointly reflect sensitivity to large errors, absolute deviation, and relative percentage error. All reported RMSE, MAE, and MAPE values are computed only over the held-out test segment after the chronological split point.

3.5. Experimental Protocol

The evaluation comprises two parts: benchmark comparisons and ablation studies. In the comparative evaluation, PADN is compared with EXP-only, Direct-MLP, CNN-LSTM, and Transformer using chronological split points at 50%, 63%, and 78% of the recorded timeline. EXP-only estimates the two parameters of the empirical link function from the development data alone and uses the last R_0 value observed in the development segment as the initial value for the test horizon. Direct-MLP uses the same development, validation, and test splits as PADN. Its normalization statistics are computed from the development segment, and its architecture is configured so that the number of trainable parameters is comparable to that of the corresponding PADN configuration. For the rolling data-driven baselines, each training sample is constructed from a rolling history window of length 24 on the normalized conditional feature sequence, and the next smoothed R_0 value is used as the one-step target. During held-out evaluation, the recorded future (t, U, T) profile is supplied as the known conditional input, but future R_0 labels are not supplied. The measured current channel is also withheld from the forecasting inputs because it is used only in the reduced-order inversion that constructs the reference labels. The ablation study removes one component at a time. The *Without phy* variant sets $\alpha = 0$, thereby removing the physics-consistency loss; the dynamic branch is retained, but without $\mathcal{L}_{\text{phy}}$ it no longer constrains the surrogate derivative and is mainly affected by residual regularization. The *Without EXP* variant sets $g_{\text{emp}}(\mathbf{x}(t); \hat{\lambda}) \equiv 0$ in the anchored dynamic relation (5), thereby removing the empirical degradation skeleton. The *Without reg* variant sets $\beta = 0$, thereby removing the residual-magnitude regularization term.

For PADN and Direct-MLP, input and target variables are normalized using statistics from the development segment. The empirical skeleton in PADN and EXP-only is evaluated using physical-unit U and T , and the derivative term is converted according to (6) before forming the physics residual in PADN. Training is performed with large temporally ordered batches to preserve chronological structure during optimization. In the present experimental setting, the PADN surrogate network and Direct-MLP both take (t, U, T) as inputs, while the PADN dynamic network additionally takes $(t, U, T, \hat{R}_0)$ as inputs. PADN is implemented with multilayer perceptrons, but the joint network-structure candidate, width, depth, activation, loss weights, and regularization settings are selected separately by validation for each cell-split-point combination rather than being fixed globally. The predefined PADN hyperparameter search space is summarized in Table 2. For each cell and chronological split point, candidate configurations were trained only on the parameter-

update subset within the development segment, and the final configuration was selected using the internal validation subset, with validation MAPE used as the primary selection criterion and RMSE/MAE monitored as supporting accuracy measures. The held-out test segment was not used to select the network-structure candidate, network width, depth, activation function, loss weights, learning rate, stopping epoch, or random seed. Direct-MLP uses validation-selected multilayer perceptron settings matched to the corresponding PADN cell-split-point combination in input, target, split, normalization, training schedule, and approximate parameter count, while excluding the physics-anchored terms. To account for stochastic variation, each candidate configuration was trained with multiple random seeds and evaluated on the validation set. Neither the held-out test labels nor the test metrics were used to select a seed. The random seed changes the parameter initialization and the final learned weights, but not the number of trainable parameters, which is fixed by the model architecture. PADN and Direct-MLP were trained using AdamW with a nominal batch size of 8192, whereas the CNN-LSTM and Transformer baselines were trained using Adam with a nominal batch size of 16,384. EXP-only has no neural optimizer and uses only two fitted empirical-link parameters. When the number of available training samples is smaller than the specified batch size, the actual batch contains all available samples. Therefore, for some split-point/model combinations, the optimization is effectively full-batch or near-full-batch. The data-driven sequence baselines use deliberately larger parameter counts than most PADN settings: CNN-LSTM has 335,489 trainable parameters, and Transformer has 282,369 trainable parameters. The PADN configurations and neural-model parameter settings are summarized in Tables 3 and 4, respectively; EXP-only is described in the protocol text and reported in the comparative performance tables as a non-neural two-parameter baseline.

Table 2. Predefined PADN hyperparameter search space used for validation-based configuration selection.

Search Item	Candidate Values or Rule
Joint network-structure candidate	wider_smooth, trend_skip, or highcap_full; each candidate defines a different surrogate/dynamic network scale and skip-connection setting
Surrogate-network depth	3 or 4 hidden layers
Surrogate-network width	64 or 128 hidden units per layer
Dynamic-network depth	2 or 4 hidden layers
Dynamic-network width	32, 64, or 256 hidden units per layer
Activation function	Tanh or SiLU
Linear skip in surrogate	Disabled or enabled
Physics-consistency weight α	0.10–1.50, including the reported values 0.10, 0.12, 0.20, and 1.50
Residual-regularization weight β	0.0005–3.0, including the reported values 0.0005, 0.001, 0.5, 1.0, and 3.0
Empirical-skeleton scale	1.0 for the Full PADN configurations in the comparative experiment
Learning rate	5×10^{-4} – 8×10^{-4}
Batch size	8192 for PADN and Direct-MLP
Maximum epochs/early-stopping patience	140–220 epochs/18–35 validation-patience epochs
Gradient clipping	5.0
Network scale and trainable parameter count	Network scale participates in validation-based selection through the joint structure, depth, and width candidates; the reported trainable parameter count is then derived from the selected architecture and covers 9894–249,094 parameters
Selection data and criterion	Development-segment internal validation subset only; primary criterion was validation MAPE, with validation RMSE and MAE monitored as supporting metrics

Table 3. PADN configurations used in experiments at the three chronological split points.

Split Point (% of Timeline)	Cell	Candidate	Parameters	Activation	α	β	g_{emp} Scale
50%	Cell-1	wider_smooth	54,722	Tanh	0.20	0.001	1.0
50%	Cell-2	trend_skip	9894	SiLU	0.12	0.0005	1.0
63%	Cell-1	highcap_full	249,094	SiLU	0.20	1.0	1.0
63%	Cell-2	highcap_full	249,094	SiLU	0.10	1.0	1.0
78%	Cell-1	highcap_full	249,094	SiLU	0.20	0.5	1.0
78%	Cell-2	highcap_full	249,094	SiLU	1.50	3.0	1.0

Table 4. Neural model parameter settings used in the comparative experiments.

Category	PADN (Proposed)	Direct-MLP	CNN-LSTM	Transformer
Architecture	Validation-selected MLP surrogate and dynamic networks with an empirical degradation skeleton, physics-consistency loss, and residual regularization	Data-only MLP surrogate mapping (t, U, T) directly to $\hat{R}_0(t)$; no empirical skeleton, physics loss, dynamic branch, or residual regularization	1D convolutional front end with 128 channels and one 192-unit LSTM layer	Two Transformer encoder layers with 128-dimensional embeddings, eight attention heads, and a 256-dimensional feed-forward block
Optimizer	AdamW	AdamW	Adam	Adam
Learning rate	5×10^{-4} – 8×10^{-4}	5×10^{-4} – 8×10^{-4}	1×10^{-3}	1×10^{-3}
Batch size	8192	8192	16,384	16,384
Epochs/ patience	140–220/18–35	140–220/18–35	200/35	200/35
Gradient clipping	5.0	5.0	5.0	5.0
Trainable parameters	9894–249,094	9729–249,408	335,489	282,369

4. Results and Discussion

4.1. Comparative Results

4.1.1. Average Forecasting Advantage with Local Exceptions

Tables 5–7 compare PADN with EXP-only, Direct-MLP, CNN-LSTM, and Transformer using the same chronological split points at 50%, 63%, and 78% of the recorded timeline. Across the six cell-split-point combinations, PADN achieves the lowest MAPE in five combinations. The exception is Cell-2 with the split point at 63%, where the two-parameter EXP-only baseline follows the local held-out trend particularly well. EXP-only also gives slightly lower RMSE and MAE for Cell-2 with the split point at 50%, although PADN remains lower in MAPE for that combination. The PADN MAPE values range from 1.873% to 3.438%, whereas the baseline MAPE values range from 2.076% to 10.520%. Averaged across the two cells, PADN has the lowest RMSE, MAE, and MAPE at every split point. Relative to the strongest non-PADN baseline at each split point, PADN reduces average MAPE from 2.927% to 2.789% at 50%, from 3.877% to 2.787% at 63%, and from 3.841% to 2.365% at 78%, corresponding to relative reductions of 4.70%, 28.12%, and 38.43%, respectively. The average RMSE values show the same ordering across the split points, decreasing from the strongest baseline values of 1.431×10^{-6} , 1.951×10^{-6} , and 1.735×10^{-6} to 1.391×10^{-6} , 1.201×10^{-6} , and 1.013×10^{-6} at the three split points.

4.1.2. Sensitivity to the Chronological Split Point Is Cell-Dependent

More development data does not produce a uniformly monotonic trend for every model and cell. For Cell-1, the PADN MAPE decreases from 2.645% to 2.136% and then to 1.873% as the chronological split point moves from 50% to 63% and then to 78% of the

recorded timeline. For Cell-2, however, the corresponding MAPE values are 2.934%, 3.438%, and 2.857%, indicating that the combination with the split point at 63% is more difficult than that at 50% despite having a longer development segment. The baselines also show clear split-dependent behavior. The cell-averaged MAPE of the empirical-skeleton-only baseline, denoted EXP-only, remains comparatively stable at 3.962%, 3.877%, and 3.841%. In contrast, Direct-MLP changes from 2.927% at 50% to 9.298% at 63% and 5.289% at 78%; CNN-LSTM rises from 5.023% to 8.149% and remains at 7.310% at 78%; and Transformer changes more moderately from 5.683% to 5.912% and 5.423%. These patterns indicate that extrapolation difficulty is determined not only by the amount of observed data but also by the cell-specific trajectory shape, the position of the chronological split point, and the representativeness of the validation segment with respect to the subsequent test-horizon dynamics.

4.1.3. Origin of the 63% Split Anomaly

The relatively poor performance for Cell-2 with the split point at 63% is largely due to the location of the chronological split point. The position of a split point determines how much data is available before testing, but forecasting difficulty also depends on whether the point falls near a change in operating regime. For Cell-2, the chronological split point at 63% falls within the 903 K operating period, before the later 923 K high-temperature regimes listed in Table 1. The validation data, therefore, mainly reflect the 903 K regime, whereas the test segment soon moves into a different range of operating conditions and follows a different local resistance-growth trend.

This shift has a larger effect on the purely data-driven baselines. Direct-MLP may fit the local development trend too closely, while the rolling one-step models are additionally affected by error accumulation during autoregressive prediction. PADN is less sensitive to this change because the empirical degradation skeleton preserves the overall degradation trend, and the residual branch accounts for deviations from it. Even so, the 63% Cell-2 case remains more challenging than the corresponding 50% case.

4.1.4. Accuracy–Efficiency Tradeoff Is Favorable

The lower average errors of PADN are not obtained by using a uniformly larger neural backbone. Table 8 reports architecture-determined trainable parameter counts for the neural models; EXP-only is marked with a dash because it is a non-neural empirical ODE baseline and is not compared by neural parameter count. The “Avg. trainable params” entry is the arithmetic average of the Cell-1 and Cell-2 trainable parameter counts at the same split point, not an average over random seeds. Thus, the larger PADN parameter counts at the 63% and 78% split points reflect the validation-selected high-capacity architecture for those split points, whereas the average at the 50% split point combines two cell-specific PADN configurations. CNN-LSTM has 335,489 trainable parameters and Transformer has 282,369, whereas PADN uses 32,308 average parameters at the 50% split point and 249,094 at the 63% and 78% split points. Direct-MLP is deliberately matched much more closely to PADN in parameter count, with 32,335 average parameters at 50% and 249,408 at 63% and 78%. EXP-only indicates that the empirical anchor alone captures meaningful degradation trend information, especially for Cell-2 at the 63% split point, but its average errors remain higher than those of PADN at all three split points. The Direct-MLP comparison is therefore not explained by a larger PADN parameter budget, and the EXP-only comparison indicates that the empirical skeleton alone is not sufficient to reproduce PADN’s average performance. Because the compared models still differ in inductive bias and training objective, this result should be interpreted as benchmark-level evidence rather than as complete causal isolation of any single component.

Table 5. Prediction error in the comparative experiment with the chronological split point at 50% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE
Cell-1	PADN	<u>1.967×10^{-6}</u>	<u>1.397×10^{-6}</u>	<u>2.645%</u>
Cell-1	EXP-only	3.021×10^{-6}	2.641×10^{-6}	4.949%
Cell-1	Direct-MLP	2.085×10^{-6}	1.535×10^{-6}	2.877%
Cell-1	CNN-LSTM	3.821×10^{-6}	3.227×10^{-6}	5.921%
Cell-1	Transformer	4.798×10^{-6}	4.029×10^{-6}	7.367%
Cell-2	PADN	8.149×10^{-7}	6.285×10^{-7}	<u>2.934%</u>
Cell-2	EXP-only	<u>7.626×10^{-7}</u>	<u>6.129×10^{-7}</u>	2.975%
Cell-2	Direct-MLP	7.762×10^{-7}	6.337×10^{-7}	2.978%
Cell-2	CNN-LSTM	1.296×10^{-6}	9.068×10^{-7}	4.124%
Cell-2	Transformer	1.259×10^{-6}	8.780×10^{-7}	3.999%

Table 6. Prediction error in the comparative experiment with the chronological split point at 63% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE
Cell-1	PADN	<u>1.421×10^{-6}</u>	<u>1.153×10^{-6}</u>	<u>2.136%</u>
Cell-1	EXP-only	3.231×10^{-6}	3.094×10^{-6}	5.678%
Cell-1	Direct-MLP	5.901×10^{-6}	5.712×10^{-6}	10.520%
Cell-1	CNN-LSTM	5.254×10^{-6}	4.809×10^{-6}	8.731%
Cell-1	Transformer	4.481×10^{-6}	4.074×10^{-6}	7.397%
Cell-2	PADN	9.813×10^{-7}	7.619×10^{-7}	3.438%
Cell-2	EXP-only	<u>6.711×10^{-7}</u>	<u>4.655×10^{-7}</u>	<u>2.076%</u>
Cell-2	Direct-MLP	2.277×10^{-6}	1.794×10^{-6}	8.076%
Cell-2	CNN-LSTM	1.984×10^{-6}	1.670×10^{-6}	7.566%
Cell-2	Transformer	1.438×10^{-6}	9.954×10^{-7}	4.427%

Table 7. Prediction error in the comparative experiment with the chronological split point at 78% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE
Cell-1	PADN	<u>1.256×10^{-6}</u>	<u>1.035×10^{-6}</u>	<u>1.873%</u>
Cell-1	EXP-only	2.583×10^{-6}	2.457×10^{-6}	4.410%
Cell-1	Direct-MLP	1.728×10^{-6}	1.534×10^{-6}	2.759%
Cell-1	CNN-LSTM	4.922×10^{-6}	4.639×10^{-6}	8.282%
Cell-1	Transformer	3.185×10^{-6}	2.661×10^{-6}	4.710%
Cell-2	PADN	<u>7.698×10^{-7}</u>	<u>6.440×10^{-7}</u>	<u>2.857%</u>
Cell-2	EXP-only	8.864×10^{-7}	7.390×10^{-7}	3.272%
Cell-2	Direct-MLP	2.231×10^{-6}	1.777×10^{-6}	7.820%
Cell-2	CNN-LSTM	1.702×10^{-6}	1.434×10^{-6}	6.338%
Cell-2	Transformer	1.702×10^{-6}	1.392×10^{-6}	6.137%

4.1.5. Trajectory-Level Errors Mirror the Metrics

Figure 7 provides the trajectory-level counterpart to the tabulated errors. In each panel, the black curve gives the extracted R_0 trajectory over the full available period, whereas the model predictions are shown only on the held-out horizon to the right of the split marker. Panels (a)–(c) correspond to Cell-1, and panels (d)–(f) correspond to Cell-2, using chronological split points at 50%, 63%, and 78% of the recorded timeline. EXP-only captures a meaningful monotonic empirical trend and is especially competitive for Cell-2 with the split point at 63%, but it can over-rise for Cell-1 and does not consistently match the two-cell average behavior. Direct-MLP tracks PADN closely with the split point at 50%, but its long-horizon trend becomes less stable in

several combinations at 63% and 78%. The CNN-LSTM and Transformer curves tend to flatten or drift relative to the extracted trajectory after the split, while PADN remains closer to the observed trend in most panels. This visual pattern is consistent with the lower average errors in Tables 5–8, and it supports the interpretation that physics-anchored training helps PADN preserve long-horizon trend behavior under the recorded operating-condition profile.

Table 8. Average comparative performance and architecture-determined trainable parameter counts across the two cells. Bold and underlined values indicate the lowest average error at each split point.

Split Point (% of Timeline)	Model	Avg. RMSE	Avg. MAE	Avg. MAPE	Avg. Trainable Params
50%	PADN	<u>1.391×10^{-6}</u>	<u>1.013×10^{-6}</u>	<u>2.789%</u>	32,308
50%	EXP-only	1.892×10^{-6}	1.627×10^{-6}	3.962%	-
50%	Direct-MLP	1.431×10^{-6}	1.084×10^{-6}	2.927%	32,335
50%	CNN-LSTM	2.559×10^{-6}	2.067×10^{-6}	5.023%	335,489
50%	Transformer	3.029×10^{-6}	2.454×10^{-6}	5.683%	282,369
63%	PADN	<u>1.201×10^{-6}</u>	<u>9.576×10^{-7}</u>	<u>2.787%</u>	249,094
63%	EXP-only	1.951×10^{-6}	1.780×10^{-6}	3.877%	-
63%	Direct-MLP	4.089×10^{-6}	3.753×10^{-6}	9.298%	249,408
63%	CNN-LSTM	3.619×10^{-6}	3.239×10^{-6}	8.149%	335,489
63%	Transformer	2.959×10^{-6}	2.535×10^{-6}	5.912%	282,369
78%	PADN	<u>1.013×10^{-6}</u>	<u>8.397×10^{-7}</u>	<u>2.365%</u>	249,094
78%	EXP-only	1.735×10^{-6}	1.598×10^{-6}	3.841%	-
78%	Direct-MLP	1.979×10^{-6}	1.656×10^{-6}	5.289%	249,408
78%	CNN-LSTM	3.312×10^{-6}	3.037×10^{-6}	7.310%	335,489
78%	Transformer	2.444×10^{-6}	2.027×10^{-6}	5.423%	282,369

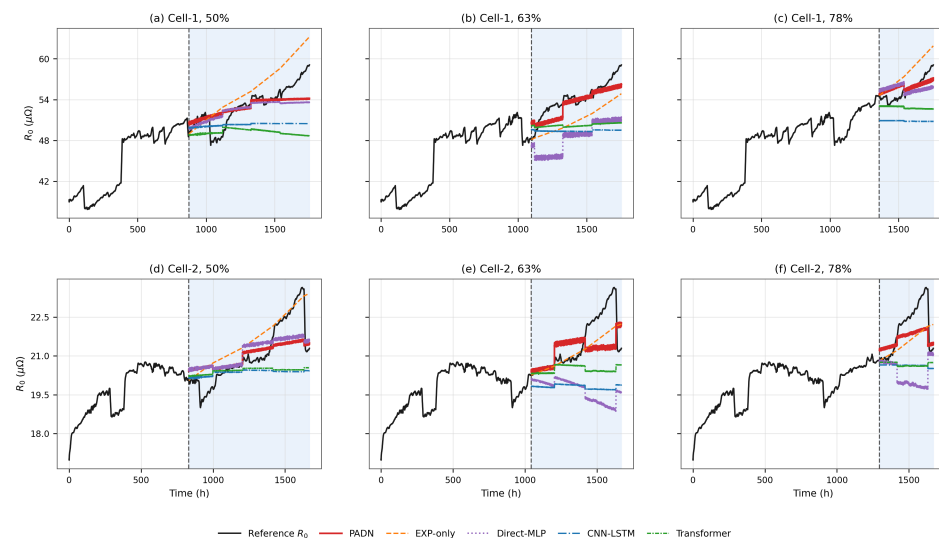


Figure 7. True and predicted equivalent ohmic-resistance health-indicator trajectories $R_0(t)$ plotted against operating time in the comparative experiment. The black curve spans the full available resistance trajectory for visual evaluation only, and the post-split reference values are not supplied to the models during inference. Model predictions are shown only on the test horizon after the vertical split marker. The light-blue shaded region indicates the prediction region after the chronological split point, without being used for any model updating, validation, or hyperparameter selection; it is used only to visualize prediction performance and compute evaluation metrics. EXP-only is a standalone empirical-anchor ODE baseline that fits the empirical-link parameters on the development segment and integrates the held-out horizon using the known future (U, T, t) profile. Direct-MLP is a data-only direct surrogate baseline using the same test-horizon inputs (t, U, T) as the PADN surrogate but without the empirical skeleton, physics-consistency loss, dynamic branch, or residual regularization. Panels (a–c) correspond to Cell-1 and panels (d–f) correspond to Cell-2; columns correspond to chronological split points at 50%, 63%, and 78% of the recorded timeline.

4.2. Ablation Results

4.2.1. All Three Components Improve Extrapolation

Tables 9–11 report the ablation results, and Table 12 summarizes the average MAPE increase across the two cells. Full PADN has the lowest RMSE, MAE, and MAPE in every cell–split-point combination, while removing the physics-consistency loss, the empirical degradation skeleton, or the residual regularization always increases the prediction error. With the split point at 50%, the average MAPE rises from 2.789% for Full PADN to 3.183–3.588% for the ablated variants. With the split point at 63%, the degradation becomes much larger, with average MAPE increasing from 2.787% to 6.246–7.082%. With the split point at 78%, the ablated variants still remain worse than Full PADN, increasing average MAPE from 2.365% to 3.711–5.184%. These results indicate that the three components make complementary contributions under the present two-cell conditional forecasting benchmark.

Table 9. Prediction error in the ablation experiment with the chronological split point at 50% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE	MAPE Increase
Cell-1	Full PADN	<u>1.967×10^{-6}</u>	<u>1.397×10^{-6}</u>	<u>2.645%</u>	–
Cell-1	Without phy	2.251×10^{-6}	1.726×10^{-6}	3.219%	0.574%
Cell-1	Without EXP	2.128×10^{-6}	1.595×10^{-6}	2.986%	0.341%
Cell-1	Without reg	2.299×10^{-6}	1.802×10^{-6}	3.350%	0.705%
Cell-2	Full PADN	<u>8.149×10^{-7}</u>	<u>6.285×10^{-7}</u>	<u>2.934%</u>	–
Cell-2	Without phy	9.525×10^{-7}	7.533×10^{-7}	3.481%	0.547%
Cell-2	Without EXP	1.156×10^{-6}	7.458×10^{-7}	3.379%	0.445%
Cell-2	Without reg	8.610×10^{-7}	8.015×10^{-7}	3.826%	0.893%

Table 10. Prediction error in the ablation experiment with the chronological split point at 63% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE	MAPE Increase
Cell-1	Full PADN	<u>1.421×10^{-6}</u>	<u>1.153×10^{-6}</u>	<u>2.136%</u>	–
Cell-1	Without phy	3.363×10^{-6}	3.178×10^{-6}	5.826%	3.691%
Cell-1	Without EXP	2.873×10^{-6}	2.683×10^{-6}	4.928%	2.792%
Cell-1	Without reg	2.717×10^{-6}	2.522×10^{-6}	4.635%	2.500%
Cell-2	Full PADN	<u>9.813×10^{-7}</u>	<u>7.619×10^{-7}</u>	<u>3.438%</u>	–
Cell-2	Without phy	2.311×10^{-6}	1.736×10^{-6}	7.783%	4.345%
Cell-2	Without EXP	2.203×10^{-6}	1.685×10^{-6}	7.563%	4.126%
Cell-2	Without reg	2.696×10^{-6}	2.117×10^{-6}	9.529%	6.091%

Table 11. Prediction error in the ablation experiment with the chronological split point at 78% of the recorded timeline. Bold and underlined values indicate the lowest error for each cell.

Cell	Model	RMSE	MAE	MAPE	MAPE Increase
Cell-1	Full PADN	<u>1.256×10^{-6}</u>	<u>1.035×10^{-6}</u>	<u>1.873%</u>	–
Cell-1	Without phy	2.019×10^{-6}	1.873×10^{-6}	3.408%	1.535%
Cell-1	Without EXP	1.495×10^{-6}	1.337×10^{-6}	2.438%	0.565%
Cell-1	Without reg	2.454×10^{-6}	1.982×10^{-6}	3.506%	1.633%
Cell-2	Full PADN	<u>7.698×10^{-7}</u>	<u>6.440×10^{-7}</u>	<u>2.857%</u>	–
Cell-2	Without phy	1.214×10^{-6}	9.650×10^{-7}	4.244%	1.387%
Cell-2	Without EXP	2.189×10^{-6}	1.798×10^{-6}	7.930%	5.073%
Cell-2	Without reg	1.125×10^{-6}	8.905×10^{-7}	3.916%	1.059%

Table 12. Average MAPE increase of ablated PADN variants across the two cells. Bold and underlined values indicate the lowest average MAPE at each split point.

Split Point (% of Timeline)	Model	Avg. MAPE	Absolute Increase	Relative Increase
50%	Full PADN	<u>2.789%</u>	—	—
50%	Without phy	3.350%	0.561%	20.108%
50%	Without EXP	3.183%	0.393%	14.093%
50%	Without reg	3.588%	0.799%	28.633%
63%	Full PADN	<u>2.787%</u>	—	—
63%	Without phy	6.804%	4.018%	144.186%
63%	Without EXP	6.246%	3.459%	124.130%
63%	Without reg	7.082%	4.295%	154.141%
78%	Full PADN	<u>2.365%</u>	—	—
78%	Without phy	3.826%	1.461%	61.769%
78%	Without EXP	5.184%	2.819%	119.183%
78%	Without reg	3.711%	1.346%	56.902%

4.2.2. Component Sensitivity Changes with the Split

The relative increase in Table 12 is computed from the unrounded average MAPE values by dividing the absolute MAPE increase by the corresponding Full PADN average MAPE and multiplying by 100%. At 50%, residual regularization is the most sensitive component on average: removing it raises the average MAPE from 2.789% to 3.588%, an absolute increase of 0.799% and a relative increase of 28.633%. At 63%, all three removals cause large degradation, and the largest average error again occurs without residual regularization, where average MAPE reaches 7.082% and the relative increase is 154.141%. The cell-level pattern is not identical, however, because Cell-1 is most affected by removing the physics-consistency loss at 63%, whereas Cell-2 is most affected by removing residual regularization. At 78%, removing the empirical degradation skeleton becomes the largest average failure mode, increasing the average MAPE from 2.365% to 5.184%, mainly because Cell-2's MAPE rises from 2.857% to 7.930%. Thus, no single component dominates every case; the empirical anchor, dynamic consistency, and residual constraint contribute differently depending on the split and cell trajectory.

4.2.3. Ablation Trajectories Are Consistent with the Failure Patterns

Figure 8 gives a qualitative view of the same component-level effects. The black curve again shows the full extracted R_0 trajectory, and ablated predictions appear only on the test horizon. Panels (a)–(c) correspond to Cell-1, and panels (d)–(f) correspond to Cell-2, following the same split-point order as Figure 7. The larger separations between Full PADN and the ablated curves in the 63% and 78% panels are consistent with the stronger MAPE increases in Table 12. Removing the empirical skeleton can bias the long-horizon trend, removing the physics-consistency loss weakens the coupling between the surrogate and residual networks, and removing residual regularization allows larger residual-rate corrections. The experiment does not include a separate hyperparameter-sensitivity sweep, so the supported claim is component-level sensitivity rather than a full loss-weight sensitivity analysis.

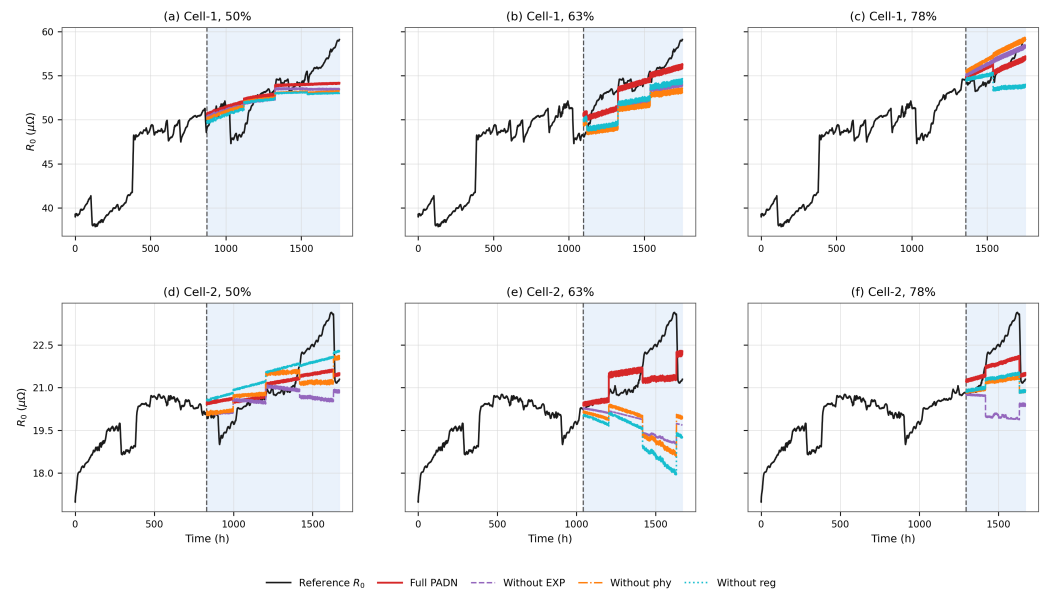


Figure 8. True and predicted equivalent ohmic-resistance health-indicator trajectories $R_0(t)$ plotted against operating time in the ablation experiment. The black curve spans the full available resistance trajectory for visual evaluation only, and the post-split reference values are not supplied to the models during inference. Ablated model predictions are shown only on the test horizon after the vertical split marker. The light-blue shaded region indicates the prediction region after the chronological split point, without being used for any model updating, validation, or hyperparameter selection; it is used only to visualize prediction performance and compute evaluation metrics. Panels (a–c) correspond to Cell-1 and panels (d–f) correspond to Cell-2; columns correspond to chronological split points at 50%, 63%, and 78% of the recorded timeline.

5. Conclusions

This study presents PADN as a gray-box framework for offline conditional forecasting of experiment-extracted and smoothed SOFC health-indicator trajectories under real variable-load operation, with the future operating-condition sequence assumed to be known. The prediction target is the extracted and smoothed equivalent ohmic-resistance trajectory, which is used here as a macroscopic equivalent health indicator rather than as a direct measurement of a single microscopic degradation mechanism. In the benchmark using the same chronological split points across methods, PADN achieves the lowest MAPE in five of six cell-split-point combinations and the lowest average RMSE, MAE, and MAPE at each split point among the compared methods. In the most demanding scenario, with the chronological split point at 50% of the recorded timeline, PADN maintains MAPE below 3% for both cells, with 2.645% for Cell-1 and 2.934% for Cell-2. The comparison with a two-parameter EXP-only baseline, a parameter-matched Direct-MLP baseline, and larger CNN-LSTM and Transformer baselines suggests that the advantage of PADN is unlikely to be explained solely by the empirical skeleton alone, direct surrogate evaluation, neural capacity, or rollout differences, although these factors remain only partly separable in the present benchmark.

The ablation results further show that the empirical degradation skeleton, physics-consistency loss, and residual-magnitude regularization play complementary roles. Removing any one of these components degrades forecasting accuracy, with especially clear deterioration when the chronological split point is at 63% or 78% of the recorded timeline. These observations support the central PADN design choice: the model should not replace physical trend information with a purely data-driven representation but should learn residual dynamics around an empirical degradation anchor while regularizing the residual correction.

This study does not establish broad generalization beyond the present benchmark. The current evidence remains limited to two single cells from one experimental platform, one label-construction pipeline, and a task setting in which future operating conditions are assumed to be known. It is therefore not formulated as direct RUL prediction or unknown-load forecasting. In particular, the target is obtained by reduced-order model inversion and subsequent smoothing, the empirical anchor is also model-informed, and the compared forecasting mechanisms are not fully isomorphic. This study therefore does not yet establish direct microscopic mechanism identification or a fully isolated causal attribution of the observed gains.

Future studies should evaluate PADN on a wider range of cell- and stack-level datasets covering more diverse operating profiles. This would provide a clearer assessment of its ability to generalize across cells and platforms. It would also be useful to include baselines with more closely matched mechanisms, examine the sensitivity of the results to health-indicator construction, and quantify prediction uncertainty. These analyses would help clarify how physical anchoring, the empirical degradation skeleton, surrogate inference, and the choice of health indicator each contribute to model performance.

Author Contributions: Conceptualization, T.Z. and J.X.; methodology, T.Z.; software, T.Z.; validation, T.Z.; formal analysis, T.Z.; investigation, Z.L.; resources, J.X.; data curation, T.Z.; writing—original draft preparation, T.Z.; writing—review and editing, T.Z.; visualization, T.Z.; supervision, J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used to support the findings of this study are available in the [SOC-Degradation] repository on GitHub: <https://github.com/chiyt14/SOC-Degradation>. These datasets were derived from the original files uploaded by the repository owner and are licensed under the terms of the repository’s license.

Acknowledgments: The authors thank the maintainers of the SOC-Degradation repository for making the variable-load SOFC aging data publicly available.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SOFC	Solid oxide fuel cell
PADN	Physics-Anchored Dual Network
EXP-only	Empirical-anchor-only ODE baseline
Direct-MLP	Direct multilayer-perceptron surrogate baseline
CNN-LSTM	Convolutional neural network–long short-term memory
RMSE	Root-mean-square error
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ODE	Ordinary differential equation
OCV	Open-circuit voltage
PINN	Physics-informed neural network

Appendix A. Notation Summary

Tables A1 and A2 summarize the principal symbols used in the PADN formulation, the health-indicator extraction procedure, and the experimental setup.

Table A1. Primary state, model, and loss symbols used in the manuscript.

Symbol	Meaning
t	Time coordinate of the health-indicator trajectory.
$\mathbf{x}(t)$	Operating-condition vector used as the conditioning input of the forecasting model.
$R_0^{\text{raw}}(t)$	Raw equivalent ohmic-resistance sequence obtained by pointwise inversion of the reduced-order cell model.
$R_0(t)$	Smoothed equivalent ohmic-resistance trajectory used as the health indicator and supervised learning target.
$R_{\text{ohm}}(T, t)$	Temperature-dependent ohmic-resistance term in the reduced-order cell model.
$\hat{R}_0(t)$	Predicted health-indicator trajectory from the surrogate network.
$\delta_d(t)$	Residual degradation-rate correction predicted by the dynamic network.
$g_{\text{emp}}(\mathbf{x}(t); \hat{\lambda})$	Empirical degradation-rate skeleton calibrated from experimental training data.
$\hat{\lambda}$	Identified parameter set of the empirical degradation skeleton.
θ_s	Trainable parameter set of the surrogate network.
θ_d	Trainable parameter set of the dynamic network.
N	Number of samples appearing in a loss or evaluation metric.
$\mathcal{L}_{\text{data}}$	Data-fitting loss between the predicted and target health-indicator trajectories.
$\mathcal{L}_{\text{phy}}$	Physics-consistency loss enforcing agreement with the anchored dynamic relation.
$\mathcal{L}_{\text{reg}}$	Residual-regularization loss limiting the magnitude of the learned residual dynamics.
α	Weight of the physics-consistency loss in the composite objective.
β	Weight of the residual-regularization loss in the composite objective.

Table A2. Primary experimental and reduced-order-model symbols used in the manuscript.

Symbol	Meaning
D_1	Initial I–V dataset used to identify the fixed parameters of the reduced-order cell model.
D_2	Periodic I–V dataset used only for a separately refitted cross-check of the extracted health indicator.
D_3	Continuous variable-load operation dataset providing the chronological sequence used for model development and testing.
ρ	Normalized chronological split point separating the earlier development segment from the later test segment.
N_{MA}	Moving-average window length used to smooth the raw health-indicator sequence.
U	Recorded cell-voltage channel used as a conditioning variable and in the reduced-order inversion.
I	Measured operating current used only in label extraction.
T	Recorded operating-temperature channel used as a conditioning variable and in the reduced-order inversion.
$E_{\text{OCV}}(T)$	Open-circuit-voltage term in the reduced-order cell model.
$c_{\text{act}}(T)$	Temperature-dependent coefficient associated with the activation-loss term.
$i_0(T)$	Temperature-dependent exchange-current-related parameter in the activation-loss term.
b_{ohm}	Arrhenius-type coefficient governing the temperature dependence of the ohmic-resistance term.
RMSE	Root-mean-square error used for trajectory-evaluation accuracy.
MAE	Mean absolute error used for trajectory-evaluation accuracy.
MAPE	Mean absolute percentage error used for trajectory-evaluation accuracy.

References

1. Singhal, S.C.; Kendall, K. *High Temperature Solid Oxide Fuel Cells: Fundamentals, Design and Applications*; Elsevier: Oxford, UK, 2003.
2. Dicks, A.L.; Rand, D.A.J. *Fuel Cell Systems Explained*, 3rd ed.; John Wiley & Sons: Chichester, UK, 2018.
3. Zarabi Golkhatmi, S.; Asghar, M.I.; Lund, P.D. A review on solid oxide fuel cell durability: latest progress, mechanisms, and study tools. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112339. [[CrossRef](#)]
4. Peng, J.; Zhao, D.; Xu, Y.; Wu, X.; Li, X. Comprehensive analysis of solid oxide fuel cell performance degradation mechanism, prediction, and optimization studies. *Energies* **2023**, *16*, 788. [[CrossRef](#)]
5. Yokokawa, H.; Tu, H.; Iwanschitz, B.; Mai, A. Fundamental mechanisms limiting solid oxide fuel cell durability. *J. Power Sources* **2008**, *182*, 400–412. [[CrossRef](#)]
6. Sasaki, K.; Haga, K.; Yoshizumi, T.; Minematsu, D.; Yuki, E.; Liu, R.; Uryu, C.; Oshima, T.; Ogura, T.; Shiratori, Y.; et al. Chemical durability of solid oxide fuel cells: Influence of impurities on long-term performance. *J. Power Sources* **2011**, *196*, 9130–9140. [[CrossRef](#)]

7. Dolenc, B.; Boskoski, P.; Stepancic, M.; Pohjoranta, A.; Juricic, D. State of health estimation and remaining useful life prediction of solid oxide fuel cell stack. *Energy Convers. Manag.* **2017**, *148*, 993–1002. [\[CrossRef\]](#)
8. Wu, X.; Ye, Q.; Wang, J. Fault diagnosis and prognostic of solid oxide fuel cells. *J. Power Sources* **2016**, *321*, 47–56. [\[CrossRef\]](#)
9. Khan, M.Z.; Mehran, M.T.; Song, R.H.; Lee, J.W.; Lee, S.B.; Lim, T.H. A simplified approach to predict performance degradation of a solid oxide fuel cell anode. *J. Power Sources* **2018**, *391*, 94–105. [\[CrossRef\]](#)
10. Mason, J.H.; Celik, I.B.; Lee, S.; Abernathy, H.; Hackett, G. Performance degradation predictions based on microstructural evolution due to grain coarsening effects in solid oxide fuel cell electrodes. *J. Electrochem. Soc.* **2018**, *165*, F64–F74. [\[CrossRef\]](#)
11. Zaccaria, V.; Tucker, D.; Traverso, A. A distributed real-time model of degradation in a solid oxide fuel cell, part I: Model characterization. *J. Power Sources* **2016**, *311*, 175–181. [\[CrossRef\]](#)
12. Linder, M.; Höcker, T.; Meier, C.; Holzer, L.; Friedrich, K.A.; Iwanschitz, B.; Mai, A.; Schuler, J.A. A model-based approach for current-voltage analyses to quantify degradation and fuel distribution in solid oxide fuel cell stacks. *J. Power Sources* **2015**, *288*, 409–418. [\[CrossRef\]](#)
13. Marra, D.; Sorrentino, M.; Pianese, C.; Iwanschitz, B. A neural network estimator of solid oxide fuel cell performance for on-field diagnostics and prognostics applications. *J. Power Sources* **2013**, *241*, 320–329. [\[CrossRef\]](#)
14. Rao, M.; Wang, L.; Chen, C.; Xiong, K.; Li, M.; Chen, Z.; Dong, J.; Xu, J.; Li, X. Data-driven state prediction and analysis of SOFC system based on deep learning method. *Energies* **2022**, *15*, 3099. [\[CrossRef\]](#)
15. Yang, K.; Liu, J.; Wang, Y.; Shi, X.; Wang, J.; Lu, Q.; Ciucci, F.; Yang, Z. Machine-learning-assisted prediction of long-term performance degradation on solid oxide fuel cell cathodes induced by chromium poisoning. *J. Mater. Chem. A* **2022**, *10*, 23683–23690. [\[CrossRef\]](#)
16. Li, M.; Wu, J.; Chen, Z.; Dong, J.; Peng, Z.; Xiong, K.; Rao, M.; Chen, C.; Li, X. Data-driven voltage prognostic for solid oxide fuel cell system based on deep learning. *Energies* **2022**, *15*, 6294. [\[CrossRef\]](#)
17. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems 30*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; pp. 5998–6008.
19. Wu, X.; Ye, Q.; Wang, J. A hybrid prognostic model applied to SOFC prognostics. *Int. J. Hydrogen Energy* **2017**, *42*, 25008–25020. [\[CrossRef\]](#)
20. Shi, W.; Zhu, J.; Han, M.; Sun, Z.; Guo, Y. Operating limitation and degradation modeling of micro solid oxide fuel cell-combined heat and power system. *Appl. Energy* **2019**, *252*, 113444. [\[CrossRef\]](#)
21. Chi, Y.; Qiu, Y.; Lin, J.; Song, Y.; Hu, Q.; Li, W.; Mu, S. Online identification of a link function degradation model for solid oxide fuel cells under varying-load operation. *Int. J. Hydrogen Energy* **2022**, *47*, 2622–2646. [\[CrossRef\]](#)
22. Wu, X.; Xu, L.; Wang, J.; Yang, D.; Li, F.; Li, X. A prognostic-based dynamic optimization strategy for a degraded solid oxide fuel cell. *Sustain. Energy Technol. Assess.* **2020**, *39*, 100682. [\[CrossRef\]](#)
23. Wu, X.; Xu, L.; Wang, J.; Yang, D.; Zhang, M.; Li, X. Discharge performance recovery of a solid oxide fuel cell based on a prognostic-based control strategy. *J. Power Sources* **2020**, *480*, 229102. [\[CrossRef\]](#)
24. Cheng, T.; Qin, H.; Cheng, Z.; Yan, D.; Jia, L.; Jiang, J.; Li, J. Fault prognosis control of solid oxide fuel cell system based on health evaluation. *Int. J. Hydrogen Energy* **2021**, *46*, 38425–38450. [\[CrossRef\]](#)
25. Shi, Y.; Zheng, H.; Jiang, W.; Song, M.; Luo, Y.; Zhang, X.; Tu, S.-T. Hybrid neural network and dynamic decay model for life prediction of solid oxide fuel cell combined heat and power systems. *J. Power Sources* **2025**, *645*, 237165. [\[CrossRef\]](#)
26. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [\[CrossRef\]](#)
27. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [\[CrossRef\]](#)
28. Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; Chen, X. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nat. Commun.* **2024**, *15*, 4332. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Ko, T.; Kim, D.; Park, J.; Lee, S.H. Physics-informed neural network for long-term prognostics of proton exchange membrane fuel cells. *Appl. Energy* **2025**, *382*, 125318. [\[CrossRef\]](#)
30. Xiong, R.; He, Y.; Sun, Y.; Jia, Y.; Shen, W. Enhanced electrode-level diagnostics for lithium-ion battery degradation using physics-informed neural networks. *J. Energy Chem.* **2025**, *104*, 618–627. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.