

Physics-Informed Machine Learning for Diagnostics and Prognostics in Solid Oxide Fuel Cells

by

Zeynab Salehi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering

University of Alberta

© Zeynab Salehi, 2026

Abstract

Solid oxide fuel cells (SOFCs) are promising energy conversion systems, but their widespread deployment is limited by complex degradation mechanisms that reduce performance and lifetime. Accurate degradation diagnosis and prognostics remain challenging due to the strong coupling between electrochemical processes and operating conditions. This thesis develops a comprehensive physics-informed and data-driven framework for SOFC health monitoring, integrating experimental degradation studies, uncertainty-aware prognostics, impedance-based diagnostics, and unified multi-task learning. Overall, the proposed methodologies provide a scalable and interpretable solution for predictive maintenance and health-aware control of SOFC systems, supporting their reliable deployment in future clean energy applications.

A series of controlled degradation experiments are conducted on anode-supported tubular SOFCs under four representative degradation mechanisms: redox cycling, fuel starvation, internal short circuit, and thermal shock, along with a long-term normal aging test. These experiments generate a diverse, multi-modal dataset containing voltage, current, polarization curves, and electrochemical impedance spectroscopy (EIS) measurements, providing a foundation for developing and validating advanced learning-based models.

First, a physics-informed machine learning framework is developed for real-time state-of-health (SOH) estimation using a temporal graph convolutional network (TGCN). Physically meaningful features are extracted from EIS data

using distribution of relaxation times (DRT) analysis, enabling the model to capture interactions among underlying electrochemical processes. The proposed approach significantly improves forecasting accuracy compared to conventional data-driven models and provides interpretable degradation tracking. Building on this, a probabilistic prognostics framework is proposed for long-horizon SOH forecasting and remaining useful life (RUL) estimation. Attention-based Informer and LSTM models are extended with uncertainty quantification, capturing both aleatoric and epistemic uncertainties to produce calibrated confidence intervals for SOH trajectories and RUL predictions. To address the challenge of shrinking forecast horizons near end-of-life, a dynamic forecasting strategy based on exponentially weighted zero-padding is introduced, enabling consistent sequence modeling and stable training.

In parallel, a diagnostic framework is developed to identify degradation mechanisms using impedance-derived features. EIS measurements are transformed into frequency-domain signatures via DRT analysis, and multiple machine learning classifiers are evaluated, including ensemble methods and temporal deep learning models. Results demonstrate that attention-based sequence models provide superior performance in distinguishing between multiple degradation modes, enabling accurate, mechanism-aware fault identification.

Finally, a unified multi-task learning (MTL) framework is proposed to integrate diagnostics and prognostics into a single model. By conditioning RUL prediction on degradation mechanism information and transferring latent representations from the classifier to the prognostic model, the proposed approach achieves significantly improved prediction accuracy and computational efficiency. This mechanism-aware prognostic framework enables more robust and reliable lifetime prediction across diverse degradation scenarios.

Preface

This thesis presents original work by Zeynab Salehi. The research was conducted using the experimental facilities at the University of Alberta Energy Mechatronics Lab (EML) led by Dr. Mahdi Shahbakhti.

The content of each chapter is fully or partially based on published or submitted papers in peer-reviewed journals and conferences [1–7]. As the lead author, I made substantial contributions to all peer-reviewed publications included in this thesis. The specific contributions of co-authors to each thesis chapter are summarized below.

In Chapter 2, experimental testing, data acquisition, and domain-specific knowledge related to fuel cells were developed in collaboration with Dr. Amir Reza Hanifi, Dr. Masood Fakouri Hasanabadi, Dr. Sajad Vafaenezhad, Saeed Hanifi, and Shakiba Sharifi.

For Chapter 3 [1], Dr. Mahdi Shahbakhti, Dr. Bob Koch, Daniel J. Smith (Cummins Inc.), and Alireza Salahi contributed through technical discussions and manuscript review. Dr. Amir Reza Hanifi and Dr. Masood Fakouri Hasanabadi provided expertise in fuel cell physics and guidance on experimental design, with Dr. Masood Fakouri Hasanabadi also contributing to data collection.

For Chapter 4 [2, 3], Dr. Mahdi Shahbakhti and Daniel J. Smith contributed through technical discussions and manuscript review. Dr. Amir Reza Hanifi and Dr. Masood Fakouri Hasanabadi provided expertise in fuel cell physics.

For Chapter 5 [4], Dr. Mahdi Shahbakhti, Dr. Sayedabbas Sobhi, and Daniel J. Smith contributed to the conceptual development of data-driven models. Dr. Amir Reza Hanifi contributed to the design of fault scenarios for exper-

imental data acquisition. Saeed Hanifi assisted with conducting experiments and collecting data.

For Chapter 6 [5], Dr. Mahdi Shahbakhti and Daniel J. Smith contributed to the conceptual development of data-driven fuel cell models. Dr. Amir Reza Hanifi provided expertise in fuel cell physics.

For the appendix chapter [6, 7], Dr. Mahdi Shahbakhti, Dr. Bob Koch, Dr. Mohammadali Tofigh, Daniel J. Smith, and Dr. Ali Kharazmi (Cummins Inc.) contributed through technical discussions and manuscript review. Dr. Amir Reza Hanifi and Dr. Sajad Vafaeenezhad provided expertise in fuel cell physics and guidance on experimental design, with Dr. Sajad Vafaeenezhad also contributing to data collection.

To my family

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Mahdi Shahbakhti, for his guidance, encouragement, and unwavering support throughout my studies. His mentorship has played a crucial role in the successful completion of this work.

I am also deeply thankful to Dr. Bob Koch for his valuable advice and support, and to Dr. Scott Dick for his insightful feedback during my research. I would also like to acknowledge Daniel J. Smith, Dr. Ali Kharazmi, Dr. Mohamadali Tofigh, and Alireza Salahi for their contributions to our joint publications. In addition, I am grateful to Dr. Amir Reza Hanifi, Dr. Sajad Vafaeenezhad, Dr. Masood Fakouri Hasanabadi, Saeed Hanifi, and Shakiba Sharifi for their significant assistance in design of experiments, data collection, and for sharing their valuable expertise in fuel cell physics.

I would like to extend my heartfelt appreciation to my beloved family for their unconditional support, patience, and encouragement throughout this journey and throughout my life. I am also grateful to my dear friends for their continuous support and for making this journey more enjoyable and meaningful.

Finally, I gratefully acknowledge the financial support provided by Cummins Inc., the Natural Sciences and Engineering Research Council of Canada (NSERC), Alberta Innovates, and the University of Alberta Future Energy Systems (FES).

Contents

1	Background and Introduction	1
1.1	Fundamental of SOFC	1
1.2	Literature Review	4
1.2.1	Degradation Modes	4
1.2.2	Diagnostic and Monitoring Tools	8
1.2.3	Degradation Prediction Methods	22
1.2.4	Degradation Mode Classification Methods	33
1.3	Research Gaps	34
1.4	Thesis Contributions and Outline	36
1.4.1	Thesis Contributions	36
1.4.2	Thesis Outline	41
2	Experimental Setup and Datasets	43
2.1	Introduction	43
2.2	Anode-supported Tubular Cells	44
2.3	Experimental Setup	45
2.4	Accelerated Degradation Experiments	47
2.4.1	Redox Cycling	47
2.4.2	Fuel Starvation	48
2.4.3	Short Circuit	49
2.4.4	Thermal Shock	50
2.5	Degradation Dataset Summary	51
3	Physics-Informed State-of-Health Estimation	55
3.1	Introduction	55
3.2	Data Processing and Feature Engineering	56
3.3	Physics-informed Machine Learning Model	63
3.4	Results and Discussion	73
3.4.1	Hyper-parameter Optimization	73
3.4.2	Model Comparison	74
3.5	Summary of Chapter	80
4	Remaining Useful Life Forecasting with Uncertainty	82
4.1	Introduction	82
4.2	RUL Forecasting Models	83
4.2.1	Dynamic Forecast Horizon	86
4.2.2	Uncertainty Quantification	87
4.3	Results and Discussions	91
4.3.1	Hyper-parameter Optimization	91
4.3.2	Model Comparison	92
4.3.3	Cross-Dataset Generalization	96
4.3.4	Sensitivity Analysis of Padding Strategy and Decay Factor	101

4.3.5	Deployment Considerations and Applications	101
4.4	Summary of Chapter	104
5	Degradation Mode Classification	106
5.1	Introduction	106
5.2	Data Processing and Class Labeling	107
5.3	Shallow Models for Degradation Mode Classification	109
5.3.1	Feature Engineering	110
5.3.2	Handling Class Imbalance via Oversampling	112
5.3.3	Random Forest Classifier	117
5.3.4	XGBoost Classifier	119
5.3.5	RF-XGB Ensemble Model	120
5.4	Deep Models for Degradation Mode Classification	120
5.4.1	Physically Informed Time-Series Augmentation and Screen- ing	121
5.4.2	LSTM Classifier	124
5.4.3	TCN Classifier	127
5.4.4	Attention-based Informer Classifier	129
5.5	Results and Discussion	131
5.5.1	Quantitative Evolution of DRT-Derived Features	131
5.5.2	Evaluation Strategy: LODO	133
5.5.3	Evaluation Metrics	133
5.5.4	Hyperparameters Tuning	135
5.5.5	Ablation Study of the Time-Series Augmentation Strategy	138
5.5.6	Models Performance and Computational Comparisons	139
5.5.7	Fold-wise Classification Uncertainty Analysis	143
5.5.8	Scope of Applicability and Deployment Considerations	145
5.6	Summary of Chapter	146
6	Degradation Mechanism-aware Prognostics	148
6.1	Introduction	148
6.2	Data Processing and Class Labeling	149
6.3	Mechanism-aware Remaining Useful Life Prediction	153
6.3.1	Single-task RUL Estimator	153
6.3.2	Degradation Mode Classifier	155
6.3.3	Feature Fusion	157
6.3.4	Asymmetric MTL	161
6.4	Results and Discussion	164
6.4.1	Evaluation Strategy and Metrics	164
6.4.2	Results and Discussion	164
6.5	Summary of Chapter	170
7	Conclusions and Future Work	172
7.1	Conclusions	172
7.1.1	Experimental Framework and Dataset Development	172
7.1.2	Physics-Informed and Probabilistic Prognostics	173
7.1.3	Impedance-Based Diagnostics	174
7.1.4	Unified Mechanism-Aware Prognostics Framework	174
7.1.5	Overall Contributions and Impact	175
7.2	Future Works	177
7.2.1	Health-aware and fault-tolerant control integration	177
7.2.2	Integration with hybrid energy management systems	177
7.2.3	Uncertainty-aware control and decision-making	178
7.2.4	Risk-aware maintenance scheduling and planning	178

7.2.5	External validation	178
7.2.6	Scaling to stack-level and fleet-level health management	178
7.2.7	Online transfer learning and adaptation	179
7.2.8	Real-time implementation	179
7.2.9	Minimum experimental dataset requirements	179
7.2.10	Closed-loop co-design of cells, experiments, and prognostics	180
References		181
Appendix A Transfer Learning for Cross-Domain SOFC Modeling		204
A.1	Tubular Cell Healthy Datasets	205
A.2	Planar Cell Healthy Datasets	205
A.3	DNN Model Development	207
A.4	Transfer Learning	208
A.5	Results and Discussion	210
A.6	Appendix Summary	216
Appendix B Benchmark SOH Forecasting Models		217
Appendix C Ph.D. Publications		220
Appendix D Thesis Files		222
D.1	Chapter 1 Source Files	222
D.2	Chapter 2 Source Files	223
D.3	Chapter 3 Source Files	224
D.4	Chapter 4 Source Files	225
D.5	Chapter 5 Source Files	226
D.6	Chapter 6 Source Files	227
D.7	Chapter 7 Source Files	228

List of Tables

2.1	Redox datasets conditions. Each test is performed in one of the eight fabricated cell. Cells N1 to N6 underwent normal Redox cycling, and cells R1 and R2 underwent random Redox cycling datasets. T_{cell} , P_{max} , $t_{\text{H}_2 \text{ cutoff}}$, and τ are cell temperature, maximum power, and hydrogen cutoff interval, and cell lifetime. All tests were conducted under constant voltage of 0.7 V and constant air volume flow rate ($\dot{V}_{\text{air}}$) of 100 SCCM. Hydrogen volume flow rate ($\dot{V}_{\text{H}_2}$) changed between 0 and 50 SCCM.	49
2.2	Summary of all degradation datasets and corresponding sample counts.	54
3.1	LSTM model hyperparameter search space and optimal values	73
3.2	CNN-LSTM model hyperparameter search space and optimal values	74
3.3	GCN-LSTM model hyperparameter search space and optimal values	74
3.4	Prediction accuracy comparison of the three SOH prediction models on R2 dataset, with look-back of 3 hours and forecast horizon of 6 hours. MAPE computed on samples with SOH > 0.05.	75
3.5	Prediction accuracy comparison of the GCN-LSTM model at different points of forecast horizon H. MAPE computed on samples with SOH > 0.05.	76
3.6	Performance comparison of SOH forecasting models. The reported p -values are obtained from one-sided paired Wilcoxon signed-rank tests comparing the trajectory-level RMSE of each benchmark model with that of the GCN-LSTM.	78
3.7	Cross-dataset validation results of GCN-LSTM trained on seven datasets and tested on the remaining one.	79
3.8	Online computational profile on NVIDIA Tesla T4 GPU for model inference ($t_{\text{inference}}$). For CNN-LSTM, convolutional feature extraction is reported separately as t_{feature} , although it is embedded within the model forward pass. For GCN-LSTM, DRT feature extraction is performed using a MATLAB implementation on a Mac 8-Core CPU, requiring 278 ms per EIS curve.	80
4.1	Search space and optimized hyperparameters for the <u>Bayesian LSTM</u> model	94
4.2	Search space and optimized hyperparameters for the <u>Bayesian Informer</u> model	95

4.3	Performance comparison of deterministic and Bayesian forecasting models for SOH trajectory prediction.	96
4.4	Computational cost comparison between Bayesian LSTM and Bayesian Informer on an NVIDIA Tesla T4 GPU. Inference time is reported for 10 MC dropout iterations.	97
4.5	LODO validation results for the Bayesian Informer model. Each fold corresponds to one dataset excluded from training and used for testing.	97
4.6	Bayesian Informer performance when trained on N1–N6 and tested on unseen datasets R1 and R2.	100
4.7	Bayesian Informer performance when trained on N1–N5 and tested on unseen datasets R1, R2, and N6.	100
4.8	Sensitivity analysis of padding strategies and decay factor λ for the Bayesian Informer model.	102
5.1	Percentage change of DRT-derived features and ohmic resistance from BOL to EOL. Values represent mean $\pm$ standard deviation across three repeated experiments for each failure mode.	133
5.2	Random Forest classifier hyperparameter search space and optimal values	135
5.3	XGBoost classifier hyperparameter search space and optimal values	136
5.4	LSTM classifier hyperparameter search space and optimal values	136
5.5	TCN classifier hyperparameter search space and optimal values	136
5.6	Informer (encoder-only) classifier hyperparameter search space and optimal values	136
5.7	RF+XGB ensemble hyperparameter search space and optimal values	137
5.8	Ablation study of the proposed augmentation strategy using the Informer classifier.	138
5.9	Performance comparison of degradation mode classifiers	140
5.10	Computational cost comparison of degradation mode classifiers. Training time is averaged across LODO folds. All experiments were conducted on an NVIDIA T4 GPU.	142
5.11	Fold-level F1 scores obtained from LOFO validation.	144
5.12	Fold-level F1 summary across the 12 LOFO folds. Mean, standard deviation (SD), and 95% confidence interval (CI) are reported for each model. The last two columns show the mean fold-wise F1 difference and paired Wilcoxon signed-rank p-value for each baseline relative to Informer.	144
5.13	Online computational profile of the diagnostic framework. Model inference time ($t_{\text{inference}}$) is measured on an NVIDIA Tesla T4 GPU. DRT feature extraction is performed using a MATLAB implementation on a Mac 8-Core CPU and requires 278 ms per EIS curve. For shallow models, engineered feature calculation adds 3.36 ms to the feature-processing time.	146
6.1	Fold-wise RUL prediction performance comparison between the baseline RUL estimator and fusion-based models under LODO cross-validation.	165
6.2	Pairwise Euclidean centroid distances between degradation modes in the latent space for different fusion strategies, averaged across LODO folds. Distances reflect the relative separability learned by the encoder.	167

6.3	Fold-wise RUL prediction performance comparison between the baseline RUL estimator and the asymmetric MTL model under LODO cross-validation. The magnitude of improvement varies across degradation modes and reflects classifier diagnostic reliability.	168
6.4	Comparison of RUL estimation strategies and their performance improvements relative to the baseline single-task RUL estimator.	168
6.5	Computational cost comparison between the single-task RUL estimator, degradation mode classifier, their sequential execution (RUL+Classifier), and the proposed asymmetric MTL framework under LODO cross-validation.	169
6.6	Standalone classifier performance of the degradation mode classifier under LODO cross-validation.	170
6.7	Classification performance of the degradation mode classifier when integrated into the asymmetric MTL framework under LODO cross-validation.	170
A.1	Prediction accuracy of model on future data of target domains (A2 and A3) with transfer learning. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.	212
A.2	Different percentage of data in target domain for TL fine-tuning for cells A2 and A3, normalization 2. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.	212
A.3	Results of transfer learning from tubular cell A1 to planar cell using normalization 2 and 10% of fine-tuning data. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.	213
B.1	Multivariate LSTM hyperparameter search space and optimal values	218
B.2	GRU model hyperparameter search space and optimal values .	218
B.3	TCN model hyperparameter search space and optimal values .	219
D.1	Chapter 1 source files	222
D.2	Chapter 2 source files	223
D.3	Chapter 3 source files	224
D.4	Chapter 4 source files	225
D.5	Chapter 5 source files	226
D.6	Chapter 6 source files	227
D.7	Chapter 7 source files	228

List of Figures

1.1	Schematic of an anode-supported tubular SOFC illustrating the cell structure, electron flow through the external circuit, oxygen ion (O^{2-}) transport through the electrolyte, and the electrochemical reactions occurring at the electrodes.	2
1.2	General workflow of SOFC health management, including data acquisition, anomaly/fault detection, degradation mode diagnostics, prognostics, and health management. The highlighted section indicates the primary focus of this thesis.	4
1.3	Overview of major SOFC health monitoring research directions, including degradation mechanisms, diagnostic and monitoring tools, and degradation prediction methodologies discussed in this thesis.	5
1.4	Delamination between the electrolyte and anode functional layer as a result of redox cycling	7
1.5	Overview of diagnostic and monitoring tools used for SOFC health assessment, including voltage-based monitoring, electrochemical impedance spectroscopy (EIS), nonlinear EIS techniques, and gas and temperature monitoring methods.	10
1.6	Schematic illustration of electrochemical impedance spectroscopy (EIS) measurements for SOFC characterization. A small-amplitude sinusoidal current perturbation is applied through the external circuit over a range of excitation frequencies, and the resulting voltage response is measured using a frequency response analyzer (FRA). The measured phase-shifted voltage and current signals are used to calculate the complex impedance and construct the Nyquist impedance spectrum.	11
1.7	EIS curves of SOFC. The intersection of the curve with real axis represents Ohmic resistance, and the distance from that point to the end of the curve is Polarization resistance. The summation of the Ohmic and Polarization resistances is the total resistance of the cell	13
1.8	ECM used for the CNLS fit of the impedance spectra of planar cell [60].	15
1.9	ECM used for the CNLS fit of the impedance spectra of tubular and microtubular cell [62].	16
1.10	DRT curves for a SOFC under redox cycling. The evolution of DRT peaks illustrates the progression of electrochemical processes and degradation over time.	20
1.11	Classification of SOFC degradation prediction methods discussed in the literature, including representative image-based, physics-based, data-driven, and hybrid approaches.	23

1.12	Schematic overview of the thesis organization. Journal publications/submissions from this thesis are shown in brackets for each part of the thesis.	42
2.1	A SEM image of a fabricated anode-supported SOFC showing the main components and the dimensions.	45
2.2	Cell configuration and experimental test bench used for SOFC electrochemical characterization.	46
2.3	Hydrogen flow rate profile applied during redox cycling experiments. The anode fuel supply was periodically switched between 50 SCCM and 0 SCCM, with the fuel-off interval increased over time to accelerate degradation.	48
2.4	Short circuit degradation experiment.	51
2.5	Thermal shock degradation protocol.	52
3.1	SOFC performance evolution during Redox cycling for dataset N4. (a) The current density profile shows repeated hydrogen interruption and recovery events at a constant operating voltage of 0.7 V. (b) The corresponding EIS and J-V curves after selected cycles illustrate progressive increases in polarization resistance and performance degradation. A unified color scheme and shared legend are used for both plots to emphasize the correspondence between electrochemical impedance and cell performance.	58
3.2	DRT evolution for the N4 dataset corresponding to the Redox cycling test shown in Figure 3.1. The DRT spectra are plotted for selected cycles using the same color scheme as the EIS and J-V curves in Figure 3.1b. The characteristic frequencies of the major peaks remain nearly constant over time, while their amplitudes increase, indicating the progressive growth of polarization losses associated with electrode processes. BOL stands for beginning of life.	61
3.3	Overview of the proposed physics-informed GCN-LSTM model: (a) graph structure linking impedance components and SOH, (b) two-layer message passing for feature embedding, and (c) temporal forecasting architecture integrating GCN features with LSTM dynamics.	69
3.4	Inputs and outputs of three studied models. L: look-back, H: forecast horizon	70
3.5	SOH forecasting results for LSTM (a), CNN-LSTM (b), and GCN-LSTM (c) on the R2 dataset.	75
3.6	Results of GCN-LSTM network in SOH forecasting with different forecast horizons.	77
3.7	Timeline of online inference workflow. Each EIS measurement requires 2 minutes (t_{EIS}), followed by feature extraction and model inference ($t_{\text{feature}} + t_{\text{inference}} < 290$ ms). The interval between successive EIS updates (t_{interval}) depends on the Redox cycling profile (10-35 minutes). Since feature extraction and inference latency are negligible compared to EIS acquisition, the limiting factor for real-time deployment is the EIS measurement itself.	80

4.1	Bayesian LSTM architecture. Stacked LSTMs with dropout layers predict future SOH trajectories from historical SOH and impedance inputs.	84
4.2	Bayesian Informer architecture. An encoder–decoder with ProbSparse attention models long-range dependencies to forecast SOH sequences.	86
4.3	Proposed prognostics workflow. EIS/J–V data are processed into SOH and impedance features, windowed with zero-padding, and fed to the Bayesian Informer model. Aleatoric and epistemic uncertainties are quantified and RUL distributions obtained via FHT analysis. The results can inform control and maintenance decisions.	90
4.4	Training and validation loss curves for the <u>Bayesian Informer</u> model.	93
4.5	Training and validation loss curves for the <u>Bayesian LSTM</u> model.	93
4.6	<u>Bayesian LSTM</u> model results	98
4.7	<u>Bayesian Informer</u> model results	99
4.8	Practical deployment workflow of the proposed SOFC prognostics and control framework. Green blocks represent the stages implemented and validated in this work, including data collection, preprocessing, model training, validation, and probabilistic SOH–RUL prediction. Yellow blocks illustrate downstream control-oriented applications enabled by the prognostics outputs, such as optimal and robust control, diagnostics, and maintenance planning. Red blocks denote the low-level controllers responsible for executing control actions in the physical system.	105
5.1	Representative DRT evolution under the four degradation mechanisms	109
5.2	Overview of the proposed time-series-specific data augmentation and screening pipeline for addressing class imbalance. The framework combines RTW and amplitude jittering, with data-driven calibration from repeated experiments. Augmented samples are subsequently screened using Mahalanobis distance analysis to ensure physical plausibility.	121
5.3	Architecture of the LSTM classifier for degradation mode identification using windowed time-series inputs $\mathbf{X}_{t-L:t} = [R_{\text{ohm}}, P_1, P_2, P_3, P_4]_{t-L:t}$	125
5.4	Architecture of the TCN classifier consisting of residual dilated causal convolution blocks, followed by global temporal pooling and fully connected layers for degradation mode prediction.	128
5.5	Architecture of the Informer classifier, including temporal embedding, multi-head ProbSparse self-attention encoding, temporal pooling, and a fully connected prediction head for degradation mode classification.	132
5.6	Confusion matrices for all evaluated classifiers. Cell values denote sample counts, while the color bar indicates normalized intensity in the range $[0, 1]$	141
6.1	Representative SOL trajectories from BOL to EOL under the four experimentally induced degradation mechanisms.	150

6.2	Overview of the baseline RUL estimation pipeline and the investigated fusion strategies for incorporating degradation mode information.	153
6.3	Proposed asymmetric MTL architecture for degradation mechanism-aware prognostics.	154
6.4	Architecture of the baseline single-task RUL estimator based on an Informer-style encoder. The multivariate input sequence, consisting of SOH and impedance-derived features, is embedded and augmented with positional encoding before being processed by stacked ProbSparse self-attention blocks. Temporal pooling is applied to obtain a fixed-length latent representation, which is mapped to a scalar RUL prediction through fully connected layers.	156
6.5	Comparison of fusion strategies for incorporating degradation mechanism information into the RUL estimator.	158
6.6	Proposed asymmetric MTL framework for joint degradation mode classification and RUL estimation. Two task-specific encoders process the same input sequence: one optimized for RUL estimation and the other for degradation mode classification. The classifier latent embedding is transferred to the RUL branch via late-stage feature concatenation.	163
6.7	Silhouette score of training embeddings versus test RMSE for baseline, early fusion, and late fusion models across LODO folds.	166
7.1	Summary of the proposed SOFC diagnostics and prognostics framework developed in this thesis, including experimental degradation testing, impedance-based feature extraction, model selection, training, and validation strategies for physics-informed and data-driven health monitoring.	176
A.1	J–V curves of three SOFC cells under 18 operating conditions. Cell temperature (T_c), hydrogen ($\dot{V}_{H_2}$), nitrogen ($\dot{V}_{N_2}$), and air ($\dot{V}_a$) flow rates are varied.	206
A.2	J–V Curves of the planar cell at 10 Operating Conditions. Cell Temperature (T_c), hydrogen ($\dot{V}_{H_2}$), nitrogen ($\dot{V}_{N_2}$), and Air ($\dot{V}_a$) flow rate conditions are changed.	207
A.3	Transfer Learning process from A1 to A2, A3, and planar cells. After training the DNN using source domain data, the trained network transfers to target domain and its weights are fine-tuned using the 1, 5, 10, and 20% of the target domain data.	209
A.4	A1: Actual vs. predicted output voltage and prediction error using DNN for test dataset. DNN model contains one input layer, 6 hidden layers, and one output layer with the number of neurons of 4/100/80/60/40/20/10/1. Experimental data corresponds to Figure A.1a.	210
A.5	Cell A1: Training and validation losses during the training process of DNN	211
A.6	Results of transfer learning from cell A1 to cell A2 (90% untrained)	213
A.7	Results of transfer learning from cell A1 to cell A3 (90% untrained)	214
A.8	Results of transfer learning from tubular cell A1 to planar cell (90% untrained)	215

B.1	Schematic representation of the benchmark forecasting models using the same multivariate input features. The inputs include historical DRT peak values, ohmic resistance, and SOH, and the models predict future SOH values over the forecast horizon.	218
B.2	SOH forecasting results for Multi-variate LSTM (a), GRU (b), and TCN (c) on the R2 dataset.	219

Abbreviations and Acronyms

Adam	Adaptive moment estimation.
AFL	Anode functional layer.
ANN	Artificial neural network.
ARIMA	Autoregressive integrated moving average.
BPNN	Back propagation neural network.
CFD	Computational fluid dynamics.
DNN	Deep neural network.
DRT	Distribution of relaxation times.
ECM	Equivalent circuit model.
EIS	Electrochemical impedance spectroscopy.
EOL	End of life.
FC	Fully connected (layer).
FFT	Fast Fourier transform.
FIB	Focused ion beam.
GA	Genetic algorithm.
GBDT	Gradient-boosted decision trees.
GCN	Graph convolutional network.
GDC	Gadolinium-doped ceria.
GRU	Gated recurrent unit.
GWO	Grey Wolf optimization.
HSMM	Hidden semi-Markov model.

IQR	Interquartile range.
J–V	Current density–voltage curve.
KF	Kalman filter.
LOFO	Leave-one-file-out cross-validation.
LSC	Lanthanum strontium cobaltite.
LS-SVM	Least squares support vector machine.
LSTM	Long short-term memory.
Macro-F1	Macro-averaged F1-score.
MAD	Median absolute deviation.
MAE	Mean absolute error.
MAPE	Mean absolute percentage error.
MLP	Multi-layer perceptron.
MOGPR	Multi-output Gaussian process regression.
MSE	Mean square error.
NNARMAX	Neural network autoregressive moving average with external inputs.
NNARX	Neural network autoregressive with external inputs.
NNOE	Neural network output error.
OC	Operating condition.
OCV	Open-circuit voltage.
PBTL	Parameter-based transfer learning.
PIML	Physics-informed machine learning.
PNO	Praseodymium nickelate.
ProbSparse	Probabilistic sparse attention (Informer).
RBF	Radial basis function.
RC	Resistor-capacitor.
ReLU	Rectified linear unit.

RF	Random forest.
RMSE	Root mean square error.
RNN	Recurrent neural network.
ROS	Random oversampling.
RTW	Random time warping.
RUL	Remaining useful life.
SCCM	Standard cubic centimeters per minute.
SEM	Scanning electron microscopy.
Seq2Seq	Sequence-to-sequence.
SMOTE	Synthetic minority over-sampling technique.
SOFC	Solid oxide fuel cell.
SOH	State of health.
SVM	Support vector machine.
SVR	Support vector regressor.
T-GCN	Temporal graph convolutional network.
Tanh	Hyperbolic tangent function.
TCN	Temporal convolutional network.
TEM	Transmission electron microscopy.
TL	Transfer learning.
TPB	Triple phase boundary.
XGB	Extreme gradient boosting (XGBoost).
YSZ	Yttria-stabilized zirconia.

List of Symbols

α	Random time warping strength
ΔV_{act} [V]	Activation loss
ΔV_{conc} [V]	Concentration loss
ΔV_{ohm} [V]	Ohmic loss
δ_i	Oversampling difficulty score
$\dot{V}_{Air}$ [SCCM]	Air flow rate
$\dot{V}_{H_2}$ [SCCM]	Hydrogen flow rate
$\dot{V}_{N_2}$ [SCCM]	Nitrogen flow rate
ϵ_a	Anode porosity
ϵ_c	Cathode porosity
ϵ_i	Random noise term
$\gamma(\tau)$ [$\Omega \cdot \text{cm}^2$]	DRT distribution function
$\hat{y}$ [V]	Predicted output voltage
$\mathcal{N}(u)$	Neighborhood of node u
$\mathcal{V}$	Set of nodes
ω [rad/s]	Angular frequency
ϕ [rad]	Voltage phase angle
ψ [rad]	Current phase angle
σ	Activation function
τ	Tortuosity
τ_{life} [h]	Cell lifetime

a_{ij}	Adjacency matrix element
b	Bias term
d	Model dimension
d^2	Mahalanobis distance
d_k	Attention key/query dimension
$d_{grain,a}$ [μm]	Anode grain size
$d_{grain,c}$ [μm]	Cathode grain size
f [Hz]	Frequency
f_A	Aggregation function
f_t	Forget gate
f_U	Update function
h_t	Hidden state at time step t
h_u	Hidden representation of node u
i	Sample index
I [A]	Current
i_t	Input gate
j	Imaginary unit
J [A/cm ²]	Current density
k	Message-passing iteration index
L_a [μm]	Anode thickness
L_c [μm]	Cathode thickness
L_e [μm]	Electrolyte thickness
L_{afl} [μm]	Anode functional layer thickness
$m_{\mathcal{N}(u)}$	Aggregated neighborhood message
n	Number of samples or nodes
P_{max} [W/cm ²]	Maximum power density

R_0 [$\Omega \cdot \text{cm}^2$]	Ohmic resistance
$r_{pore,a}$ [μm]	Anode pore size
$r_{pore,c}$ [μm]	Cathode pore size
T_{cell} [$^{\circ}\text{C}$]	Cell temperature
$t_{H_2\ cutoff}$ [min]	Hydrogen cutoff duration
u	Node index
V [V]	Cell voltage
V_0 [V]	Open-circuit voltage
W_{neigh}	Neighbor weight matrix
W_{self}	Self-loop weight matrix
x_u	Input feature vector of node u
$X_{ave,a}$ [μm]	Anode grain contact size
$X_{ave,c}$ [μm]	Cathode grain contact size
y [V]	Measured output voltage
Z [$\Omega \cdot \text{cm}^2$]	Total impedance
z_u	Final node embedding
Z_{pol} [$\Omega \cdot \text{cm}^2$]	Polarization impedance

Chapter 1

Background and Introduction

1.1 Fundamental of SOFC

A fuel cell is a type of electrochemical conversion device that transforms the chemical energy from a fuel such as hydrogen into electricity through a chemical reaction with an oxidizing agent such as air [8]. Solid oxide fuel cell or SOFC is a common type of fuel cell with low pollutant emissions and high thermal efficiency that can operate with different fuels at 600-800°C. SOFCs are used in many applications such as combined heat and power units, transportation, stationary applications, and others [9, 10].

The SOFC structure includes a dense electrolyte sandwiched between a porous anode and cathode, with an external circuit linking the anode and cathode for current collection. Within the cathode, oxygen gas is converted into oxygen ions, O^{2-} , through a reaction with electrons released from the anode. This process is known as oxygen reduction. The oxygen ions then pass through the electrolyte to reach the anode, where they react with hydrogen to produce steam and electrons. A schematic of tubular SOFC and reactions is shown in Figure 1.1. The SOFC operating parameters are hydrogen ($\dot{V}_{H_2}$) and air ($\dot{V}_a$) flow rates, cell temperature (T_c), current density (J), and voltage (V). Some of the SOFC design parameters are the anode (L_a), cathode (L_c), anode functional layer (L_{afl}), and electrolyte thicknesses (L_e), and anode porosity (ϵ_a).

Operating SOFCs in galvanostatic mode, allows the control of the current

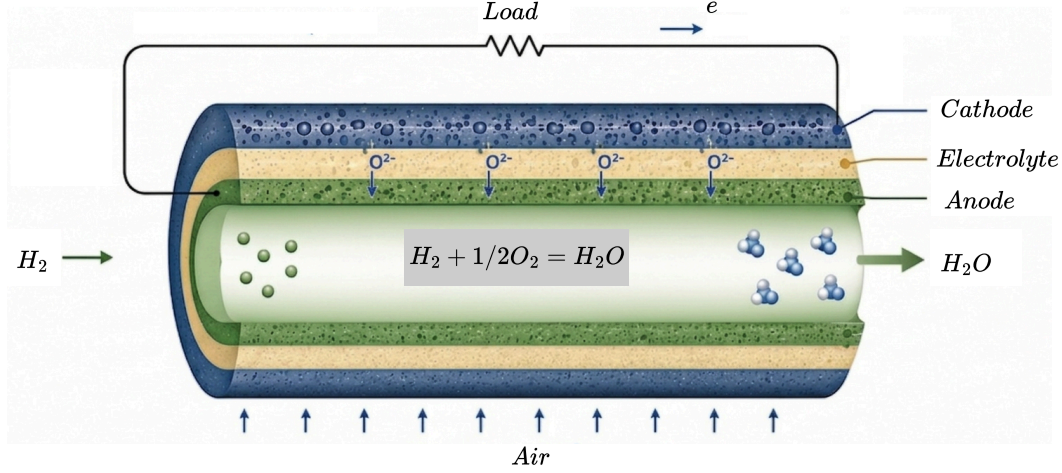


Figure 1.1: Schematic of an anode-supported tubular SOFC illustrating the cell structure, electron flow through the external circuit, oxygen ion (O^{2-}) transport through the electrolyte, and the electrochemical reactions occurring at the electrodes.

load applied to the cell, and the output voltage of the SOFC can be measured, providing insights into the cell's performance [11]. When the electrical current of the fuel cell is equal to zero, the output voltage is a maximum value called open-circuit voltage (OCV) and defined as V_0 . During operation, the actual voltage of the cell is less than this amount due to the overpotential polarizations and internal resistances. There are three types of overpotential polarizations: Ohmic (ΔV_{ohm}), activation (ΔV_{act}), and concentration (ΔV_{cons}) polarization, which are related to electrical conduction, electrochemical reaction at the electrode/electrolyte interface, and gas diffusion respectively [12]. The cell voltage, V , is:

$$V = V_0 - (\Delta V_{act} + \Delta V_{ohm} + \Delta V_{conc}) \quad (1.1)$$

While activation polarization is the dominant resistance in electrodes, depending on their microstructure, concentration polarization strongly affects thick electrode performance, such as in some anode-supported SOFCs [13]. Particularly, at high-temperature operation (more than 700°C), concentration polarization is crucial because at these temperatures, ohmic and activation polarizations will decrease significantly, while concentration polarization will

be constant or will slightly increase [13, 14].

Voltage in SOFC is influenced by a combination of factors, including cell temperature, hydrogen flow rate, air flow rate, and current density. These factors are interdependent and contribute to the overall performance of the cell as [15]:

- **Cell Temperature:** Higher temperatures generally increase reaction rates, improving voltage and efficiency, but excessive temperatures can lead to material degradation, reducing performance over time.
- **Hydrogen Flow Rate:** Adequate hydrogen flow ensures that the electrochemical reaction is supplied with sufficient fuel. A drop in hydrogen flow can lead to fuel starvation, reducing voltage, while excessive flow can introduce inefficiencies.
- **Air Flow Rate:** Oxygen supply through air flow is critical for the cathode reaction. Inadequate air flow can result in oxygen starvation, causing the voltage to drop, while too much air flow can lead to cooling effects and reduced efficiency.
- **Current Density:** As current density increases, voltage tends to decrease due to resistive and polarization losses. This relationship follows the typical J-V curve for fuel cells, where higher currents result in greater losses.

In addition to these primary factors, voltage is also influenced by fuel composition, system pressure, and long-term cell degradation [16, 17].

The main commercialization challenges for SOFC in future energy systems include reducing degradation, identification and mitigation of degradation, and consequently increasing lifespan [18]. Therefore, predicting degradation performance is crucial for adjusting the performance of SOFC systems to minimize degradation. Implementing an effective health control strategy is the foundation for predictive maintenance. This allows properly designed correction actions to prevent the destruction of SOFC systems, and reduces cost.

Figure 1.2 illustrates the general workflow of SOFC health management considered in this work. The process begins with data acquisition using sensor and electrochemical measurements such as EIS. The collected signals are then used for anomaly and fault detection to determine whether the system is operating under normal or abnormal conditions. After detecting abnormal behavior, fault classification identifies the type of fault during the early stage of operation. The next stage, degradation mode diagnostics, focuses on identifying the dominant degradation mechanisms or degrading components associated with the observed behavior. Prognostics then estimates the future degradation trajectory, including SOH evolution and RUL prediction, which ultimately supports health management decisions such as maintenance scheduling, operational adjustment, and resource planning.

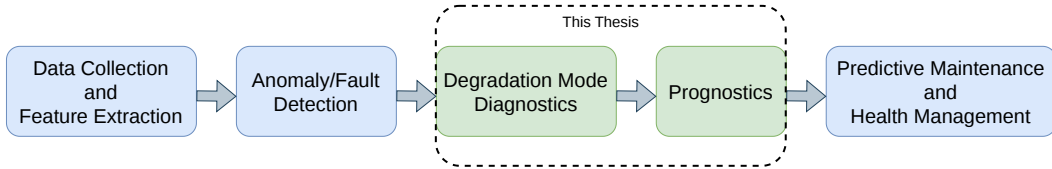


Figure 1.2: General workflow of SOFC health management, including data acquisition, anomaly/fault detection, degradation mode diagnostics, prognostics, and health management. The highlighted section indicates the primary focus of this thesis.

It is therefore essential to (1) identify suitable degradation prediction methods, (2) select accurate and interpretable diagnostic tools, and (3) understand lifetime issues and dominant degradation modes. In the following sections, the literature is reviewed and discussed in three categories: (1) degradation modes and lifetime issues, (2) SOH diagnostic tools, and (3) degradation prediction and classification methods which is shown in Figure 1.3.

1.2 Literature Review

1.2.1 Degradation Modes

Failures occurring within SOFCs can generally be classified into two main categories: (i) physical failures that directly affect cell performance, such as

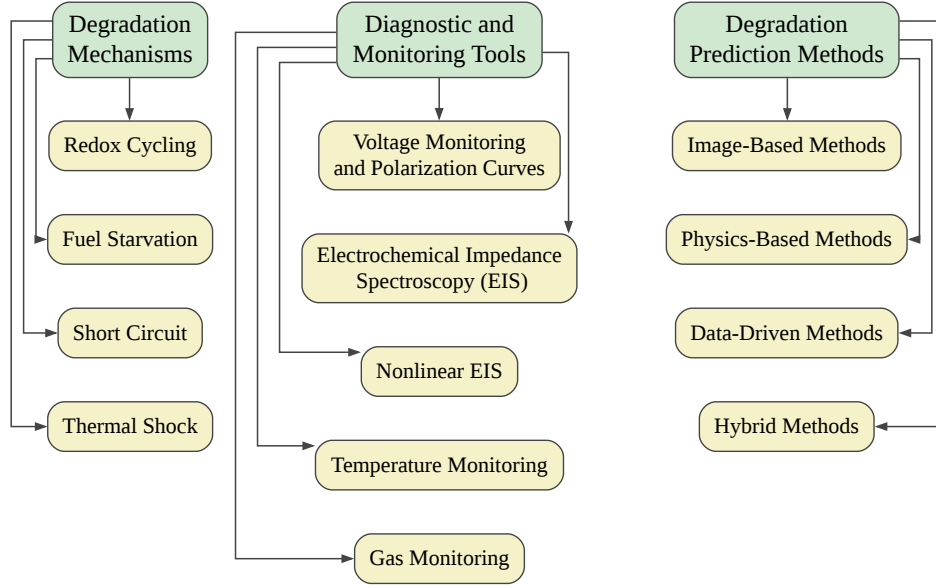


Figure 1.3: Overview of major SOFC health monitoring research directions, including degradation mechanisms, diagnostic and monitoring tools, and degradation prediction methodologies discussed in this thesis.

increases in internal resistance and electrochemical losses, and (ii) morphological failures that lead to structural damage and irreversible changes in the cell microstructure. Morphological degradation ultimately manifests as physical performance deterioration that can be observed through electrochemical measurements [19, 20].

The occurrence and progression of these failures are influenced by two primary factors: (i) the materials used in the cell components and (ii) the operating conditions. Material-related degradation typically develops gradually over time and is largely irreversible. Mitigating this type of degradation generally requires the development of improved materials or modifications in cell design and fabrication [19, 20]. In contrast, operating conditions can induce degradation mechanisms that may be partially or fully reversible if appropriate corrective actions are taken.

In this dissertation, the primary focus is on failures associated with operational conditions, which can lead to reversible or partially reversible degradation ef-

fects. Understanding and identifying these operational failures is essential for enabling effective monitoring, control, and mitigation strategies during SOFC operation. To strategically prioritize degradation mechanisms for experimental investigation, it is essential to assess their impact on SOFC performance (severity rating), the probability of their occurrence (occurrence rating), and the feasibility of implementing the mitigation strategies (feasibility rating) [21]. Based on comprehensive literature review and insights from extensive field and cell testing, four critical degradation modes have been identified as priorities: (1) redox-related degradation caused by fuel interruptions [22], (2) fuel starvation due to insufficient fuel flow or excessive current demand [21], (3) internal short-circuit failures [23], and (4) degradation driven by thermal gradients and thermal transients [24].

Redox Cycling. SOFC anodes are generally produced by sintering a mixture of ceramic powders, which are then subjected to a reduction process to create a cermet material. This process typically involves combining a ceramic material with high oxygen ion conductivity, such as yttria-stabilized zirconia (YSZ) or samaria-doped ceria, with a metal oxide selected to obtain the desired metallic component after reduction [25]. Nickel is the most extensively employed metallic component due to good electronic conductivity, high catalytic activity, and remarkable electrochemical activity for the reaction of oxide ions with the fuel gas at triple-phase boundary (TPB) [26]. When fuel is depleted during operation, nickel may undergo reoxidation, leading to a substantial increase in volume—about 69%—as it reverts from bulk nickel to nickel oxide form [27]. This reduction and oxidation process is known as a Redox cycle. Since nickel oxide must be reduced back to nickel metal before operation, and this Redox process recurs when switching between different modes like power generation and standby, or during interruptions in fuel supply [28], multiple redox cycles can happen during the operational life of a SOFC. These cycles cause significant degradation over time, leading to delamination and mechanical damages to the fuel cell (Figure 1.4), making Redox cycle degradation a

key focus of research in SOFCs.

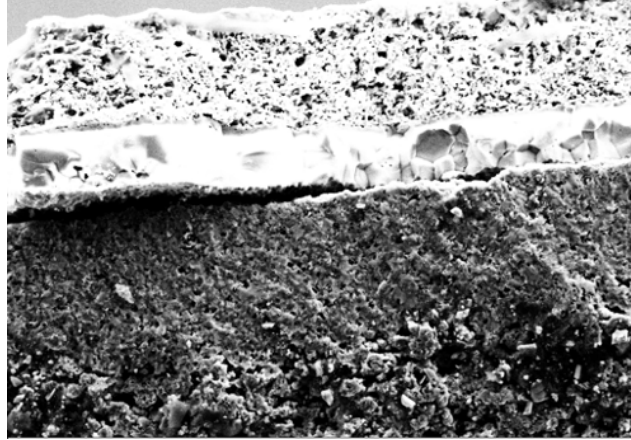


Figure 1.4: Delamination between the electrolyte and anode functional layer as a result of redox cycling

Fuel Starvation. Fuel starvation occurs when the fuel supply to the anode becomes insufficient relative to the electrochemical load imposed on the cell. Under such conditions, the available hydrogen or fuel species cannot sustain the electrochemical reactions occurring at the TPB. As a result, the cell potential locally drops and oxygen ions arriving at the anode may react with the nickel catalyst instead of the fuel, leading to partial oxidation of the nickel phase and degradation of the anode microstructure [29]. In addition, fuel starvation can lead to the formation of large local overpotentials, which may trigger reverse currents and cause severe localized heating within the cell [21]. This phenomenon can result from several operational issues, including insufficient fuel flow rate, uneven fuel distribution within stacks, or rapid load changes. If prolonged, fuel starvation may cause irreversible damage such as nickel oxidation, electrode delamination, and accelerated degradation of electrochemical performance [30]. Consequently, detecting and mitigating fuel starvation is critical for maintaining stable SOFC operation.

Internal Short Circuit. Internal short circuits in SOFCs occur when an unintended electronic conduction path forms between the anode and cathode through the electrolyte or other cell components. Ideally, the electrolyte should

be purely ionic conducting and electronically insulating; however, defects such as cracks, pinholes, or manufacturing imperfections may allow electrons to pass directly between the electrodes [23]. In addition, metallic particle migration, contamination, or excessive local heating can also create conductive bridges within the electrolyte layer [31]. The presence of an internal short circuit reduces the open-circuit voltage and leads to parasitic currents that bypass the external load, thereby reducing the overall efficiency of the cell. These leakage currents can also produce localized Joule heating, which may further aggravate material degradation and accelerate structural damage in the electrolyte or electrodes [32]. As a result, internal short circuits represent a significant reliability concern, particularly in long-term operation of SOFC stacks.

Thermal Shock. Thermal shock refers to the mechanical stress and potential structural damage that occur when a SOFC experiences rapid temperature changes. Due to the multilayer ceramic structure of SOFCs, different components such as the electrolyte, electrodes, and interconnects possess distinct thermal expansion coefficients. Rapid heating or cooling can therefore induce large thermal gradients and mechanical stresses within the cell structure [24]. These stresses may lead to cracking of the electrolyte, delamination of electrode layers, or degradation of sealing materials [33]. Thermal shock events may occur during startup and shutdown cycles, abrupt load changes, or failures in thermal management systems. Repeated exposure to such temperature fluctuations can weaken the structural integrity of the cell and significantly reduce its operational lifetime [34]. Consequently, careful thermal management and controlled heating and cooling procedures are essential for maintaining the mechanical stability and long-term durability of SOFC systems.

1.2.2 Diagnostic and Monitoring Tools

In this section, an overview of various diagnostic tools is presented that facilitate the early detection of degradation, thereby helping to mitigate irreversible deterioration. Below, we provide a detailed description and discussion of the

most relevant conventional and non-conventional tools for assessing the SOH based on current and voltage measurements. These tools possess several desirable characteristics: (i) they are applicable in both laboratory and real-world operating environments, (ii) they are time- and cost-efficient, and (iii) they can be readily implemented in different systems. The specific tools discussed in this section are shown in Figure 1.5 and include polarization curves, electrochemical impedance spectroscopy, the advanced methodology of distribution of relaxation times, and total harmonic distortion analysis.

Particular emphasis is placed on understanding electrochemical impedance spectroscopy (EIS), as it serves as a fundamental framework for many subsequent SOH monitoring techniques. Additionally, its advanced methodology, the distribution of relaxation times (DRT), is explained in detail, as it enables the identification of distinct frequency domains associated with the various electrochemical processes occurring within SOFC systems.

Voltage Monitoring

The investigation of electrochemical systems typically begins with conventional monitoring tools such as voltage measurements. Voltage monitoring provides insight into the performance of the cell at a specific operating point. However, to obtain information about SOFC behavior across a range of operating conditions, polarization curve measurements are commonly employed. The polarization curve is one of the most widely used techniques for the electrochemical characterization of fuel cells and electrolysis cells.

Polarization measurements can be performed using two different approaches: (i) a step-by-step method in which individual operating points are selected sequentially, or (ii) a continuous adjustment of the imposed current or voltage value. The selection of the testing speed represents a compromise between experimental duration and the requirement to obtain meaningful and reliable results. Two distinct methods for obtaining current density profiles are described in [35].

In the first approach, the current density is gradually increased from OCV

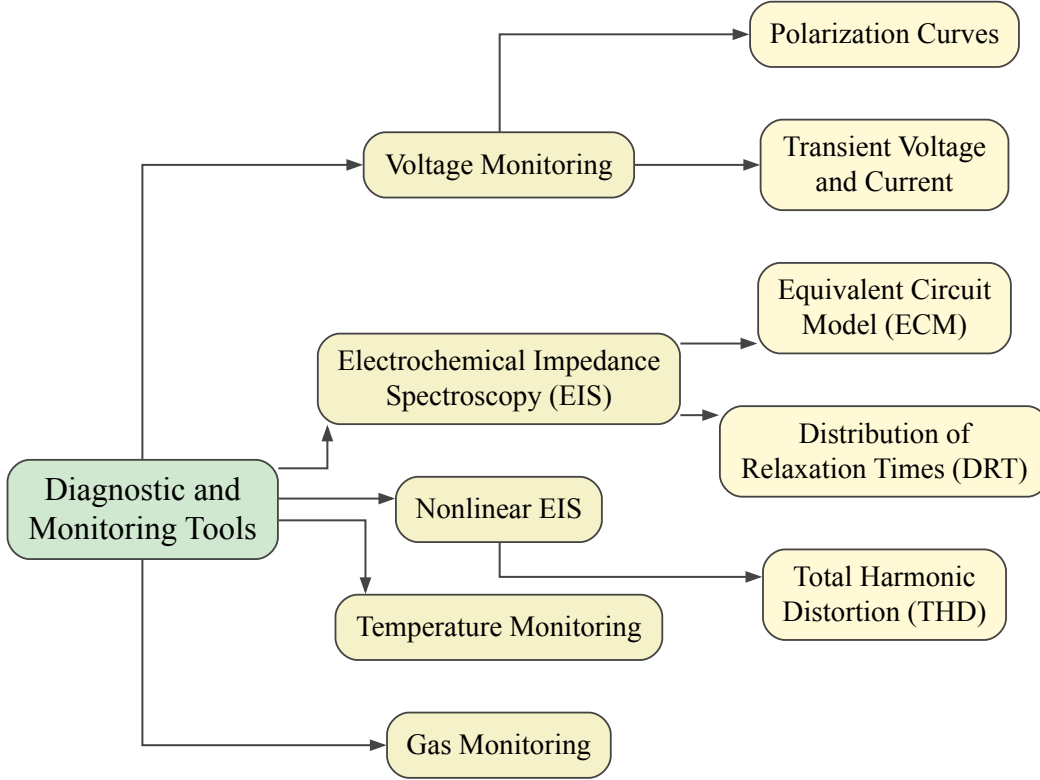


Figure 1.5: Overview of diagnostic and monitoring tools used for SOFC health assessment, including voltage-based monitoring, electrochemical impedance spectroscopy (EIS), nonlinear EIS techniques, and gas and temperature monitoring methods.

to the maximum current density corresponding to approximately 0.7 V in discrete steps. At each current density level, the system is allowed to stabilize and measurements are typically performed for approximately 10 minutes to ensure thermochemical equilibrium. In contrast, the second method utilizes a rapid hold-time acquisition technique. This approach aims to prevent significant temperature fluctuations by increasing the current density rapidly, typically within 10 seconds. Detailed procedures for properly conducting polarization curve measurements are available in technical reports, particularly those described in [36, 37].

The characteristic shape of polarization curves provides preliminary insights into the different loss mechanisms occurring within the cell. Nevertheless, for a more precise understanding of the distribution of losses, their physical origins,

and the identification of potential early-stage degradation modes, more advanced measurement techniques such as linear and nonlinear EIS are required.

Electrochemical Impedance Spectroscopy (EIS)

Electrochemical impedance spectroscopy (EIS) is a well-established non-destructive technique used to analyze the processes and mechanisms occurring within electrochemical systems [38]. This technique is also commonly referred to as impedance spectroscopy or AC impedance spectroscopy. In EIS measurements, a sinusoidal perturbation in current or voltage is applied to a system operating under steady DC conditions, and the corresponding system response is measured, as shown in Figure 1.6

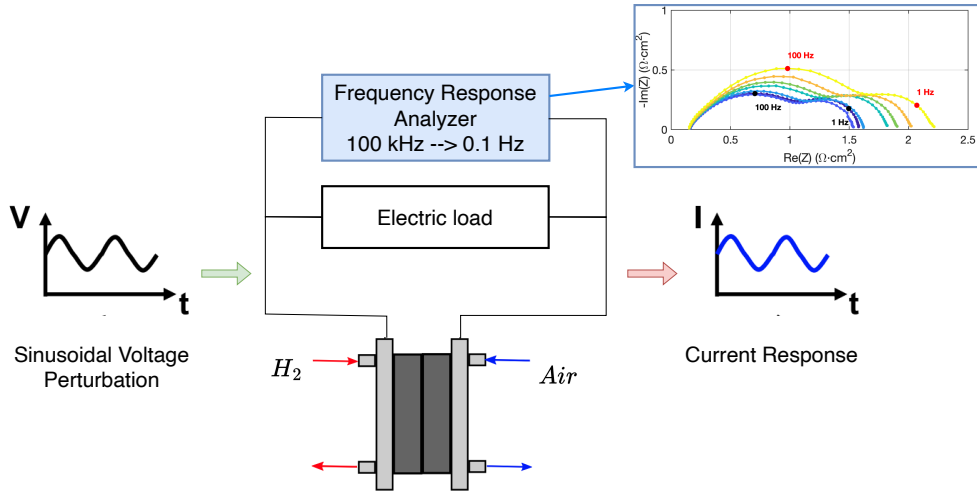


Figure 1.6: Schematic illustration of electrochemical impedance spectroscopy (EIS) measurements for SOFC characterization. A small-amplitude sinusoidal current perturbation is applied through the external circuit over a range of excitation frequencies, and the resulting voltage response is measured using a frequency response analyzer (FRA). The measured phase-shifted voltage and current signals are used to calculate the complex impedance and construct the Nyquist impedance spectrum.

Typically, these measurements are conducted over a frequency range between 1 MHz and 100 mHz. The exact frequency range can be adjusted depending on the processes that are expected to be analyzed. For example, if very slow electrochemical processes are of interest, the lower frequency limit may be extended to approximately 1 mHz. Conversely, the upper frequency limit

may be restricted by inductive effects present in the measurement system, which can significantly reduce the maximum usable frequency. Ultimately, the impedance is calculated as the vector ratio of voltage to current over the specified frequency range.

$$Z(j\omega) = \frac{u(t)}{i(t)} = \frac{U}{I} \exp(j(\phi - \psi))$$

Depending on the system configuration, the measurements can be conducted in either (i) a pseudo-potentiostatic mode, where the voltage is controlled and the current response is measured, or (ii) a pseudo-galvanostatic mode, where the current is controlled and the voltage response is measured. In the case of SOFC systems, galvanostatic mode is typically employed.

When the system is operated in potentiostatic mode, the applied voltage can be expressed as $u = u_{DC} + u_{AC}$, where u_{DC} represents the DC component and u_{AC} denotes the AC perturbation signal. The resulting current response can then be expressed as a Fourier series in the following form [39].

$$i(t) = i_0 + i_1 \sin(\omega t + \psi_1) + i_2 \sin(2\omega t + \psi_2) + \dots + i_n \sin(n\omega t + \psi_n)$$

Alternatively, when the system is operated in galvanostatic mode, the applied current can be written as $i = i_{DC} + i_{AC}$. The resulting voltage response of the system can then be obtained using the following expression [39]. An example of EIS curve of SOFC can be seen in Figure 1.7.

$$u(t) = u_0 + u_1 \sin(\omega t + \xi_1) + u_2 \sin(2\omega t + \xi_2) + \dots + u_n \sin(n\omega t + \xi_n)$$

To ensure accurate and reliable impedance measurements, several requirements must be satisfied. These include steady-state conditions, causality, time-invariance, and linearity [40]. Prior to conducting impedance measurements, it is essential to confirm that both electrodes have reached stable operating conditions. This stability is achieved when the system reaches thermochemical equilibrium. If the measured response signal follows a sinusoidal waveform and corresponds directly to the applied input signal, the system is considered causal. Furthermore, the system function of the electrochemical system under

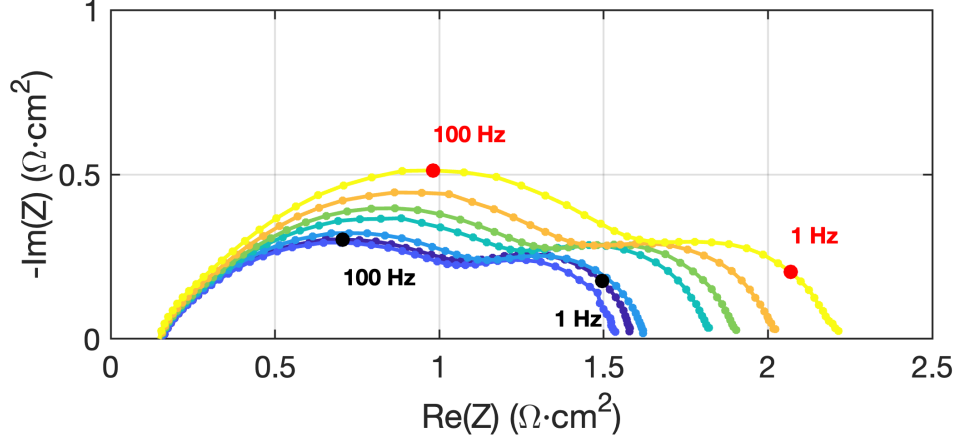


Figure 1.7: EIS curves of SOFC. The intersection of the curve with real axis represents Ohmic resistance, and the distance from that point to the end of the curve is Polarization resistance. The summation of the Ohmic and Polarization resistances is the total resistance of the cell

investigation is only indirectly influenced by the time domain, which indicates time-invariant behavior. Finally, the linearity condition must be satisfied. To maintain linearity, the amplitude of the perturbation signal must remain sufficiently small so that the system response remains within a pseudo-linear regime [41–43].

However, if the perturbation amplitude is excessively small, the signal-to-noise ratio may become poor. Therefore, it is necessary to select a sufficiently large amplitude to reduce the influence of noise while still preserving the linear behavior of the system [44]. If the linearity condition is violated and nonlinear behavior appears, nonlinear analysis methods must be applied to interpret the measured data.

Both current and voltage signals can be used as perturbation inputs in EIS measurements. Various current amplitudes have been used in previous studies, including 20 mA [45], 10–40 mA [42], 60 mA [46–49], and even 100 mA [50]. In more recent studies, the AC perturbation amplitude has been adjusted proportionally to the DC load, commonly at a value of approximately 4% of the DC current, while ensuring that the linearity condition remains satisfied [51, 52]. Similarly, several voltage perturbation amplitudes have been proposed in the literature, such as 10 mV [53, 54], 12 mV [55], 14 mV [54], and 50 mV [56].

In addition to the conventional approach of applying a constant amplitude perturbation during a measurement, the use of a random binary signal (RBS) has been proposed as an efficient technique for impedance measurements [57, 58]. This method relies on random variations in the perturbation amplitude at different time intervals, which are then analyzed across different frequencies. Another alternative approach involves the use of multi-sine excitation signals, which enable the characterization of the cell within a significantly shorter time frame [59].

Equivalent Circuit Model (ECM). To simplify the interpretation of impedance spectra, it is common practice to employ equivalent circuit models (ECMs) that fit the measured impedance data. In ECMs, various electrical components are interconnected to represent the behavior of the electrochemical system under investigation. These components are intended to represent physical and chemical processes occurring within the system, such as electrochemical reactions, electron transfer, mass transport, and electrolyte resistance [44]. The elements used in ECMs can generally be categorized into two groups: (i) lumped electrical elements, including resistors (R), capacitors (C), and inductors (L), which represent ideal electrical behavior, and (ii) distributed or frequency-dependent elements such as the constant phase element (CPE), Warburg diffusion element (W), Gerischer element (G), and other bounded diffusion elements that account for non-ideal electrochemical processes occurring within the system [44].

Over time, numerous equivalent circuit models have been developed to analyze the performance of SOFC systems and to improve the understanding of the underlying electrochemical processes. The most common approach used to determine the parameters of these models is the complex nonlinear least squares (CNLS) fitting method.

The equivalent circuit proposed in [60], shown in Figure 1.8, consists of five sequential impedance elements: (i) an ohmic resistor (R_0) representing ohmic losses, (ii) two serial RQ elements (R_{2A} and R_{3A}) corresponding to gas diffu-

sion and activation polarization in the anode functional layer, (iii) a Gerischer element (R_{2C}) representing cathode activation polarization, (iv) a generalized finite-length Warburg element (G-FWS) (R_{1A}) representing gas diffusion polarization within the anode substrate, and (v) an additional RQ element (R_{1C}) that becomes significant only at very low cathode oxygen partial pressures ($p(O_{2, cat}) < 0.05$ atm). In another study by the same research group [61], an identical equivalent circuit was employed. However, in that work the final RQ element was associated with processes occurring during reformat operation.

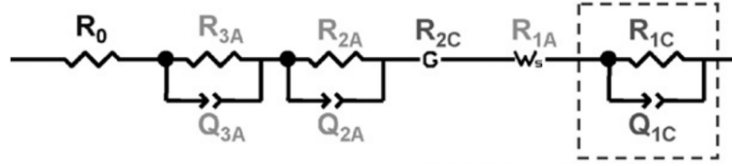


Figure 1.8: ECM used for the CNLS fit of the impedance spectra of planar cell [60].

In their study, Sumi et al. [62] employed an alternative equivalent circuit model, illustrated in Figure 2, which primarily consists of R and Q (CPE) elements. In this model, the inductive element L represents the high-frequency noise caused by the measurement lead wires, while the R_0 element corresponds to the ohmic resistance of the cell. The remaining RQ elements represent electrochemical processes occurring within tubular and microtubular SOFCs. Specifically, RQ-1 and RQ-4 are associated with the characterization of charge transfer (activation) and gas diffusion (concentration) processes occurring at the cathode. The elements RQ-2 and RQ-3 partially overlap and are attributed to the activation polarization processes occurring at the anode. Finally, the fifth RQ element, RQ-5, represents concentration losses within the anode. Notably, the same equivalent circuit model used by Sumi et al. [62] was also adopted by Nielsen and Mogensen [63] to investigate the degradation of LSM:YSZ air electrodes induced by steam exposure.

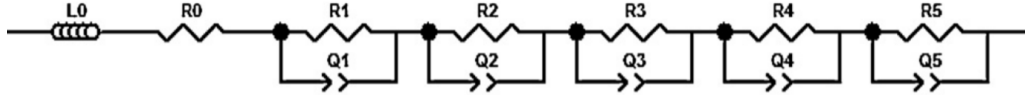


Figure 1.9: ECM used for the CNLS fit of the impedance spectra of tubular and microtubular cell [62].

Distribution of Relaxation Times (DRT). The Distribution of Relaxation Times (DRT) method provides researchers with a powerful analytical tool for extracting the most relevant time constants from measured EIS data and associating them with specific electrochemical mechanisms. The fundamental assumption of the DRT approach is that each frequency point in the impedance spectrum corresponds to a particular physical or electrochemical process occurring within the system. In most cases, these processes are represented by a series of RC elements or, more recently, RQ elements that approximate the system response. An optimization procedure is then used to determine the minimum number of elements required to accurately represent the processes occurring within an operating SOFC system [64].

One of the main advantages of the DRT approach is that it reduces user subjectivity compared to equivalent circuit modeling. Unlike ECM fitting, which often requires assumptions regarding the structure of the circuit model, DRT analysis provides a more objective representation of the underlying electrochemical processes [65]. For this reason, DRT analysis can also serve as an initial step in identifying an appropriate equivalent circuit model, thereby offering a more systematic approach to interpreting impedance spectra.

In the work reported in [66], the authors discuss two different approaches for interpreting impedance spectra when comparing ECM and DRT methodologies. In the first approach, a simple equivalent circuit composed of several RC elements is used as the basis for performing DRT analysis without requiring prior knowledge of the system's physical behavior. In the second approach, DRT analysis is combined with experimentally measured impedance spectra to guide the development of an appropriate equivalent circuit model capable

of representing complex electrochemical processes.

In summary, the DRT method provides valuable insights into physical processes that may not be directly observable through experimental measurements alone. Compared with conventional impedance analysis methods, DRT enables a more detailed interpretation of electrochemical mechanisms and more effectively captures the complexity of the processes occurring within SOFC systems. Consequently, DRT is increasingly recognized as an important analytical technique in impedance spectroscopy.

The theoretical foundation of the DRT method was introduced in 1992 by Jürgen Weese, who developed a numerical framework for analyzing electrical systems using this approach. This methodology was implemented in a program known as FTIKREG [67]. The Fredholm integral equation of the first kind represents an ill-posed problem, meaning that the solution may not be unique or may be highly sensitive to measurement noise. As a result, special numerical techniques must be employed to obtain stable solutions. The code developed by Weese [67] has been made available to the research community and highlights the complexity involved in computing accurate DRT spectra. The successful application of this method requires careful selection of several numerical parameters that must remain consistent with the physical characteristics of the electrochemical system. Among these parameters, the regularization parameter λ plays a particularly important role.

To obtain reliable solutions, appropriate regularization techniques must be applied, with Tikhonov regularization being one of the most commonly used approaches [67]. Regularization methods serve two main purposes: they help stabilize ill-posed problems and prevent overfitting by incorporating additional constraints into the optimization procedure [68]. The selection of an appropriate regularization parameter not only smooths the noisy measurement data but also strongly influences the accuracy of the resulting DRT spectra.

In addition to the original Tikhonov regularization method proposed in [67], several alternative regularization approaches have been investigated in the literature [69]. The research group led by Ciucci has been actively studying

different regularization and discretization techniques aimed at improving the interpretation of DRT deconvolution results. Their work has demonstrated that alternative frameworks, including Bayesian statistical approaches [70] and Lasso regularization [68], can also be effectively applied as alternatives to classical Tikhonov regularization. Furthermore, this research group has developed a MATLAB-based software package called *DRTtools*, which enables researchers to compute DRT spectra directly from impedance data.

To gain a deeper understanding of the individual electrochemical processes occurring in SOFC systems, experimental investigations often begin at the level of individual components such as electrodes or electrolytes. These studies may then progress to symmetric cells before finally examining complete single cells or stacks. For example, in the work presented in [66], the authors investigated the influence of impurity concentrations in polycrystalline proton-conducting electrolytes. Using DRT analysis, they successfully distinguished between bulk resistance and grain boundary resistance. Their results showed a significant increase in the DRT area corresponding to relaxation times below 1 second, indicating inhibited ion transport due to the formation of silica-based phases at grain boundaries.

Similarly, DRT analysis has been applied to LSCF-based symmetrical cells to identify several electrochemical processes, including interfacial reactions and diffusion-related resistances [71]. In another study, [72] investigated the use of SFCM as an air electrode material for proton-conducting SOFCs. Through DRT analysis, four distinct peaks were observed. The low-frequency peak below 1 Hz was attributed to concentration polarization caused by gas diffusion. A peak between 10 and 100 Hz was associated with oxygen reduction and surface diffusion toward the triple-phase boundary. The peak observed between 100 Hz and 1 kHz was related to oxygen incorporation into the lattice structure. Finally, peaks in the high-frequency range between 1 kHz and 10 kHz were attributed to overlapping processes such as coupled charge transfer, ionic transport through the air electrode, and electrical contact resistance between the electrolyte and grain boundaries. These observations are consistent with

several previous studies [73–76].

In another investigation, Soltanizade et al. [77] demonstrated that DRT analysis could successfully identify the specific influence of gas diffusion limitations that hinder the oxygen reduction reaction. Importantly, these effects could not be clearly distinguished using conventional EIS analysis alone.

Clematis et al. [78] also emphasized the importance of DRT analysis in studying the catalytic activity of BSCF- and LSCF-based air electrodes. By analyzing EIS data through DRT methods, they identified two distinct kinetic regimes associated with surface oxygen exchange reactions and bulk diffusion processes. These regimes were not clearly distinguishable using conventional impedance analysis.

In their study, Sumi et al. [22] performed a detailed DRT analysis of micro-tubular SOFCs composed of Ni-GDC/YSZ/GDC/LSCF materials to identify key electrode processes and detect degradation mechanisms. Their results identified five important frequency ranges that should be monitored to better understand both fuel electrode and air electrode processes. These frequency ranges were categorized as follows: (i) air electrode processes occurring between 5–20 Hz and 10–100 kHz, and (ii) fuel electrode processes occurring within 20 Hz–10 kHz, 1–10 Hz, and below 1 Hz.

Additional results reported by the same research group in [62] indicated that peaks within the frequency ranges of 0.37–5 kHz and 4.8–7.8 Hz were influenced by fuel electrode sintering processes and hydrogen partial pressure. Furthermore, variations in peaks between 10 Hz and 10 kHz were associated with the sintering temperature of the air electrode [79]. Gas diffusion processes occurring within the fuel electrode substrate were observed at lower frequencies ranging from approximately 0.5 Hz to 15 Hz [61]. The exact frequency ranges depended strongly on the type of fuel being used. For example, diffusion peaks associated with CO/CO₂ fuels were observed between 0.5 and 1 Hz, whereas diffusion processes associated with H₂/H₂O fuels occurred approximately ten times faster and therefore appeared within the frequency range of 5–10 Hz.

In [80], the authors summarized findings from several previous studies [81–84]

for fuel-electrode-supported SOFCs and identified six distinct electrochemical processes: (i) P1 ($f < 1$ Hz), representing fuel electrode diffusion for re-formate fuels and adsorption/desorption processes at the air electrode; (ii) P2 ($f = 1 - 10$ Hz), associated with gas conversion processes; (iii) P3 ($f = 10 - 100$ Hz) and P4 ($f = 100 - 1000$ Hz), corresponding to oxygen transport-related processes in the fuel electrode; (iv) P5 ($f = 1 - 10$ kHz), representing charge transfer processes in the fuel electrode; and (v) P6 ($f > 10$ kHz), which currently has no clearly identified physical interpretation.

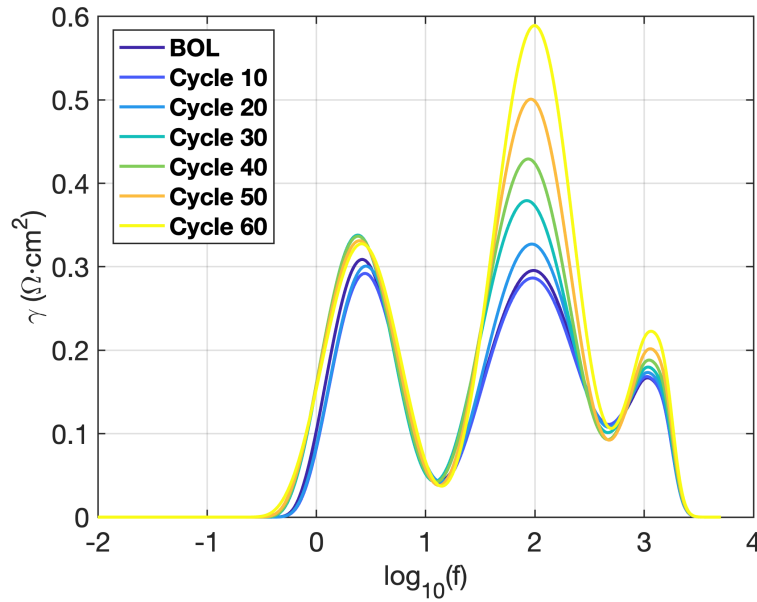


Figure 1.10: DRT curves for a SOFC under redox cycling. The evolution of DRT peaks illustrates the progression of electrochemical processes and degradation over time.

Non-Linear Electrochemical Impedance Spectroscopy (NLEIS)

Under certain operating conditions, the assumption of system linearity may no longer hold, and the electrochemical system may exhibit nonlinear behavior. Two primary factors can lead to the emergence of such nonlinear responses: (i) the application of perturbation signals with excessively large amplitudes and (ii) specific electrochemical processes that induce state-of-charge (SOC) deactivation or morphological changes in the electrodes, which subsequently

distort the system response.

When nonlinear behavior occurs, conventional linear impedance analysis becomes insufficient for accurately interpreting the system dynamics. In such situations, nonlinear system analysis must be applied to the measured data in order to separate the contributions arising from the fundamental linear response and those originating from higher harmonic components. These approaches provide additional insight into the nonlinear electrochemical processes occurring within the system [85, 86]. Such information is particularly valuable for achieving a comprehensive understanding of the electrochemical behavior of the system, since conventional linear impedance spectroscopy may not always be capable of capturing all relevant mechanisms, especially under certain operating conditions.

Consequently, in situations where nonlinear responses are present, the use of nonlinear electrochemical impedance spectroscopy (NLEIS) or total harmonic distortion (THD) analysis becomes necessary. These techniques enable the identification and analysis of higher-order harmonics in the system response, thereby providing valuable information regarding nonlinear electrochemical processes. Despite the ability of NLEIS and THD to provide insights beyond those obtainable through traditional linear EIS, relatively few studies have employed these approaches for the analysis of SOFC systems to date [87–90].

Gas Analysis and Temperature Monitoring

In addition to the electrochemical diagnostic techniques described previously, several other monitoring methods can be employed to assess the current condition and operating state of SOFC systems. Among these methods, gas composition analysis and temperature monitoring are particularly important, as they can be readily applied to large-scale cells and extended to the monitoring of stacks and complete systems.

When multiple monitoring techniques are used simultaneously, the ability to interpret the ongoing electrochemical reactions and system behavior is significantly enhanced [91, 92]. For example, combining voltage monitoring, elec-

trochemical impedance spectroscopy, gas analysis, and temperature measurements at multiple locations within the system provides highly detailed information about the processes occurring within the fuel cell. This integrated approach enables the identification of a wide range of operational issues and degradation mechanisms.

Several studies have demonstrated that the combination of these monitoring techniques provides a powerful framework for diagnosing operational problems and understanding system performance [48, 93–97]. By integrating information obtained from electrical, thermal, and chemical measurements, it becomes possible to obtain a more comprehensive picture of the system state and to identify deviations from normal operating conditions more effectively.

1.2.3 Degradation Prediction Methods

Predicting degradation in SOFC systems is an advanced monitoring and diagnostic capability aimed at evaluating the condition of critical components and forecasting their performance deterioration. Accurate degradation prediction provides a foundation for effective maintenance planning and helps reduce operational costs by minimizing unplanned maintenance. Consequently, reliable degradation prediction is essential for optimizing SOFC system performance and enabling timely health management strategies that mitigate further performance loss.

Currently, degradation prediction approaches for SOFC systems can generally be categorized into four groups: image-based methods, physics-based methods, data-driven methods, and hybrid methods. Figure 1.11 shows these groups.

Image-based Methods

Image analysis methods provide a robust and reliable approach for predicting performance degradation in SOFC systems. By characterizing microscopic features as well as mechanical and thermal properties, image analysis enables quantitative assessment of degradation mechanisms within the cell. In this approach, internal structural information of the SOFC is first captured using

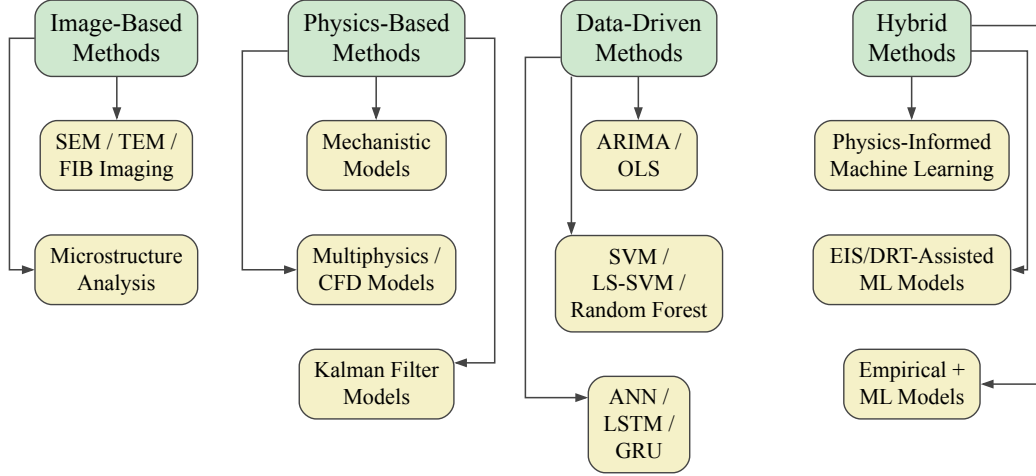


Figure 1.11: Classification of SOFC degradation prediction methods discussed in the literature, including representative image-based, physics-based, data-driven, and hybrid approaches.

imaging techniques such as scanning electron microscopy (SEM), transmission electron microscopy (TEM), and focused ion beam (FIB). The resulting images are then processed and analyzed to extract quantitative microstructural information, including phase fractions, grain sizes, particle size distributions, compositional shape factors, spatial organization of phases, and other descriptive features. These microstructural characteristics are subsequently used to evaluate the degradation behavior of the SOFC system. Compared with model-based and data-driven performance prediction approaches, image analysis offers a direct and effective means of quantifying degradation phenomena [98].

Ananyev et al. [99] proposed a method for detecting degraded regions in SOFCs by analyzing the line autocorrelation function of microstructural images. This technique successfully identified performance degradation associated with chromium poisoning and nickel agglomeration within the SOFC system. Similarly, Simwonis et al. [100] introduced an image analysis technique to investigate the degradation of Ni/YSZ anodes in SOFCs. Their results showed a 33% reduction in the initial electrical conductivity after 4000 hours of exposure to an Ar/4% H₂/3% H₂O atmosphere at 1000 °C. This conductivity loss was attributed primarily to the aggregation of metal particles within

the anode microstructure. Yan et al. [101] developed a cost-effective and efficient image analysis approach for evaluating degradation in SOFC anode microstructures. Their method was capable of covering a larger analysis area than the conventional FIB-SEM technique while reducing the dependency on specialized equipment.

Despite its advantages, the image-based approach has several limitations. One major drawback is the high cost associated with the specialized imaging equipment required for microstructural characterization, which can significantly increase the overall cost of system performance analysis. In addition, these imaging techniques typically require the SOFC system to be disassembled or operated offline in order to observe the internal structure. As a result, image-based methods are generally limited to offline degradation analysis rather than real-time monitoring of SOFC performance.

Physics-based Methods

Physics-based modeling approaches describe SOFC degradation using mathematical formulations that represent the underlying physical and electrochemical processes within the cell. By incorporating material properties, transport phenomena, and electrochemical reactions, these models aim to simulate the evolution of SOFC performance and predict long-term degradation behavior. Such models can capture trends in voltage decay and provide insights into the mechanisms governing performance loss.

Several studies have developed mechanistic or multiphysics models to describe degradation in SOFC systems. For example, Dolenc et al. [102] proposed a nonlinear SOFC stack model with lumped parameters combined with a Kalman filter (KF) to forecast stack degradation. Their results showed that the model could adapt to variations in degradation rates and maintain strong predictive capability even under changing operating conditions. Similarly, Gemmen et al. [103] developed a mechanistic degradation model based on polarization losses in SOFCs. In their work, the area specific resistance (ASR) and degradation rate (DR) were introduced as key parameters for character-

izing stack degradation. Experimental results indicated that ASR remains relatively insensitive to certain variations in operating conditions, making it suitable for comparing design and material changes, whereas DR provides a more appropriate indicator for evaluating long-term efficiency degradation.

A number of studies have focused on modeling degradation associated with microstructural changes in electrodes, particularly particle coarsening. Mason et al. [104] developed a multiphysics model to investigate performance degradation caused by electrode coarsening in SOFCs. Their results demonstrated that higher operating temperatures accelerate degradation when coarsening is the dominant mechanism, and that an optimal operating temperature exists that maximizes the remaining lifetime of the cell. Similar degradation mechanisms have also been investigated through multiphysical modeling frameworks. For instance, Fu et al. [105] developed a multi-physics field model to analyze degradation in direct methane reforming SOFCs caused by Ni particle coarsening. Their model quantified the influence of coarsening on parameters such as Ni particle percolation probability, effective triple-phase boundary (TPB) area, electron conductivity, maximum power density, and degradation rate. In a related study, Zhu et al. [106] employed a transient multiphysics model to analyze the effect of Ni coarsening on SOFC performance. Their findings showed that increased operating temperatures accelerate Ni particle growth and intensify performance degradation, while higher steam-to-carbon ratios further exacerbate coarsening effects. To overcome the high computational cost associated with multiphysics simulations, the authors trained an artificial neural network (ANN) using data generated from the multiphysics model, enabling faster prediction and optimization of SOFC operating conditions.

Other modeling approaches have investigated degradation mechanisms from a system-level perspective. Zaccaria et al. [107] developed a real-time one-dimensional SOFC model capable of predicting the influence of operating parameters on output voltage while incorporating degradation effects. Nakajo et al. [108] proposed a simulation-based framework to analyze lifetime degradation of an SOFC system under different operating strategies, including con-

stant power output and dynamic dispatch operation. Their results indicated that dynamically dispatched systems exhibit improved durability and reduced degradation compared with systems operating continuously at full power.

Several studies have also focused on modeling degradation caused by microstructural evolution in the anode. Khan et al. [109] developed a predictive model to analyze performance degradation caused by Ni grain growth in Ni-scandia stabilized zirconia (SSZ) anodes. By combining microstructural characterization using SEM with electrochemical measurements obtained through EIS, the study demonstrated that particle size evolution and triple-phase boundary (TPB) reduction follow a power-law coarsening model. The resulting electrochemical model accurately predicted the increase in anode polarization resistance over time.

In addition to electrochemical degradation, mechanical degradation has also been investigated through multiphysics modeling approaches. Fang et al. [110] developed a comprehensive model that integrates solid mechanics with coupled transport phenomena including mass, heat, flow, and charge transport. Their simulations provided detailed insights into stress development, creep strain rate, and mechanical failure probability in planar SOFCs. The study highlighted that mismatches in thermal expansion coefficients and temperature gradients are significant factors influencing long-term mechanical stability.

Furthermore, detailed computational models have been developed to analyze the influence of fuel contaminants on SOFC performance. For example, Sezer et al. [111] developed a three-dimensional computational fluid dynamics (CFD) model to simulate degradation of planar SOFCs exposed to phosphine contamination. The model parameters were calibrated using experimental data obtained from button cell experiments, enabling accurate prediction of performance degradation caused by contaminant exposure.

Despite their ability to provide valuable physical insight into degradation mechanisms, physics-based models also present several limitations. The predictive accuracy of these models is often constrained by incomplete understanding of degradation mechanisms occurring in electrodes, electrolytes, inter-

connects, and seals [112]. In addition, modeling challenges arise from material inhomogeneity, fabrication variations, and operational uncertainties. Because these models require numerous simplifying assumptions, they may produce overly idealized representations of degradation behavior [113]. Moreover, detailed multiphysics and CFD-based models [111] are computationally intensive, which limits their applicability for real-time monitoring and prognostics of SOFC degradation.

Data-driven Methods

Data-driven approaches construct predictive models directly from sensor measurements and operational data, without explicitly incorporating physical knowledge of the underlying system. These models learn the relationship between system inputs and outputs from historical data, enabling the prediction of future system states such as voltage, SOH, or RUL. Due to their ability to capture complex nonlinear relationships, data-driven methods are widely used for degradation prediction and fault diagnosis in SOFC systems.

Several machine learning models have been proposed to predict SOFC degradation using operational data. For example, Rao et al. [114] developed a multi-step prediction model using a two-layer long short-term memory (LSTM) network trained on data from degrading SOFC systems. The model was designed to handle sequential inputs and generate multi-step forecasts of system states. When compared with the traditional autoregressive integrated moving average (ARIMA) model and recursive LSTM prediction strategies, the proposed multi-step LSTM model demonstrated superior performance in predicting SOFC degradation trends over multiple future time steps.

Other studies have explored classical machine learning techniques for fault identification and prognostics. Wu et al. [115] employed a least squares support vector machine (LS-SVM) classifier to identify degradation modes in SOFC systems. This work was further extended by incorporating a hidden semi-Markov model (HSMM) to estimate the remaining useful life of the system [116]. By integrating the HSMM-based prognostics framework with an

empirical degradation model, the method was able to estimate the SOH and predict RUL under both cathode and anode poisoning conditions.

Various neural network architectures have also been investigated for voltage prediction and degradation forecasting. Li et al. [117] compared several recurrent neural network (RNN) architectures, including LSTM, gated recurrent unit (GRU), encoder-decoder LSTM, and encoder-decoder GRU models. Their results indicated that encoder-decoder architectures improved prediction accuracy compared with standard recurrent networks, with the encoder-decoder GRU model achieving the highest accuracy among the tested architectures. Similarly, Marra et al. [118] developed an artificial neural network (ANN) model designed for on-field diagnosis in SOFC-based distributed power generation systems. By carefully selecting input variables representing the temporal evolution of stack performance, the ANN model achieved improved prediction accuracy and generalization capability in forecasting voltage variations caused by degradation.

Other neural-network-based prognostic models have been proposed for SOFC systems operating under different conditions. Wu et al. [24] developed an Elman neural network model to predict future voltage values and detect potential failures based on residual voltage patterns. By combining the predictive model with the dynamic response of the SOFC system, the approach successfully identified root causes of voltage fluctuations and heat exchanger failures during different operational stages, including start-up, long-term operation, and hot-standby periods. In another study, Sonderskov et al. [119] utilized voltage, current, and temperature measurements collected from field-installed SOFC stacks to estimate a SOH metric. An LSTM-based recurrent neural network was then trained to forecast SOH values up to six months ahead.

Dynamic neural network models have also been explored for degradation trend prediction and RUL estimation. Cheng et al. [120] investigated three types of dynamic neural network models: neural network autoregressive with external inputs (NNARX), neural network autoregressive moving average with external inputs (NNARMAX), and neural network output error (NNOE). These mod-

els were used to forecast degradation trajectories and estimate the remaining useful life of SOFC systems. Comparative analysis showed that the NNARX model provided higher prediction accuracy and produced smaller relative errors in RUL estimation compared with NNARMAX and NNOE models.

In addition to purely data-driven approaches, several studies have combined machine learning techniques with degradation indicators derived from experimental measurements. For instance, Xu et al. [121] proposed using degradation resistance as an indicator of SOH and applied an Unscented Kalman Filter (UKF) for real-time SOH estimation. Similarly, Yang et al. [122] combined machine learning techniques with the DRT method to predict SOFC degradation caused by chromium poisoning. Their results demonstrated that the hybrid approach achieved strong predictive performance and showed good agreement with experimental measurements. Several studies have integrated statistical techniques with degradation indicators derived from experimental measurements. For instance, Xu et al. [121] proposed using degradation resistance as an indicator of SOH and applied an Unscented Kalman Filter (UKF) for real-time SOH estimation. Similarly, Yang et al. [122] combined machine learning techniques with the DRT method to predict SOFC degradation caused by chromium poisoning. Their results demonstrated that this hybrid approach achieved strong predictive performance and showed good agreement with experimental measurements. A number of studies have also utilized electrochemical characterization techniques to extract degradation indicators that can be used for prognostic analysis. For example, Muñoz et al. [123] employed EIS, polarization curves, DRT analysis, and ECM to forecast the degradation behavior of SOFCs. Their results indicated that these electrochemical diagnostics can successfully predict the performance degradation of individual cells operating under predefined conditions. Similarly, Gazzarri et al. [124] investigated cathode electrode delamination by analyzing variations in EIS parameters, particularly series resistance and polarization resistance, to estimate the resulting degradation in SOFC performance.

Data-driven prognostic approaches have also been investigated for SOFC sys-

tems operating under faulty conditions, with the objective of predicting the RUL of the system. Wu et al. [125] proposed a prognostic method based on phase-space trajectory similarity to estimate the RUL of SOFC systems. Building on this prognostic framework, a multi-objective dynamic optimization strategy was introduced to balance power generation efficiency and degradation rate throughout the predicted remaining life. This approach achieved a reduction in voltage degradation of up to 37.30%. Sheng et al. [126] proposed a framework in which a degradation resistance model combined with a Kalman filter was used for model-based RUL prediction, while a recursive multi-step LSTM network was used for data-driven prognostics. Comparative analysis showed that the LSTM-based approach achieved higher prediction accuracy and was better suited for RUL prediction compared with the Kalman filter-based method. [127] proposed a hybrid framework that couples a dynamic decay model with a neural network, achieving less than 5% prediction error and providing a practical approach to predict life span of SOFC systems. Complementing these, [128] estimated a local linear trend using ordinary least squares (OLS) and derived a closed-form probability density function (PDF) for RUL, but did not provide uncertainty bands for the full SOH trajectory. Despite the promising results achieved by data-driven methods, several limitations remain. First, the performance of these models strongly depends on the availability of large quantities of high-quality training data. However, collecting degradation datasets for SOFC systems is often costly and time-consuming. Second, the accuracy of the models depends on whether the training data sufficiently capture the degradation behavior of the system. Finally, data-driven approaches require careful selection of model architectures and hyperparameters to achieve reliable predictions, which may limit their practicality in some applications.

Hybrid Methods

Hybrid approaches can be broadly categorized into two groups. The first combines physics-based and data-driven models at the model or output level,

where a physics model provides baseline predictions that are refined using machine learning techniques. The second, commonly referred to as physics-informed machine learning (PIML), embeds physical knowledge directly into the machine learning framework. Rather than treating physics and learning as separate modules, PIML incorporates domain knowledge throughout the learning process to improve prediction accuracy, robustness, interpretability, and generalization, particularly when experimental data are limited.

Hybrid model coupling. Hybrid approaches combine physics-based modeling with data-driven techniques in order to leverage the advantages of both methodologies. By integrating physical knowledge of the system with data-driven learning capabilities, these approaches can improve prediction accuracy while maintaining physical interpretability. Several studies have explored such hybrid frameworks for SOFC monitoring, fault diagnosis, and degradation prognostics.

In [129], a machine-learning framework was developed for the identification and early mitigation of critical events in SOFC systems through the detection of subtle system anomalies. The proposed approach employs a hybrid architecture consisting of two complementary components: a data-driven model based on gradient-boosted decision trees and a computational model representing the underlying physics of the SOFC system. By combining these components, the method enables reliable detection of abnormal system behavior at an early stage.

Zu et al. [130] proposed a hybrid degradation prediction method that integrates an empirical degradation model with a data-driven learning algorithm. This combined approach captures the long-term degradation trends observed in aging data while simultaneously preserving short-term nonlinear degradation characteristics. As a result, the hybrid framework significantly improves prediction accuracy compared with using either empirical models or purely data-driven methods alone.

Similarly, Gallo et al. [21] developed an online aging estimation algorithm that

combines an envelope dynamic model with electrochemical impedance spectroscopy measurements. This hybrid framework enables real-time monitoring of SOFC stacks, allowing both fault detection during operation and prediction of the RUL of the system.

Physics-informed machine learning (PIML). Physical knowledge can be incorporated into machine learning models in several ways.

- Observational bias introduces physics during data generation through experiments, simulations, preprocessing, or physically meaningful feature extraction.
- Learning bias embeds physical laws into the optimization process by modifying the loss function, adding constraints, or regularizing the training objective.
- Inductive bias integrates prior physical knowledge into the model architecture itself, for example through graph structures, conservation-aware network designs, or physics-guided feature representations.
- Discrepancy bias combines analytical physics models with machine learning by allowing the learning algorithm to model only the unknown or unmodeled dynamics, thereby compensating for discrepancies between simplified physical models and real-world system behavior.

In this thesis, the proposed prognostic framework follows the observational and inductive bias paradigms. Physics-based information extracted from electrochemical impedance spectroscopy through DRT analysis is incorporated as physically meaningful features, while the graph representation of impedance components encodes electrochemical relationships within the learning architecture. These physics-informed elements improve both model interpretability and prediction accuracy compared with purely data-driven approaches.

1.2.4 Degradation Mode Classification Methods

Simulation modeling and machine learning have been increasingly adopted for SOFC monitoring, diagnosis, and fault isolation. Early studies combined physics-based system models with data-driven classifiers to enable fault recognition under varying operating conditions. For example, Costamagna et al. proposed a model-based diagnostic framework that integrates a quantitative SOFC model with support vector machines (SVM) to detect and isolate multiple fault types across a range of system operating regimes [131]. Polverino et al. employed a dynamic SOFC system model to emulate increased polarization losses, enabling the construction of a signature matrix for on-site fault diagnosis [132]. More recently, Le et al. introduced a simulation-informed machine learning approach that leverages physics-based EIS simulations of SOFC stacks under normal aging and several key degradation modes to train an SVM classifier for failure detection [133].

Beyond simulation-driven approaches, machine learning has also been applied directly to experimental and operational SOFC datasets. Zhu et al. developed a fault diagnosis framework based on principal component analysis combined with random forest (PCA-RF), demonstrating high accuracy in identifying critical faults such as heat exchanger rupture and hydrogen shortage in SOFC power generation systems [134]. Wu et al. investigated degradation mode identification using least-squares SVM (LS-SVM) classifiers [115]. In addition, Wu et al. developed an Elman neural network model to forecast voltage trajectories and detect potential degradations based on residual voltage patterns [24]. By incorporating system dynamics, this approach was shown to capture voltage fluctuation root causes and heat exchanger failures across multiple operating phases, including start-stop events, long-term operation, and hot-standby conditions.

1.3 Research Gaps

Although significant progress has been made in SOFC health monitoring, degradation analysis, and prognostics, several important research gaps remain. These gaps limit the development of reliable predictive maintenance and supervisory control strategies for SOFC systems. The major research gaps identified through the literature review and experimental experience are summarized as follows:

- **Limited comprehensive studies across multiple degradation mechanisms:** Most existing studies focus on a single degradation mechanism or specific operational condition. A comprehensive and systematic investigation that addresses multiple dominant SOFC degradation modes within a unified framework remains lacking.
- **Lack of controlled degradation experiments under diverse operating conditions:** Experimental studies that systematically investigate degradation under controlled stress conditions are limited. In particular, accelerated degradation experiments operating at different stress levels are needed to better understand how operating conditions influence degradation evolution and system durability.
- **Limited availability of diverse experimental degradation datasets:** The development of reliable data-driven prognostic models requires large and diverse datasets. However, publicly available SOFC degradation datasets that include multiple degradation modes and full-lifespan trajectories are extremely limited.
- **Limited use of physics-informed feature engineering in machine learning models:** Most data-driven approaches rely primarily on voltage and current signals, which provide limited insight into underlying electrochemical degradation processes. Incorporating physics-informed indicators extracted from EIS, such as DRT-based features, can improve

model interpretability, reliability, and generalization for SOFC health monitoring and prognostics.

- **Limited RUL prediction frameworks with uncertainty quantification:** Reliable prognostics require not only accurate point estimates but also well-calibrated uncertainty bounds to support risk-aware maintenance and operational decision-making. However, most existing RUL prediction models do not explicitly quantify aleatoric and epistemic uncertainties associated with long-term degradation forecasting.
- **Challenges in accurate long-horizon SOH forecasting under dynamic prediction horizons:** Existing prognostic models often focus on short-term forecasting and struggle to maintain accuracy over long degradation horizons toward end-of-life. In addition, as SOFC systems age, the remaining prediction horizon continuously shrinks, making conventional fixed-horizon forecasting strategies less effective and potentially unstable during training. Developing models capable of accurately forecasting the full SOH trajectory while handling dynamic prediction horizons is therefore essential for reliable long-term prognostics and predictive maintenance.
- **Lack of benchmarking studies for degradation mode classification using impedance-derived features:** Although EIS provides rich diagnostic information about internal electrochemical processes, systematic benchmarking of machine learning methods for degradation mode classification using impedance-derived indicators, such as DRT-based features, remains limited.
- **Lack of integrated diagnostic-prognostic frameworks for multi-degradation mode systems:** Most existing prognostic models are degradation-mechanism agnostic and therefore may exhibit reduced prediction accuracy across different degradation modes. Integrating degradation mode identification with RUL prediction within a unified multi-

task learning framework can improve prognostic accuracy by enabling shared learning between diagnostic and prognostic tasks.

1.4 Thesis Contributions and Outline

1.4.1 Thesis Contributions

Chapter 2 Contributions

- Nearly identical anode-supported tubular SOFCs are fabricated and multiple full-lifespan degradation experiments are conducted, generating five independent degradation trajectories.
- A comprehensive experimental study of Redox-induced SOFC degradation is carried out under both regular and random Redox cycling intervals, resulting in eight datasets collected at different operating temperatures and degradation severities.
- The Redox cycling experiments produce a large-scale multi-modal degradation dataset consisting of 2,143,000 samples, including transient voltage/current density signals, polarization curves (J-V), and detailed EIS measurements.
- The resulting Redox cycling dataset is publicly released to support future research in SOFC health monitoring and prognostics.
- Additional experimental campaigns are conducted to investigate three other critical SOFC degradation mechanisms: fuel starvation, internal short-circuit faults, and thermal-stress-induced degradation, generating three degradation datasets for each degradation mode.
- In addition to the fault-induced degradation datasets, a long-duration dataset representing normal aging behavior is collected to serve as a baseline reference for healthy system operation.
- Throughout the experiments, transient voltage and current measurements together with repeated EIS and J-V measurements are performed

to monitor degradation evolution, enabling the extraction of mechanism-related features through DRT analysis.

Chapter 3 Contributions

- A novel physics-informed framework for real-time SOH estimation in SOFC systems is proposed, integrating electrochemical impedance spectroscopy (EIS) measurements with machine learning. The proposed approach reduces forecasting error by 37% compared to a baseline LSTM model relying solely on voltage and current measurements.
- A physics-structured Temporal Graph Convolutional Network (TGCN) is developed, where graph nodes correspond to electrochemical processes identified through DRT peaks, and graph connections encode the physical interdependencies among degradation phenomena in SOFCs. This design improves both predictive accuracy and model interpretability.
- The study introduces the first long-horizon SOH forecasting framework for SOFC systems under Redox cycling degradation, addressing one of the most critical degradation mechanisms in SOFC operation.
- A physics-informed machine learning (PIML) approach is developed that fuses voltage, current density, and impedance-derived features, enabling accurate and robust SOH estimation during dynamic degradation conditions.
- The proposed framework demonstrates strong generalization across diverse operating conditions, validated through cross-dataset experiments involving multiple Redox cycling severities and operating temperatures, achieving $\text{RMSE} = 0.070\text{--}0.085$ and $R^2 = 0.94\text{--}0.95$.
- The proposed model achieves real-time inference capability suitable for online deployment, with total computation time for DRT feature extraction and model prediction remaining below 300 ms, which is negligible relative to the 2-minute EIS acquisition interval.

Chapter 4 Contributions

- The entire future SOH trajectory is forecasted, not just time-to-failure, using Bayesian deep learning models based on LSTM and Informer architectures. Trajectory forecasts provide rich information on health evolution toward EOL estimation, enabling proactive control and maintenance well before failure. To comprehensively evaluate forecasting performance, deterministic baseline models including temporal convolutional networks (TCN), GRU, and a deterministic Informer are also implemented and compared with the Bayesian LSTM and Bayesian Informer frameworks.
- Aleatoric uncertainty is captured by adding a variance output head trained with a Gaussian negative log-likelihood loss, and epistemic uncertainty is estimated via Monte Carlo (MC) dropout. This results in establishing well-calibrated confidence intervals (CIs) on SOH trajectories. Furthermore, by extracting the first hitting times (FHT) of SOH threshold crossings across MC dropout trajectories, an RUL distribution is derived.
- An innovative method is developed to address the shrinking forecast horizon that naturally occurs as an SOFC approaches its EOL. Unlike conventional approaches that retrain separate models for each horizon, this method introduces zero-padding with exponentially decaying loss weights on the padded tail. This strategy maintains a consistent forecast sequence length across all samples while guiding the model to converge to $\text{SOH} = 0$ at EOL. In addition, sensitivity analyses are conducted to evaluate the impact of the padding strategy and the decay factor on model performance.
- Under identical data and training protocols, the Bayesian attention-based Informer achieves approximately 90% lower forecast error, 6% higher coverage, and about 55% lower sharpness than a Bayesian LSTM

model, demonstrating the advantage of attention mechanisms for long-horizon degradation forecasting.

- Model robustness and generalization capability are assessed using a leave-one-dataset-out cross-validation strategy, where models are trained on multiple degradation datasets and evaluated on unseen datasets to verify their ability to generalize across different degradation severities and operating conditions.

Chapter 5 Contributions

- A comprehensive framework is developed for identifying SOFC degradation modes using EIS measurements. Instead of relying on voltage or current signals, the method uses impedance-derived diagnostic indicators to distinguish between different degradation mechanisms.
- Electrochemical impedance spectra collected during degradation are processed using the DRT method to extract physically interpretable features representing electrochemical processes within the cell. These DRT-derived trajectories, together with ohmic resistance, provide compact and mechanism-informed indicators that form the basis of the classification dataset.
- Six representative classification models are evaluated to benchmark degradation mode identification performance. These include three shallow ensemble methods. Random Forest (RF), XGBoost (XGB), and an RF+XGB ensemble, as well as three temporal deep learning models: LSTM, TCN, and the attention-based Informer architecture.
- The study provides a systematic comparison between classical feature-based ensemble learners and sequence models capable of capturing short- and long-range temporal dependencies in impedance evolution.
- This work presents the first experimental benchmark that utilizes DRT-derived impedance features to train and evaluate modern time-series

classifiers, including attention-based architectures, for multi-degradation mode identification in SOFC systems.

Chapter 6 Contributions

- A degradation mechanism-aware prognostics framework is proposed based on an attention-driven sequence model that jointly estimates the active degradation mode and the RUL of SOFC systems while capturing long-range degradation dynamics.
- A learning framework is developed that conditions RUL estimation on degradation mechanism information using an attention-based Informer architecture. The Informer model leverages its long-range attention mechanism to capture temporal dependencies in degradation trajectories, while the mechanism-aware conditioning enables the model to exploit correlations between degradation modes and degradation patterns for improved RUL prediction.
- A systematic investigation of early and late fusion strategies is conducted to determine the most effective stage for incorporating degradation mode information into the RUL estimation pipeline.
- The separability of degradation mechanisms within the learned feature space is analyzed using silhouette scores and pairwise class centroid distances, providing quantitative and physically interpretable insights into degradation behavior.
- An asymmetric MTL design is introduced in which latent embeddings from the degradation mode classifier are transferred to the RUL estimation branch, leading to up to a 43% reduction in root mean square error (RMSE) compared with a mechanism-agnostic baseline.
- The proposed asymmetric MTL framework demonstrates improved computational efficiency, achieving higher predictive accuracy while reducing

training and inference costs by approximately 8% compared with sequential classification and RUL estimation pipelines.

1.4.2 Thesis Outline

Figure 1.12 illustrates the overall structure and workflow of the thesis, highlighting the relationships between the different chapters and research components. The work begins in Part I with the introduction, identification of research gaps, and the development of a comprehensive experimental framework, including degradation datasets generated under multiple degradation mechanisms. Building on this foundation, Part II focuses on prognostics, where physics-informed and data-driven models are developed for SOH and RUL forecasting, incorporating impedance-derived features and uncertainty quantification. Part III addresses diagnostics by developing and benchmarking models for degradation mode classification using electrochemical impedance signatures. These two research streams are subsequently integrated in Part IV through a mechanism-aware prognostics framework based on multi-task learning, enabling joint degradation mode identification and RUL prediction. Finally, Part V presents the conclusions and outlines future research directions. The figure also emphasizes the progressive flow of information, demonstrating how experimental data and intermediate insights are leveraged across chapters to construct increasingly advanced and integrated modeling frameworks.

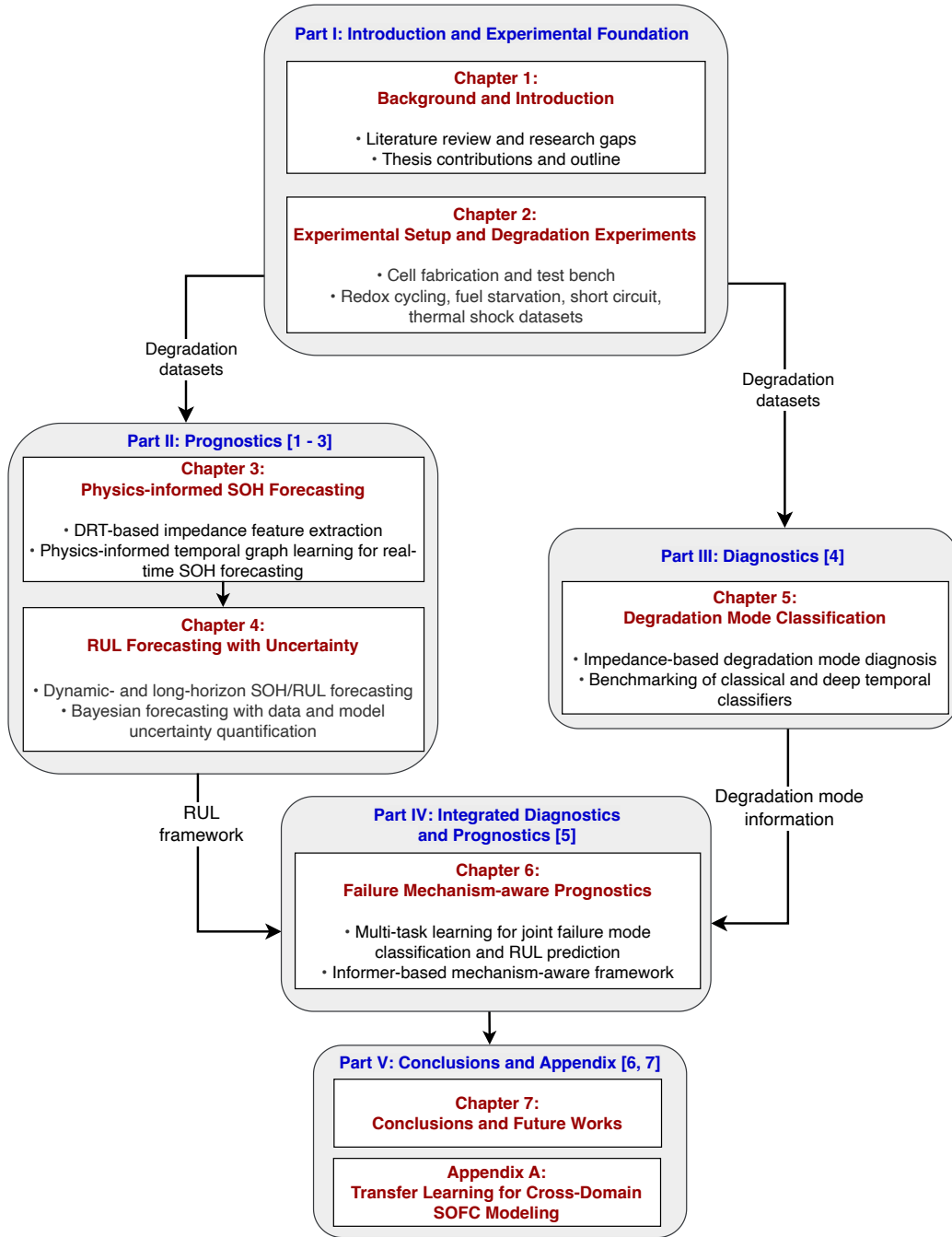


Figure 1.12: Schematic overview of the thesis organization. Journal publications/submissions from this thesis are shown in brackets for each part of the thesis.

Chapter 2

Experimental Setup and Datasets

2.1 Introduction

This chapter presents the experimental foundation of the thesis, including the design, fabrication, and testing of SOFCs, as well as the systematic generation of degradation datasets used throughout the subsequent chapters. Since the proposed diagnostic and prognostic frameworks are data-driven and physics-informed, the quality, diversity, and reliability of the experimental data play a critical role in ensuring model accuracy, interpretability, and generalizability. First, the fabrication process and structural characteristics of the anode-supported tubular SOFCs are described, followed by a detailed overview of the experimental setup used for electrochemical characterization. Standard performance measurements, including J-V curves, OCV, and EIS, are introduced to establish the baseline testing procedures.

The chapter then presents a comprehensive set of controlled degradation experiments designed to replicate key degradation mechanisms encountered in SOFC operation. These include redox cycling, fuel starvation, short circuit faults, and thermal shock. Each degradation mode is imposed through carefully designed protocols to accelerate aging while preserving the underlying physical degradation processes. The experiments are conducted on multiple nominally identical cells to ensure repeatability and to capture variability across degradation trajectories.

Finally, the resulting datasets are summarized, including both transient multivariate time-series measurements and steady-state electrochemical characterization data. The collected data capture the coupled electrochemical, thermal, and operational dynamics of SOFC degradation, forming a large-scale, high-fidelity dataset that supports degradation mode classification, SOH estimation, and RUL prediction in the following chapters.

2.2 Anode-supported Tubular Cells

This study uses anode-supported SOFCs that offer high performance and stability. The cell designs, formulations, and experimental setups were fully developed at the Energy Mechatronics Laboratory (EML) at the University of Alberta. In tubular cells, the anode supports were fabricated using a slip-casting method with a NiO-YSZ (yttria-stabilized zirconia, 8 mol% Y_2O_3) anode composition. Graphite (40 vol%, average particle size of $4.5\ \mu\text{m}$; Sigma-Aldrich) was used as a pore former to generate porosity in the anode support. After slip casting, the anode supports were pre-sintered at $1000\ ^\circ\text{C}$. Subsequently, a Ni-YSZ anode functional layer (AFL) and a YSZ electrolyte layer were coated onto the supports and co-sintered at $1400\ ^\circ\text{C}$. A porous scaffold of $\text{Ce}_{0.9}\text{Gd}_{0.1}\text{O}_{2-\delta}$ (gadolinium-doped ceria, GDC) was deposited on the electrolyte surface and sintered at $1250\ ^\circ\text{C}$. The porous scaffold was then infiltrated with a $\text{Pr}_2\text{NiO}_{4+\delta}$ (PNO) catalyst to form the cathode. Copper and silver mesh current collectors were attached to the anode and cathode surfaces, respectively, for electrochemical performance testing. The tubular cells were placed in a tubular furnace and reduced using a nitrogen/hydrogen gas mixture for 24 h, followed by gradual replacement of nitrogen with humidified hydrogen (50 sccm) to ensure complete reduction of NiO to metallic Ni within the anode microstructure. Detailed information on the fabrication procedure of these cells can be found in our previous works [135, 136]. The SEM image of a fabricated anode-supported SOFC, shown in Figure 2.1, illustrates the main components, including the anode support, anode functional layer, electrolyte,

and cathode, along with their respective dimensions.

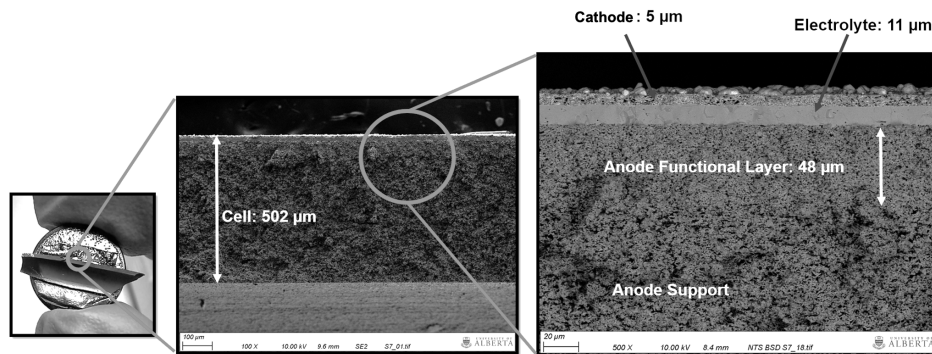
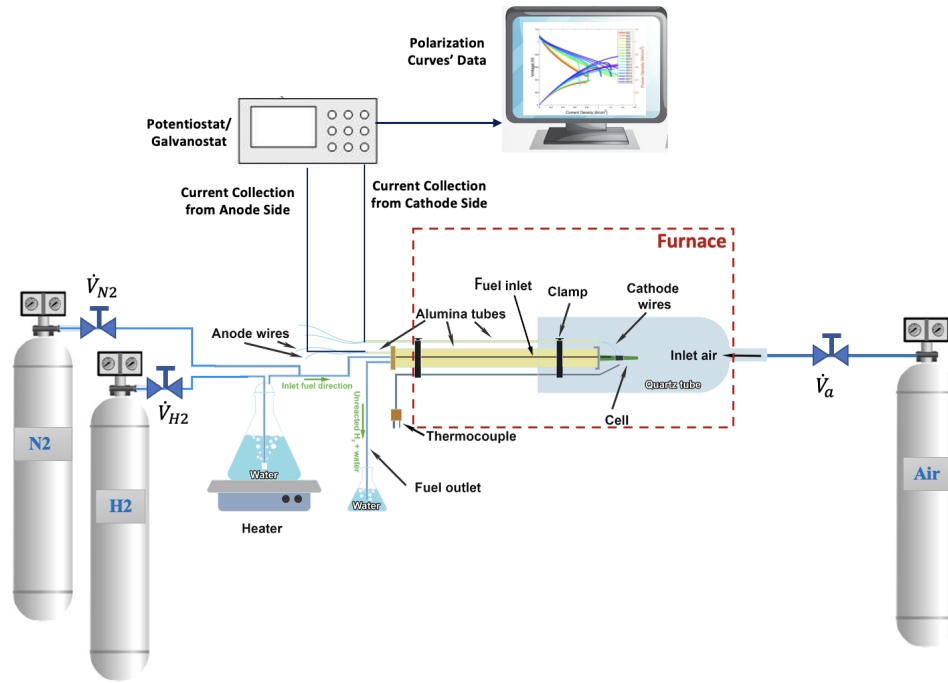


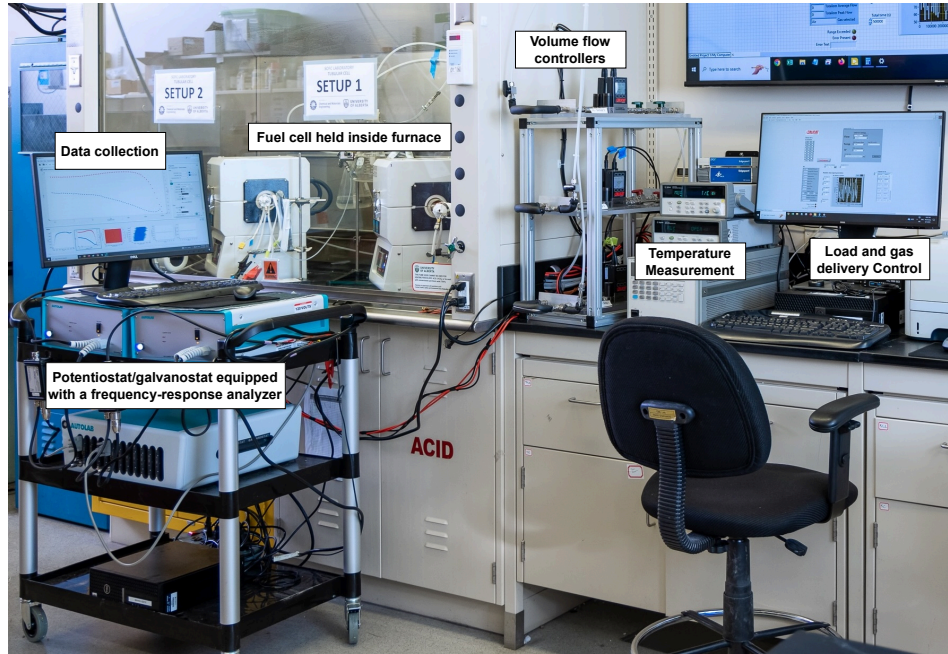
Figure 2.1: A SEM image of a fabricated anode-supported SOFC showing the main components and the dimensions.

2.3 Experimental Setup

In order to conduct the electrochemical performance tests on the fuel cells, copper and silver mesh current collectors were attached to the anode and cathode surfaces, respectively. Next, silver wires were connected to the surface of each mesh and these wires were used to measure the cell performance. The cell was placed inside the hot zone of a furnace. Initial reduction of NiO as well as sealing of ceramabond took place by gradually increasing the temperature of the furnace to 750 °C while dwelling at 90 and 260 C for 1 h. Meanwhile, nitrogen and hydrogen are mixed with a 43:7 volume ratio and fed to the anode side. Air is provided to the cathode by a quartz tube. The flow rate of the gases is controlled using digital volume flow controllers. The temperature of the cell is monitored using a thermocouple close to the cathode active area. Performance measurements such as J-V, OCV, and EIS of the cell was measured using Solartron SI 1287A Potentiostat/Galvanostat with the Solartron 1255A Frequency Response Analyzer module. Figure 2.2a illustrates the setup schematic, and Figure 2.2b shows the photograph of the experimental setup. To measure the J-V curve, the current density is raised gradually from zero to the maximum current density in specific levels and the voltage response is recorded. In EIS measurements, a sinusoidal current (I) or voltage (V) pertur-



(a) Schematic of the anode-supported tubular SOFC mounted inside the furnace with gas supply lines, electrical connections, and temperature measurement.



(b) The experimental setup, including the potentiostat/galvanostat with frequency response analyzer, volume flow controllers, and the load and gas delivery control system.

Figure 2.2: Cell configuration and experimental test bench used for SOFC electrochemical characterization.

bation is applied to the fuel cell operating under the OCV condition, and the resulting voltage or current response is measured. The EIS measurements are typically conducted over a frequency (f) range of 1 MHz to 100 mHz, but can be adjusted based on the anticipated processes to be analyzed. In this study, EIS curves are measured in frequency range of 100 KHz to 0.1 Hz. To maintain system linearity, the signal amplitude must be sufficiently low to achieve a pseudo-linear response from the system [41–43]. Nonetheless, excessively small amplitudes can lead to a poor signal-to-noise ratio, necessitating the selection of a sufficiently high amplitude value to minimize signal noise. Various sinusoidal voltage amplitudes have been proposed in the existing literature, such as 10 mV [53, 54], 12 mV [55], 14 mV [54], and 50 mV [56]. Here, the sinusoidal voltage amplitude is 20 mV.

2.4 Accelerated Degradation Experiments

2.4.1 Redox Cycling

To characterize the Redox-induced degradation phenomenon, a standard cyclic protocol was designed and applied to eight single tubular SOFCs. All cells were fabricated under identical conditions and exhibited nearly identical initial performance, confirmed by baseline J-V and EIS measurements collected at OCV. Prior to testing, each cell underwent a 72-hour stabilization period at a constant current density of 0.3 A/cm² to ensure steady electrochemical performance before initiating the Redox cycles.

The Redox cycling tests were conducted at a constant operating voltage of 0.7 V, while the corresponding current density was continuously monitored as an indicator of performance degradation. The air volume flow rate was maintained at $\dot{V}_{\text{air}} = 100$ SCCM throughout the tests, and the hydrogen flow rate alternated between full fuel supply ($\dot{V}_{\text{H}_2} = 50$ SCCM) and complete cutoff ($\dot{V}_{\text{H}_2} = 0$ SCCM), as shown in Figure 2.3. In the regular Redox tests (N1-N6), each pattern was repeated for 10 consecutive cycles before transitioning to the next pattern. If the 10% power degradation threshold, a commonly used EOL

criterion [102, 126], was not reached, the hydrogen cutoff duration ($t_{\text{H}_2 \text{ cutoff}}$) was sequentially extended to 25 and then 30 minutes, while maintaining the full hydrogen duration at 5 minutes; each extended pattern was also repeated for 10 cycles. Additionally, two tests (R1 and R2) employed random Redox cycling, in which the $t_{\text{H}_2 \text{ cutoff}}$ was randomly selected between 5 and 30 minutes in each cycle. The total cycle duration was kept constant at 40 minutes, such that the full hydrogen supply period equaled $(40 - t_{\text{H}_2 \text{ cutoff}})$ minutes. After each cutoff period, once hydrogen flow was restored, the cell was returned to OCV for EIS and J-V characterization. To investigate temperature effects, the regular Redox procedures (N1-N6) were repeated at different temperatures (T_{cell}) of 650, 700, and 750 °C, whereas the randomized Redox cycles (R1 and R2) were conducted only at 750 °C. A comprehensive list of the Redox datasets is presented in Table 2.1, describing the test conditions and cell lifetimes (τ).

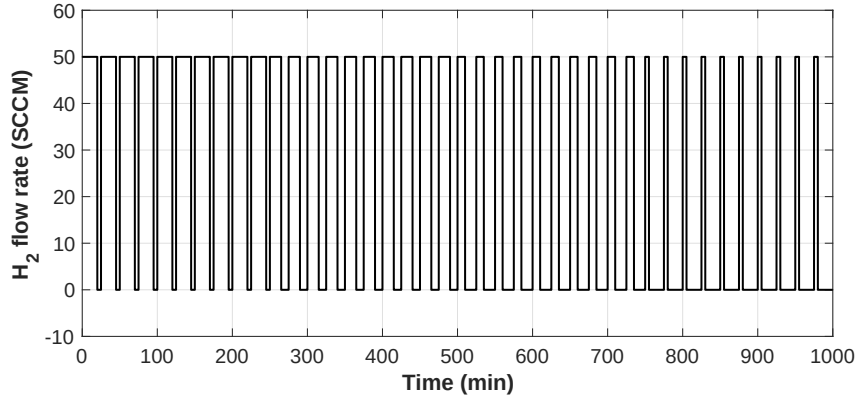


Figure 2.3: Hydrogen flow rate profile applied during redox cycling experiments. The anode fuel supply was periodically switched between 50 SCCM and 0 SCCM, with the fuel-off interval increased over time to accelerate degradation.

2.4.2 Fuel Starvation

Fuel starvation experiments were conducted galvanostatically at a constant current density of 0.3 A/cm², with the furnace temperature maintained at 750°C and an air flow rate of 100 SCCM. Each starvation cycle consisted of two operating phases: a nominal operation period with a hydrogen flow rate of 50 SCCM for 30 min, followed by a hydrogen-deficient phase at 5 SCCM

Table 2.1: Redox datasets conditions. Each test is performed in one of the eight fabricated cell. Cells N1 to N6 underwent normal Redox cycling, and cells R1 and R2 underwent random Redox cycling datasets. T_{cell} , P_{max} , $t_{\text{H}_2 \text{ cutoff}}$, and τ are cell temperature, maximum power, and hydrogen cutoff interval, and cell lifetime. All tests were conducted under constant voltage of 0.7 V and constant air volume flow rate ($\dot{V}_{\text{air}}$) of 100 SCCM. Hydrogen volume flow rate ($\dot{V}_{\text{H}_2}$) changed between 0 and 50 SCCM.

Cell	T_{cell} ($^{\circ}\text{C}$)	P_{max} (W/cm^2)	$t_{\text{H}_2 \text{ cutoff}}$ (min)	τ (h)
N1	650	0.8	[5, 10, 15, 20, 25, 30]	26.9
N2	650	0.8	[5, 10, 15, 20, 25, 30]	27.5
N3	700	0.6	[5, 10, 15, 20, 25, 30]	11.7
N4	700	0.6	[5, 10, 15, 20, 25, 30]	47.0
N5	750	0.4	[5, 10, 15, 20, 25, 30]	46.9
N6	750	0.4	[5, 10, 15, 20, 25, 30]	42.5
R1	750	0.8	Random [5 to 30]	54.6
R2	750	0.8	Random [5 to 30]	49.7

for 15 min to intentionally induce fuel starvation. Upon completion of each starvation segment, the hydrogen flow was restored to 50 SCCM and the cell was switched to OCV conditions. The cell was held at OCV for approximately 10 min in total, including the time required to acquire J–V and EIS measurements (approximately 3 min), to ensure adequate electrochemical stabilization before and after characterization. This OCV settling procedure was applied consistently across all degradation experiments. The hydrogen flow rate profile follows the same cyclic structure as in the redox cycling tests shown in Figure 2.3; however, the fuel-deficient segment was fixed at 15 min, and the anode inlet flow was switched between 50 SCCM (nominal) and 5 SCCM (starvation) instead of a complete shutoff.

2.4.3 Short Circuit

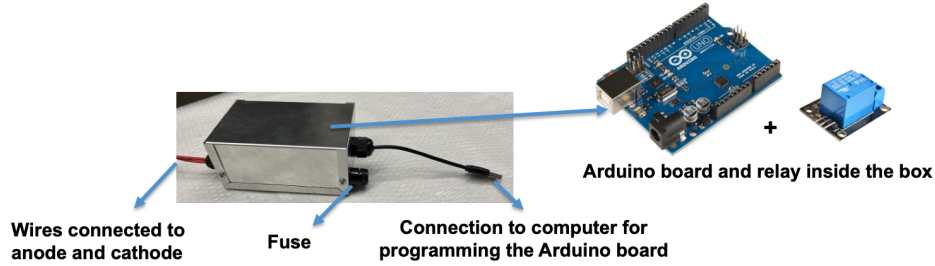
Short circuit degradation was introduced through repeated external shorting events designed to emulate internal short-circuit faults. The experiments were conducted at a constant current density of $0.3 \text{ A}/\text{cm}^2$, with an air flow rate of

100 SCCM and a furnace temperature of 750°C. Each cycle consisted of 10 min of steady galvanostatic operation, during which the anode and cathode were momentarily connected for 5 s to induce a controlled short circuit. Following each shorting event, the cell was returned to OCV conditions and held for stabilization before diagnostic J–V and EIS measurements were collected, consistent with the OCV settling procedure used in the other degradation experiments.

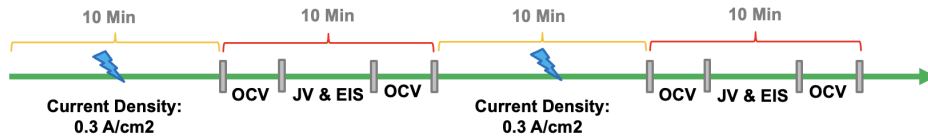
To ensure repeatability and operational safety, a customized external short-circuit device was developed for these experiments. The device consisted of a relay module enclosed within a protective housing and was controlled by an Arduino microcontroller. Two high-current leads extended from the enclosure and were connected directly to the anode and cathode terminals of the fuel cell. The relay was actuated through programmed commands sent from a host computer to the Arduino board, enabling precise control of the short-circuit duration. During each degradation cycle, the relay was closed for exactly 5 s, directly connecting the anode and cathode to induce the short-circuit condition. For safety, an inline fuse was incorporated into the circuit; if the short-circuit current exceeded a predefined safe threshold, the fuse interrupted the circuit to protect both the fuel cell and the test equipment. An image of the designed device and the short circuit timing procedure are shown in Figure 2.4. The figure illustrates the hardware implementation, including the enclosed relay-Arduino unit with safety fuse and external connections, as well as the programmed timing sequence used to impose controlled short-circuit events during degradation testing.

2.4.4 Thermal Shock

Thermal shock experiments were performed manually to impose severe and rapid temperature transients representative of abnormal thermal events. Each thermal shock cycle began with the cell stabilized at 750°C for 10 min, after which it was removed from the furnace and exposed to ambient air. The cell cooled from 750°C to 30°C over approximately 5 min, with an initial rapid cool-



(a) Custom designed external short circuit device used to impose controlled 5 s shorting events, consisting of an Arduino-controlled relay housed in a protective enclosure with an inline fuse for safety.



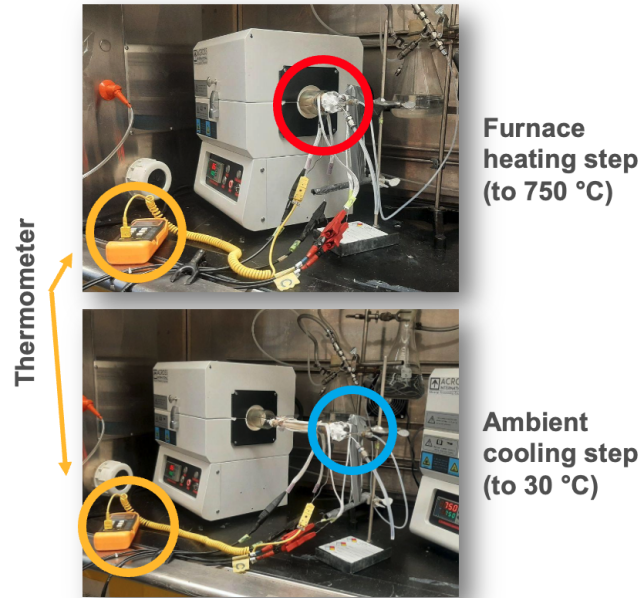
(b) Short circuit test timing and characterization sequence, showing galvanostatic operation at 0.3 A/cm² followed by transitions to OCV for J-V and EIS acquisition.

Figure 2.4: Short circuit degradation experiment.

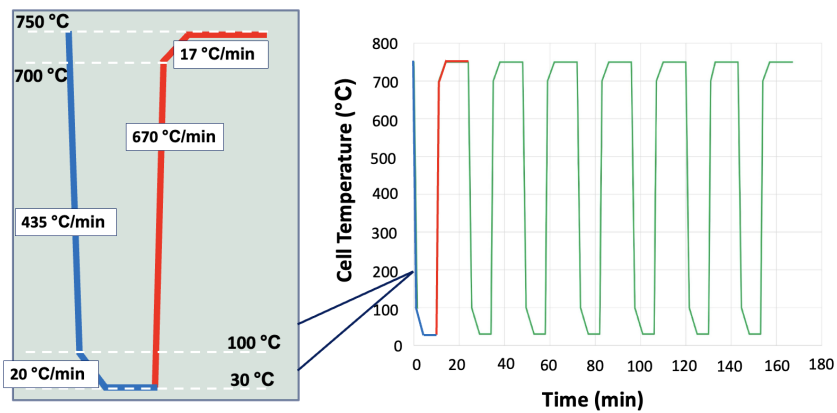
ing stage from 750°C to about 100°C at an average rate of 435°C/min, followed by a slower cooling stage from 100°C to 30°C at approximately 20°C/min. After a 10 min dwell at 30°C, the cell was reinserted into the furnace and reheated to 750°C over approximately 4 min, with heating rates of 670°C/min from 30°C to 700°C and 17°C/min from 700°C to 750°C. Once thermal equilibrium was re-established, the cell was transitioned to OCV and held for stabilization prior to acquiring J-V and EIS measurements, following the same procedure used in the other degradation tests. The thermal shock protocol is illustrated in Figure 2.5, including both the manual insertion/removal procedure used to generate repeated thermal shocks and the measured transient temperature profile.

2.5 Degradation Dataset Summary

Collectively, the redox cycling, fuel starvation, short circuit, and thermal shock experiments produced a comprehensive dataset comprising over 5 million sam-



(a) Photographs of the manual thermal shock procedure, where the cell is inserted into the furnace hot zone (heating step) and then withdrawn to ambient conditions (cooling step) to impose rapid temperature gradients.



(b) Measured temperature profile during repeated heating and cooling transients.

Figure 2.5: Thermal shock degradation protocol.

ples. This dataset captures both transient operational dynamics and steady-state electrochemical characteristics, including J–V curves and EIS measurements, and forms the foundation for the subsequent degradation mode classification analysis.

The transient data consist of multivariate time-series measurements, where each time step includes five simultaneously recorded signals: cell voltage, current, temperature, hydrogen flow rate, and air flow rate. These signals provide a comprehensive representation of the system’s dynamic behavior under varying operating conditions and degradation mechanisms.

Table 2.2 summarizes the composition of the dataset across all experiments. It reports the number of transient time-step samples, along with the number of J–V and EIS characterization scans collected throughout each degradation trajectory. The corresponding sample counts for J–V and EIS data reflect the number of measurements extracted from each scan.

Table 2.2: Summary of all degradation datasets and corresponding sample counts.

Dataset	Transient samples	J–V Curves (samples)	EIS Curves (samples)	Dataset samples
Redox - N1	387,361	40 (2,400)	40 (2,440)	392,201
Redox - N2	396,997	50 (3,000)	50 (3,050)	402,047
Redox - N3	168,483	25 (1,500)	25 (1,525)	171,508
Redox - N4	676,799	61 (3,660)	61 (3,721)	684,180
Redox - N5	675,360	61 (3,660)	61 (3,721)	682,741
Redox - N6	612,005	61 (3,660)	61 (3,721)	619,386
Redox - R1	786,242	70 (4,200)	70 (4,270)	794,712
Redox - R2	715,674	55 (3,300)	55 (3,355)	722,339
Starvation - 1	162,292	15 (900)	15 (915)	164,107
Starvation - 2	172,804	16 (960)	16 (976)	174,740
Starvation - 3	158,403	15 (900)	15 (915)	160,218
ShortC - 1	22,187	10 (600)	10 (610)	23,397
ShortC - 2	47,112	20 (1,200)	20 (1,220)	49,532
ShortC - 3	38,404	16 (960)	16 (976)	40,340
Thermal - 1	97,179	15 (900)	15 (915)	98,994
Thermal - 2	99,376	15 (900)	15 (915)	101,191
Thermal - 3	83,632	13 (780)	13 (793)	85,205
Total	5,283,312	647 (38,820)	647 (39,518)	5,361,650

Chapter 3

Physics-Informed State-of-Health Estimation¹

3.1 Introduction

Building on the experimental datasets and degradation scenarios introduced in Chapter 2, this chapter focuses on the development of an accurate and interpretable framework for SOH estimation in SOFCs. Reliable SOH estimation is a critical component of health monitoring systems, as it enables early detection of degradation and supports predictive maintenance strategies. However, accurately capturing SOFC degradation remains challenging due to the complex interactions among electrochemical processes and the limitations of conventional data-driven models that rely primarily on voltage and current measurements.

To address these challenges, this chapter proposes a physics-informed machine learning approach that integrates EIS data with conventional operational signals for improved SOH forecasting. In particular, physically meaningful features are extracted from impedance measurements using DRT analysis, providing a compact representation of the underlying electrochemical processes governing degradation. These features are then incorporated into a temporal graph convolutional network (TGCN), where the model structure is designed to reflect the interactions among different degradation-related phenomena.

The proposed framework enables the model to capture both temporal depen-

¹This chapter is partially based on the following publication: [1].

dencies and physics-based relationships within the system, resulting in improved predictive accuracy and interpretability compared to traditional sequence models. The effectiveness of the approach is demonstrated using multiple Redox cycling datasets, where the model achieves significantly lower forecasting error and strong generalization across varying operating conditions. In addition, the computational efficiency of the model supports real-time deployment for online SOH monitoring.

Overall, this chapter establishes a physics-informed foundation for data-driven SOH estimation in SOFC systems by combining impedance-based feature extraction with advanced temporal learning models. The proposed methodology addresses key limitations of existing approaches and provides a robust basis for subsequent developments in long-horizon prognostics and uncertainty-aware lifetime prediction presented in the following chapters.

3.2 Data Processing and Feature Engineering

Figure 3.1a presents an example of the regular Redox cycling pattern for cell N4. To measure the J-V curve, the current density is raised gradually from zero to the maximum current density in specific levels and the voltage response is recorded. In EIS measurements, a sinusoidal current (I) or voltage (V) perturbation is applied to the fuel cell operating under the OCV condition, and the resulting voltage or current response is measured.

$$Z(j\omega) = \frac{|V(j\omega)|}{|I(j\omega)|} \exp(j(\phi - \psi)) \quad (3.1)$$

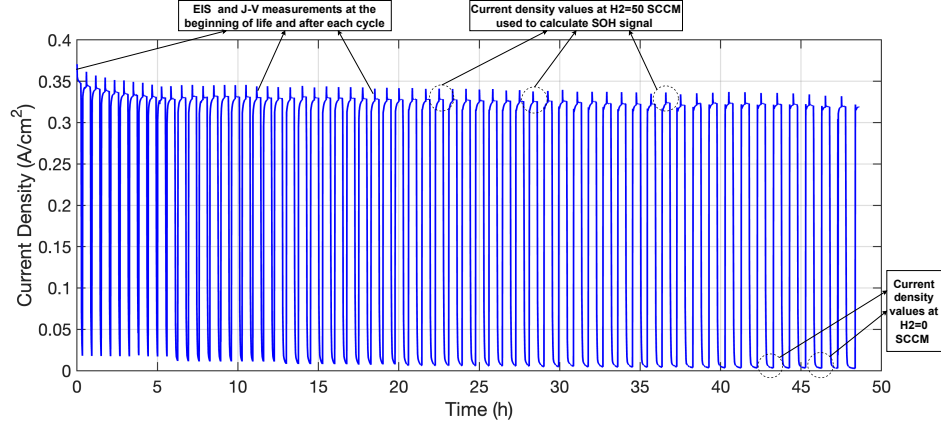
Where, j , ω , ϕ , and ψ are imaginary unit, angular frequency, phase of the voltage signal, and phase of the current signal, respectively. The ratio of the voltage to current ($Z(j\omega)$), defined in Eq. (3.1), can be represented in a Nyquist plot. The EIS measurements are typically conducted over a frequency (f) range of 1 MHz to 100 mHz, but can be adjusted based on the anticipated processes to be analyzed. In this study, EIS curves are measured in frequency range of 100 KHz to 0.1 Hz. To maintain system linearity, the signal amplitude must be sufficiently low to achieve a pseudo-linear response from the system [41–43].

Nonetheless, excessively small amplitudes can lead to a poor signal-to-noise ratio, necessitating the selection of a sufficiently high amplitude value to minimize signal noise. Various sinusoidal voltage amplitudes have been proposed in the existing literature, such as 10 mV [53, 54], 12 mV [55], 14 mV [54], and 50 mV [56]. Here, the sinusoidal voltage amplitude is 20 mV.

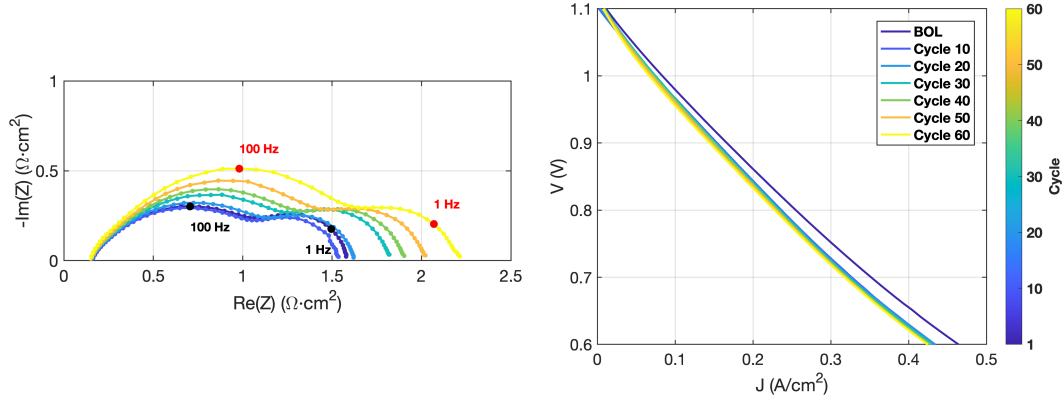
Figure 3.1 shows the evolution of EIS and J-V curves during the N4 Redox cycling experiment. In the EIS data (Figure 3.1b), the first semicircle increases with each cycle. After the 10th cycle, when the hydrogen cutoff interval was extended, the increase in impedance becomes more noticeable, showing an acceleration in degradation mechanisms. Similarly, the J-V curves (Figure 3.1b) show a consistent decrease in output power over time, confirming the progressive performance loss due to Redox cycling.

It should be noted that the curves in Figure 3.1 correspond to in-house fabricated tubular anode-supported SOFCs that were specifically operated under accelerated redox cycling conditions at a constant voltage of 0.7 V. These laboratory-scale cells were designed to generate a controlled and repeatable degradation dataset rather than to represent commercial performance benchmarks. As a result, the measured power densities ($0.4\text{-}0.8\text{ W/cm}^2$ at $650\text{-}750\text{ }^\circ\text{C}$) are lower than those of commercial SOFCs, which typically exhibit $0.8\text{-}2\text{ W/cm}^2$ [18, 137] with degradation rates below 1% per 1000 h and an EOL criterion of 10% power loss [138]. The accelerated redox protocol used here induces rapid degradation to capture clear impedance evolution patterns until the EOL condition is reached, and the resulting data are utilized for data-driven modeling.

Although the absolute power of our cells is lower, both impedance and power variations are clearly observable. Moreover, the SOH is defined as a normalized quantity relative to the cell’s initial performance, making it independent of the absolute power scale. Therefore, the proposed framework is applicable across different cell configurations and power ratings. In systems such as commercial SOFCs, where degradation progresses more slowly under normal aging conditions, the total test duration may need to be extended to observe



(a) Current density variation during Redox cycling for dataset N4 operated at a constant voltage of 0.7 V.



(b) Representative EIS and J-V curves collected at the beginning of life (BOL) and after selected Redox cycles. The legend and colorbar apply to both curves, indicating the corresponding Redox cycle number. The curves color is transitioning from blue (beginning of life) to yellow (end of life).

Figure 3.1: SOFC performance evolution during Redox cycling for dataset N4. (a) The current density profile shows repeated hydrogen interruption and recovery events at a constant operating voltage of 0.7 V. (b) The corresponding EIS and J-V curves after selected cycles illustrate progressive increases in polarization resistance and performance degradation. A unified color scheme and shared legend are used for both plots to emphasize the correspondence between electrochemical impedance and cell performance.

the evolution of the impedance spectra, while J-V and EIS measurements can be acquired less frequently. The method requires only two types of input data: (i) time-aligned J-V and EIS measurements collected periodically, and (ii) reference J-V or power data used to compute the SOH and define the EOL threshold.

Since the Redox test is performed at constant voltage $V=0.7$ V, by knowing the current density values during the test, along with its initial value, the SOH can be determined using [119, 121]:

$$\text{SOH} = \frac{J - J_E}{J_B - J_E}, \quad \text{valid for } H_2 = 50\text{SCCM} \quad (3.2)$$

The terms J , J_B , J_E represent the actual current density, the current density at the beginning of life, and the current density at EOL, respectively. An EOL criterion for performance is defined as a 10% decrease from the initial performance value signifies the EOL of the SOFC, $J_E = 0.9 \times J_B$ [102, 126]. By utilizing the J and J_B as a reference, the SOH value at each stage, ranging between zero and one, can be determined.

By analyzing the EIS curves, degradation mechanisms can be identified, offering valuable insights for designing suitable accommodation measures. However, these curves are complex and challenging to interpret. To interpret EIS curves, a commonly-used method is to employ an ECM that fits the measured impedance data. Then, ECM various components are attributed to certain behavior of the SOFC. Each of these components corresponds to physical and chemical phenomena such as electrochemical reactions, charge and mass transfer, and electrolyte resistance [44]. An optimization algorithm can be employed to determine the minimum number of elements necessary to accurately represent the processes occurring within a SOFC [64]. The DRT is also a valuable tool for deconvoluting phenomena with different time constants from measured impedance data. By using the DRT method, distinct peaks can be identified at specific frequencies, known as the characteristic frequencies corresponding to specific phenomena [22, 71, 72, 78, 80]. After DRT conversion, peak analysis needs to be performed to separate impedances. By multiplying the ratio of the

peak area to the DRT curve area by the overall cell polarization impedance, the contribution of that peak to the total polarization impedance can be determined. These peaks are represented by a series of resistor-capacitor (RC) elements or, more recently, resistor-constant phase element (R-CPE) elements in an ECM. DRT analysis can serve as an initial step in identifying a suitable ECM [65] or it can be applied independently to extract EIS features.

Figure 3.2 shows the DRT curves over time for N4 Redox dataset. The DRT calculations were conducted using the **DRTtools** Matlab code provided in [139], employing radial basis function (RBF) discretization and Tikhonov regularization. Choosing an appropriate regularization parameter (λ) in Tikhonov inversion is crucial for accurately resolving overlapping relaxation processes. Excessively large λ values over-smooth the DRT, merging distinct peaks and obscuring physically meaningful features, while too small λ values amplify measurement noise and introduce false peaks [66]. In this study, a fixed value was used for all EIS-to-DRT conversions ($\lambda = 10^{-3}$), consistent with prior work conducted by our research group using identical SOFC geometry, materials, and measurement setup [135]. That study demonstrated that this λ provides well-resolved peaks at the expected characteristic frequencies, representing the optimum for our cell configuration and setup. By maintaining λ constant, any variations observed in the DRT peaks can be attributed to the physical effects of Redox degradation rather than differences in numerical inversion parameters.

$$Z(\omega) = R_0 + Z_{\text{Pol}}(\omega) = R_0 + \int_0^\infty \frac{\gamma(\tau)}{1 + j\omega\tau} d\tau \quad (3.3)$$

The **DRTtools** code calculates γ function (Eq. (3.3)) for each of the EIS curves in Figure 3.1b, with the results shown in Figure 3.2. R_0 is ohmic resistance and the area under the γ function corresponds to the polarization impedance of the cell (Z_{Pol}). As shown in Figure 3.2, the shape and characteristic frequency of the peaks remain relatively constant, changes in peak heights are attributed to the corresponding impedance changes over time.

The identification of DRT peaks in this work follows the systematic approach previously established by members of our research group for the same tubular

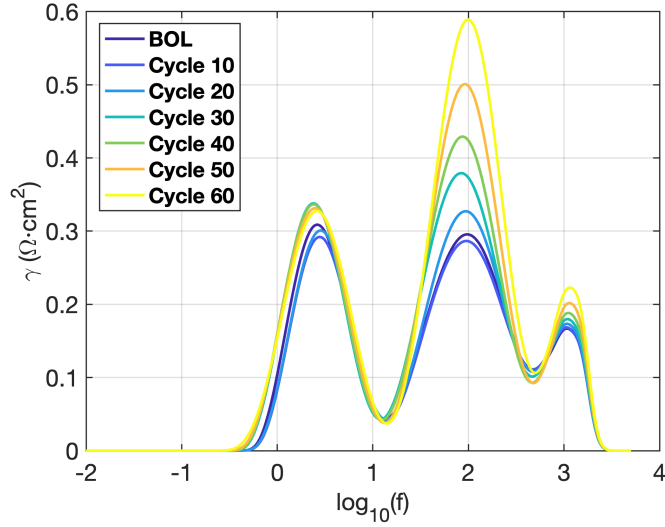


Figure 3.2: DRT evolution for the N4 dataset corresponding to the Redox cycling test shown in Figure 3.1. The DRT spectra are plotted for selected cycles using the same color scheme as the EIS and J-V curves in Figure 3.1b. The characteristic frequencies of the major peaks remain nearly constant over time, while their amplitudes increase, indicating the progressive growth of polarization losses associated with electrode processes. BOL stands for beginning of life.

SOFC platform [135]. In that study, all DRT inversion parameters were fixed including the regularization coefficient, discretization, and measurement setup to ensure that variations in peak behavior reflected true electrochemical effects rather than numerical artifacts. The physical origins of the peaks were then determined by sequentially varying one experimental factor at a time, such as cathode material, gas composition, flow rate, or temperature, while keeping all other parameters constant. Peaks that changed under cathode modification or oxygen partial pressure were attributed to cathode activation and gas diffusion processes, whereas those insensitive to such changes but temperature-dependent were assigned to charge transfer or ionic transport. This controlled perturbation methodology enabled clear attribution of each DRT feature to a distinct physical mechanism. Since the present study employs the same cell geometry, materials, and DRT configuration, and the DRT spectra exhibit similar characteristic frequency ranges, we rely on that work for assigning the observed peaks to their underlying processes.

For further validation, several studies in the literature investigating anode-supported tubular SOFCs have reported DRT peaks in comparable frequency ranges and assigned similar physical meanings [22, 62, 79, 140]. Based on these consistent observations, the DRT peaks in this study are interpreted as follows (in order of decreasing frequency):

- The first peak, with the highest characteristic frequency in the range of $10^3 - 10^4$ Hz, is associated with the oxygen ion transfer in the fuel and air electrodes.
- The second peak, with a characteristic frequency in the range of $10^2 - 10^3$ Hz, represents the charge transfer occurring at the anode/electrolyte interface, often referred to as anode activation polarization.
- The third peak, characterized by a frequency range of $10^1 - 10^2$ Hz, is indicative of charge transfer processes taking place at the cathode, referred to as cathode activation polarization.
- The fourth/fifth peak, with a characteristic frequency of $10^{-2} - 10^1$ Hz, is related to gas diffusion within the fuel and air electrodes and is commonly referred to as electrode concentration polarization.

By measuring the current density, SOH is calculated using Eq. (6.1). EIS measurements are taken at the start and after each Redox cycle. Since the hydrogen cutoff intervals at Redox cycles are different, there are different intervals between the SOH and EIS data samples. To provide uniformly sampled time-series data for training, a correlation between the extracted features from EIS and SOH is used to interpolate between the data points. For interpolating the temporal gaps in time-series data, a multi-output Gaussian process regression (MOGPR) [141, 142], is used here. MOGPR was selected due to its ability to model multivariate time-series correlations and provide uncertainty estimates, enhancing robustness in sparse data scenarios. MOGPR estimates mean and standard deviation at each requested point; however, only the mean

values, captured at a sample rate of 30 second, are used to create the time-series signal for training the forecasting model.

Most data-driven studies on SOH forecasting for SOFCs have been conducted using a single experimentally obtained degradation dataset from one SOFC system for both model training and testing [24, 114, 117, 120, 126]. These datasets represent a partial degradation segment rather than the full life of the cell. Some other works have employed a few more datasets, such as five in [115, 116], but these also cover limited degradation periods and do not reach the EOL. In comparison, the present study utilizes eight independent datasets, which is above typical practice, to demonstrate the generalization and effectiveness of the proposed method across operating conditions and degradation severities within the same mechanism (Redox). The eight datasets were obtained at three operating temperatures (650, 700, and 750 °C) and under both normal and randomly generated redox cycling intervals, covering the complete life cycle of each cell from beginning to end of the cell life.

3.3 Physics-informed Machine Learning Model

Through periodic measurements of SOFC output power and impedance, the SOH of a cell can be determined, along with the underlying mechanisms contributing to power degradation. Using the historical SOH and impedance data, it is possible to forecast the SOH until the next scheduled inspection. To integrate the impedance information and SOH, a multivariate time-series forecasting model is needed to consider both cross-series correlations and temporal dependencies. Physical knowledge can help guiding the model to capture the cross-series relationships [143]. Variety of PIML strategies exist to integrate physical knowledge into machine learning models applied in energy systems. In this study, physical knowledge was incorporated through the model architecture rather than through an explicitly physics-constrained loss function. This choice was motivated by the limited availability of sufficiently established and generalizable mathematical formulations describing SOFC degradation

dynamics under the investigated operating conditions. Consequently, defining a reliable degradation-physics residual for inclusion in the loss function would require assumptions that may not adequately represent the underlying electrochemical processes. In contrast, the selected impedance-derived features have clear physical interpretations and represent interacting electrochemical phenomena whose evolution is coupled during degradation. Therefore, a graph-based architecture was adopted to encode these physically meaningful dependencies and allow the model to learn the interactions among the features. This formulation introduces physics through inductive bias in the network structure while avoiding the imposition of potentially incomplete or oversimplified degradation equations.

A GCN-LSTM model is adapted here to make it possible to incorporate the physics in designing the network. This GCN-LSTM model contains a GCN that captures the cross-series correlations followed by LSTM network to model the temporal relationships. The GCN is a graph-structured network including the nodes connected through edges. Each node in the graph represents one of the time-series such as impedances and SOH. The edges between these nodes are established based on the physical relationships of the phenomena represented by the nodes, and helps the network to capture cross-series correlations effectively.

The information is passing between the connected nodes, called neighbors, through message-passing iterations or layers. At each message-passing layer of a GCN, $h_u^{(k)}$ is the *hidden embedding* corresponding to each node $u \in \mathcal{V}$ and will be updated based on the aggregated information from the node's neighborhood, $\mathcal{N}(u)$, expressed as follows [144, 145]:

$$\begin{aligned} h_u^{(k+1)} &= f_U^{(k)} \left(h_u^{(k)}, f_A^{(k)} \left(\{ h_v^{(k)} \mid \forall v \in \mathcal{N}(u) \} \right) \right) \\ &= f_U^{(k)} \left(h_u^{(k)}, m_{\mathcal{N}(u)}^{(k)} \right) \end{aligned} \quad (3.4)$$

The f_U and f_A are Update and Aggregate, arbitrary differentiable functions. $m_{\mathcal{N}(u)}$ is the message that is aggregated from u 's graph neighborhood $\mathcal{N}(u)$. Superscripts are used to demonstrate the message-passing iterations, sometimes known as message-passing layers. Considering the iteration k , the f_A

takes the set of embedding of the neighbor's information as input and calculate the message $m_{\mathcal{N}(u)}^{(k)}$. Then, f_U combines this message with the previous embedding of the node, $h_u^{(k-1)}$, to generate the updated embedding $h_u^{(k)}$. The input features of all nodes are the initial embeddings at $k = 0$. It means $h_u^{(0)} = x_u, \forall u \in \mathcal{V}$. After applying k message-passing layers, the output of the final layer is the embedding for each node, $h_u^{(k)} = z_u, \forall u \in \mathcal{V}$. What is defined in Eq. (3.4) is an abstract framework for message-passing process and can be translated into something that can be implemented [144, 145]:

$$h_u^{(k)} = \sigma \left(W_{\text{self}}^{(k)} h_u^{(k-1)} + W_{\text{neigh}}^{(k)} \sum_{v \in \mathcal{N}(u)} h_v^{(k-1)} + b^{(k)} \right) \quad (3.5)$$

where $W_{\text{self}}^{(k)}, W_{\text{neigh}}^{(k)} \in \mathbb{R}^{d^{(k)} \times d^{(k-1)}}$ are trainable matrices, σ is activation function, and $b^{(k)}$ is the bias term. The message-passing frameworks relies on linear operations followed by a single element-wise non-linearity similar to standard multi-layer perceptron (MLP). It can be equivalently defined through the f_U and f_A functions [144, 145]:

$$m_{\mathcal{N}(u)} = \sum_{v \in \mathcal{N}(u)} h_v \quad (3.6)$$

$$f_U(h_u, m_{\mathcal{N}(u)}) = \sigma(W_{\text{self}} h_u + W_{\text{neigh}} m_{\mathcal{N}(u)}) \quad (3.7)$$

$$m_{\mathcal{N}(u)} = f_A^{(k)}(h_v^{(k)}, \forall v \in \mathcal{N}(u)) \quad (3.8)$$

The fundamental concept behind the GCN message-passing framework is quite simple: with each iteration, nodes gather information from their immediate neighbors. As iterations continue, the information within each node embedding expands to include details from more distant parts of the graph. In the first iteration ($k = 1$), a node embedding captures information from its 1-hop neighborhood, which includes features from directly connected neighbors. By the second iteration ($k = 2$), the embedding extends to cover the 2-hop neighborhood. Generally, after k iterations, a node embedding contains insights from its k -hop neighborhood, progressively integrating information from farther reaches of the graph [146].

The connections between the nodes are established based on the physical relationships and locations of the electrochemical phenomena represented by the nodes. When degradation occurs in a given region (such as the anode or cathode), all impedance elements associated with that region are affected. For example, edges are formed between nodes representing charge transfer, oxygen ion transfer, and gas diffusion within the cathode, as these processes are physically interconnected, and the same applies to anode-related impedances. The ohmic impedance is associated with the electrolyte and current collectors and arises from ionic and electronic transport across these components. Finally, because all these localized impedances contribute collectively to overall performance degradation, they are connected to the SOH node. The detailed connections are as follows:

- **Nodes 1, 2, and 4** are connected because they all represent phenomena occurring within the anode region. Node 1 (oxygen ion transfer), node 2 (charge transfer), and node 4 (gas diffusion) are physically co-located processes; hence, degradation in one affects the others.
- **Nodes 1, 3, and 4** are connected as they correspond to phenomena in the cathode. Node 1 (oxygen ion transfer), node 3 (charge transfer), and node 4 (gas diffusion) interact through oxygen partial pressure and reaction kinetics at the cathode-electrolyte interface, influencing each other during degradation.
- **Node 5** (ohmic resistance) represents the electrolyte and current collectors and is not directly linked to other impedance nodes. It mainly reflects bulk ionic and electronic conduction, independent of electrode-specific reactions.
- **Node 6** (SOH) is connected to all other nodes (1-5) because each localized impedance process contributes to the overall degradation of cell performance and long-term health evolution.

These connections are defined using an $n \times n$ adjacency matrix, composed of zeros and ones. Here, n represents the number of nodes, and if node i and node j are connected, the matrix element a_{ij} is set to one; if they are not connected, a_{ij} is zero. Figure 3.3 illustrates the designed graph along with a simplified schematic of the message-passing process. When deciding the number of message-passing layers or iterations, it's crucial to consider the system's complexity and the interactions between ongoing phenomena.

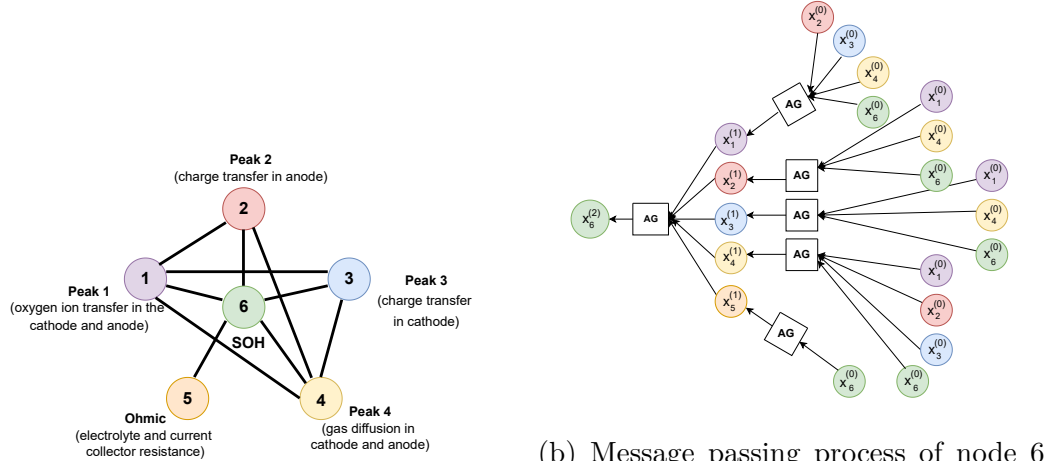
Two message-passing iterations are adopted to allow each node to exchange information with both its direct and second-order neighbors, enabling indirect electrochemical interactions to be represented without explicitly connecting every node pair. For example, node 2 (anode activation) and node 3 (cathode activation) are not directly connected, and their interaction occurs indirectly through node 6 (SOH), which represents the overall cell performance. This reflects the fact that changes in activation behavior at one electrode influence the total voltage and therefore the operating state of the opposite electrode. Similarly, node 5 (ohmic resistance) is not directly connected to the activation or concentration nodes but interacts with them indirectly through its effect on the cell voltage. Physically, higher ohmic resistance reduces the effective voltage across the electrodes, forcing both the anode and cathode to operate at higher overpotentials to sustain the same current. This accelerates activation and concentration polarization losses. Thus, although the ohmic process is treated as an independent node, its influence propagates through the network via the SOH node and through message-passing layers. This two-layer design allows the model to capture both local and global dependencies within the SOFC. In other words, while the primary graph connections are defined based on component-level relationships (e.g., anode, cathode, electrolyte), the two message-passing layers enable the model to also account for cross-component electrochemical phenomena. Increasing the number of layers beyond two can lead to over-smoothing, where node representations become indistinguishable in such a dense graph and physically meaningful separations begin to blur. It may also cause overfitting given the limited dataset size. Thus, two message-

passing iterations provide the optimal balance between physical relevance, information propagation, and numerical stability.

While this study focuses on Redox cycling, the proposed DRT-to-graph framework is inherently extensible to other degradation mechanisms. In cases such as cathode degradation, the electrochemical processes may overlap or evolve differently, leading to shifts in characteristic frequencies and merging of adjacent peaks [82]. These phenomena may alter both the number and interrelation of DRT peaks, requiring corresponding modifications in the graph design, such as redefining nodes or updating edge connections to reflect new physical dependencies. In cases where an additional polarization process arises (for example, when another gas such as methane is added to hydrogen), a new concentration polarization node may be introduced and connected to the anode-related nodes. In the present work, the graph structure remains static, and only node features (peak amplitudes) evolve over time. For degradation mechanisms involving dynamic interactions or peak merging, dynamic graph neural network frameworks [147] can be employed, where both node values and connectivity evolve over time. All DRT analyses in this study were performed using a fixed regularization parameter (λ) reported in Section 3.2, ensuring that any changes in DRT features stem from the degradation process itself rather than mathematical conversion.

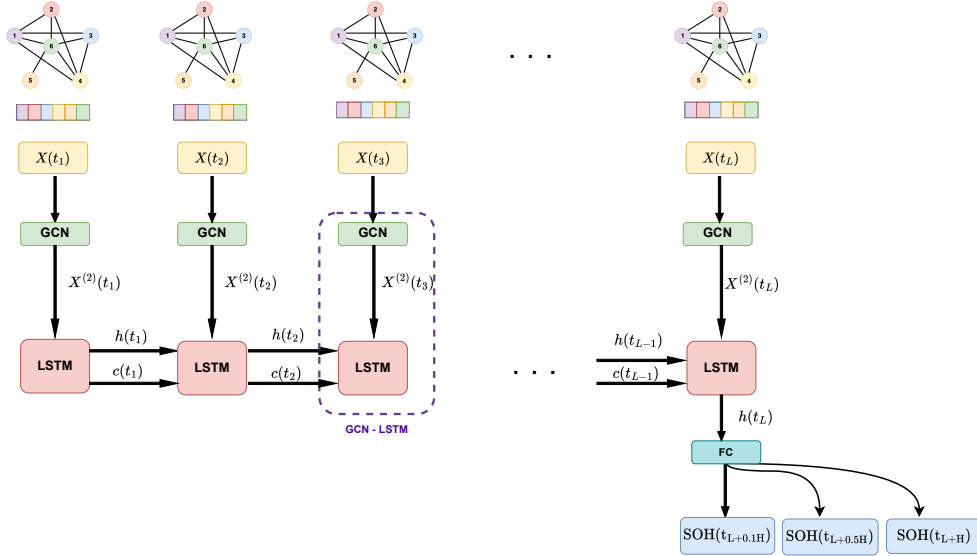
This DRT-to-graph mapping is conceptually generalizable to SOFCs with different geometries or materials as well. Geometry and material composition influence electrochemical time constants, potentially shifting characteristic frequencies or altering the number of observable DRT peaks [77]. If the same electrochemical processes remain distinguishable, the existing graph structure can be kept and retrained on the new data to capture updated node features. However, if peak merging, splitting, or new processes emerge due to geometry or material differences, the graph topology should be revised accordingly.

GCN-LSTM can consider both the cross-series correlations and temporal relationships. Figure 3.3c shows the GCN-LSTM framework implemented here, including two main components: a GCN and an LSTM. LSTMs leverage gated



(a) Graph structure containing polarization impedances associated with DRT peaks, ohmic resistance, and SOH as nodes. The connections are designed based on the physical phenomena associated with each node.

(b) Message passing process of node 6 as an example. AG: Aggregate function. $x_i^{(0)}$: initial values of the node features. $x_i^{(1)}$: updated values of the node features after first message passing layer. $x_i^{(2)}$: updated values of the node features after the second message passing layer.



(c) Schematic of proposed GCN-LSTM to forecast SOH of SOFCs. h : LSTM hidden state, c : LSTM cell state, X : Input feature vector, $X^{(2)}$: Embedded feature vector after two message passing iterations, SOH: Output vector, L: look-back, H: forecast horizon.

Figure 3.3: Overview of the proposed physics-informed GCN-LSTM model: (a) graph structure linking impedance components and SOH, (b) two-layer message passing for feature embedding, and (c) temporal forecasting architecture integrating GCN features with LSTM dynamics.

connections, enabling them to learn which features within a sequence are essential and should be retained, while discarding less relevant ones. Unlike recurrent neural networks (RNNs), LSTMs are less prone to the vanishing or exploding gradient problem and are designed to process long-term dependencies within sequential data [148]. The combined model uses 3-hour historical time series data as input and employs the GCN to calculate the embedded features. Then, the embedded time-series data is fed into the LSTM model, where temporal relationships are captured. Ultimately, the results are derived via the fully connected (FC) layer. The output focuses on predicting SOH values at specific future time steps. These predictions include points at 10%, 50%, and the full extent of the forecast horizon which is 6 hours. The model is trained using datasets of N1 to N6 and R1, then tested on dataset R2, which has an approximate cell lifetime of 50 hours.

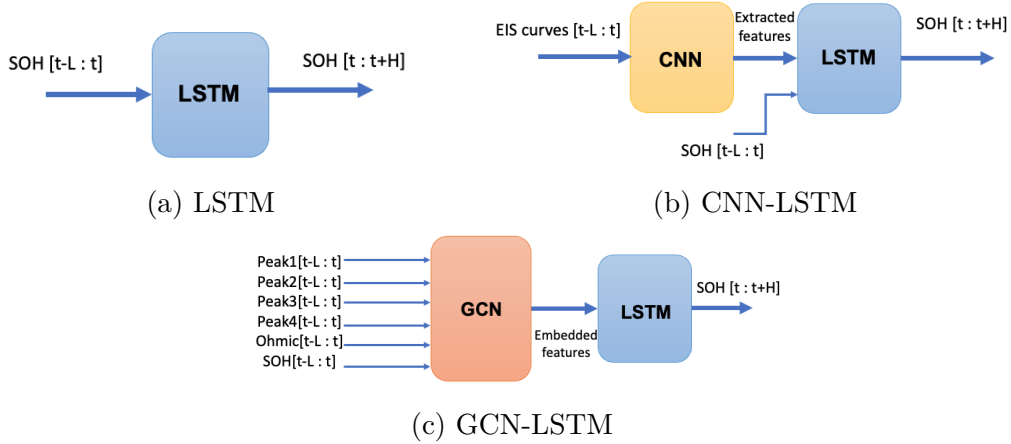


Figure 3.4: Inputs and outputs of three studied models. L: look-back, H: forecast horizon

As shown in Figure 3.4, time-series forecasting models can be employed for this purpose, utilizing SOH alone or the combined dataset of SOH and impedance. Two other forecasting models are introduced considering different scenarios of how to use SOH and impedances data and compared to GCN-LSTM model. The first model (Figure 3.4a) uses only historical SOH data to forecast future SOH values. An LSTM model is trained on this sequential data, enabling it to predict future values based on past observations. The next

model (Figure 3.4b) integrates raw EIS data with SOH for forecasting. This provides the model with rich information, potentially increasing its forecasting accuracy compared to the LSTM model that relies solely on SOH data. A convolutional neural network (CNN) is used to extract features from the raw EIS data. These extracted features are then combined with the SOH signal, resulting in a multivariate time-series data, then fed to an LSTM network to capture the temporal relationships.

The CNN-LSTM and GCN-LSTM models differ in input format but were trained using the same datasets and temporal alignment settings to ensure a fair comparison. Both models were developed from the same underlying SOH and EIS data. In the GCN-LSTM, DRT-derived features (ohmic resistance and peaks 1-4) were interpolated using MOGPR as described in Section 3.2. In the CNN-LSTM, raw EIS spectra were upsampled to match the same time steps as SOH before being processed by the CNN. Both models used the same look-back window, ensuring that each LSTM received equivalent temporal information. This setup isolates the architectural difference between CNN and GCN, allowing for a meaningful performance comparison.

While CNNs are powerful tools for learning local patterns in multidimensional data, they are inherently designed to capture spatial locality like correlations between neighbor elements in a grid such as pixels in an image or frequency bins in a spectrum. In contrast, the DRT-derived peaks and ohmic resistance are not organized on a uniform spatial grid but correspond to distinct electrochemical phenomena distributed across the physical structure of the SOFC. These include oxygen ion transfer (peak 1), charge transfer (peaks 2 and 3), gas diffusion (peak 4) in the anode and cathode, and ohmic resistance. The relationships among these processes are physically defined, motivating the use of a graph representation to model their interactions. As shown in Figure 3.3a, edges in the graph are constructed based on electrochemical understanding and the physical location of the processes within the cell. For instance, nodes corresponding to phenomena co-located in the cathode or anode are connected. This physically informed topology enables the GCN to learn not only from

impedance features but also from their defined physical interdependencies, allowing the network to capture mechanism-level correlations that govern SOH evolution.

A brief overview of the LSTM and CNN-LSTM architectures is provided below.

LSTM: LSTMs can be applied in sequence-to-sequence (Seq2Seq) modeling. As shown in Figure 3.4a, a sequence of past SOH values is fed into a 1-layer LSTM network. Unlike typical Seq2Seq models that output entire sequence, specific points in the future sequence are chosen as the LSTM’s direct output. These predictions include points at 10%, 50%, and the full extent of the forecast horizon. To enable forecasting at various future time steps, a FC layer follows the LSTM layer, providing predictions for the SOH at those specific points.

CNN-LSTM: In the CNN-LSTM model, feature extraction precedes the windowing process. The EIS data is first over-sampled to ensure uninformed and equal number of samples as the SOH data. The raw EIS data represented as a two-dimensional matrix of size 60×2 (real and imaginary components at 60 frequency points), is treated as an image-like input. This matrix is fed into a sequence of two convolutional layers, each containing 32 filters with a kernel size of 3×3 and employing the ReLU activation function. These convolutional layers are followed by two max-pooling layers with a pooling size of 2×2 , facilitating the reduction of data dimensionality and capturing local features within the EIS data. A flattening layer then transforms the pooled feature maps into a one-dimensional vector, which is subsequently fed into a FC layer to generate a 5-dimensional output vector. As shown in Figure 3.4b, this vector, containing the extracted features from the EIS data, is then concatenated with the corresponding SOH value at a specific time step, resulting in a 6-dimensional vector that incorporates both SOH and the learned EIS features. This process of feature extraction and concatenation is repeated for each time step in the EIS and SOH sequences. By stacking these resulting 6-dimensional vectors, a sequential data structure is created that serves as input

to the LSTM network. The LSTM leverages this input, containing past values of both SOH and the learned EIS features, to predict future SOH values.

3.4 Results and Discussion

3.4.1 Hyper-parameter Optimization

The hyperparameters for the LSTM, CNN-LSTM, and GCN-LSTM models were selected through a random search strategy within defined search spaces. To simplify the optimization process, the **Adam** optimizer and **ReLU** activation function were selected as fixed parameters for all models, as these are widely adopted for time-series forecasting tasks. The look-back window length, batch size, and learning rate values optimized in the baseline LSTM model were kept constant for the CNN-LSTM and GCN-LSTM models. The forecast horizon can be adjusted by the user based on the required prediction accuracy, as detailed in section A.5. Additionally, the number of graph convolutional (message-passing) layers in the GCN-LSTM was set to two, guided by the underlying physics of SOFC degradation phenomena, as explained in section A.3.

Tables 5.4-3.3 summarize the remaining hyperparameters, their respective search spaces, and the optimized values for each model. For each case, the total number of possible hyperparameter combinations and the number of random search trials—set to 30% of the total combinations—are also reported.

Table 3.1: LSTM model hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of LSTM layers	{1, 2}	1
Number of LSTM units	{32, 64, 128}	32
Learning rate	{1e-2, 1e-3, 1e-4}	1e-3
Batch size	{16, 32, 64}	32
Look-back window length	{240, 360, 480}	360

Total combinations: 162. Random search trials: 49.

Table 3.2: CNN-LSTM model hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of Conv layers	{1, 2, 3}	2
Kernel size	{(3×3), (5×5)}	(3×3)
Pooling size	{(2×2), (3×3)}	(2×2)
Number of LSTM layers	{1, 2}	1
Number of LSTM units	{32, 64, 128}	64
Total combinations: 72. Random search trials: 22.		

Table 3.3: GCN-LSTM model hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Aggregation function	["sum", "mean", "max"]	"mean"
Update function	["concat", "add"]	"add"
Number of LSTM layers	{1, 2}	1
Number of LSTM units	{32, 64, 128}	64
Total combinations: 36. Random search trials: 11.		

3.4.2 Model Comparison

While the RUL estimation is essential to determine when to replace a cell, having a model to forecast the SOH at specific future time horizons, well before reaching EOL, is equally valuable. It allows operators to optimize operating conditions, schedule maintenance in advance, and mitigate sudden performance drops.

The SOH forecasting models are trained using datasets N1 to N6 and R1, and evaluated on the unseen dataset R2. A look-back window of 3 hours and a forecast horizon of 6 hours are applied across all models. The forecasting results are illustrated in Figure 3.5, and the corresponding root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) values are summarized in Table 3.4.

The GCN-LSTM network demonstrates the best forecasting performance among the evaluated models. Across all error metrics (RMSE, MAE, and

Table 3.4: Prediction accuracy comparison of the three SOH prediction models on R2 dataset, with look-back of 3 hours and forecast horizon of 6 hours. MAPE computed on samples with SOH > 0.05.

Model	RMSE	MAE	MAPE (%)	R^2	% RMSE ↓	% MAE ↓
LSTM (baseline)	0.134	0.108	14.5	0.84	—	—
CNN-LSTM	0.117	0.094	12.6	0.88	13	13
GCN-LSTM	0.085	0.067	8.9	0.94	37	38

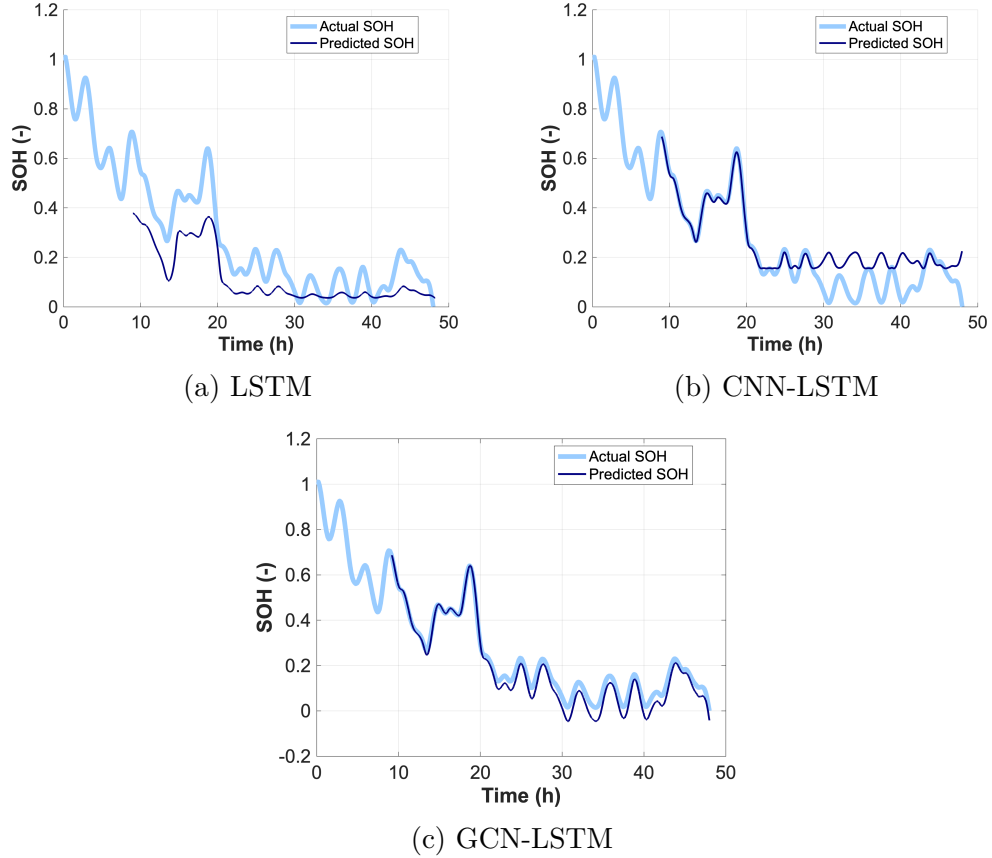


Figure 3.5: SOH forecasting results for LSTM (a), CNN-LSTM (b), and GCN-LSTM (c) on the R2 dataset.

MAPE), both impedance-aware models outperform the LSTM: CNN-LSTM yields 13% lower RMSE/MAE/MAPE, while GCN-LSTM achieves 37 - 39% reductions relative to LSTM. Consistently, R^2 increases from 0.84 for the LSTM model to 0.88 for CNN-LSTM and 0.94 for GCN-LSTM, indicating the strongest fit for GCN-LSTM. Moreover, GCN-LSTM improves upon CNN-

LSTM by 27 - 29% across the error metrics. These results highlight that DRT-based impedance features combined with a physics-informed graph-structured temporal architecture substantially improve SOH prediction.

The improved forecasting accuracy obtained by the physics-informed GCN-LSTM model is attributed to its incorporation of physical knowledge regarding SOFC degradation mechanisms. Each node in the developed graph-structured network represents specific physical phenomena inside the cell such as oxygen ion transfer, charge transfer, and mass transport, extracted from EIS data through DRT analysis. The connections between the nodes capture physically meaningful correlations among different phenomena and degradation mechanisms. Thus, the forecasting improvements show the model’s capability to learn and generalize underlying degradation physics rather than just relying on data-driven correlations.

As the forecast horizon decreases, the GCN-LSTM model yields progressively better results. The model is designed to predict SOH values at specific future time points of 0.6, 3, and 6 hours, corresponding to 10%, 50%, and 100% of the forecast window. The prediction outcomes for these different horizons are presented in Figure 3.6, with the corresponding RMSE, MAE, MAPE, and R^2 values reported in Table 3.5. This comparison reveals a trade-off between forecast accuracy and prediction horizon length, enabling users to select a forecasting strategy aligned with their specific operational requirements.

Table 3.5: Prediction accuracy comparison of the GCN-LSTM model at different points of forecast horizon H . MAPE computed on samples with SOH > 0.05 .

H (h)	RMSE	MAE	MAPE (%)	R^2
0.6	0.057	0.045	6.1	0.97
3	0.077	0.061	8.0	0.95
6	0.085	0.067	8.9	0.94

To ensure a fair comparison with the proposed GCN-LSTM framework, additional baseline models including multivariate LSTM, gated recurrent unit

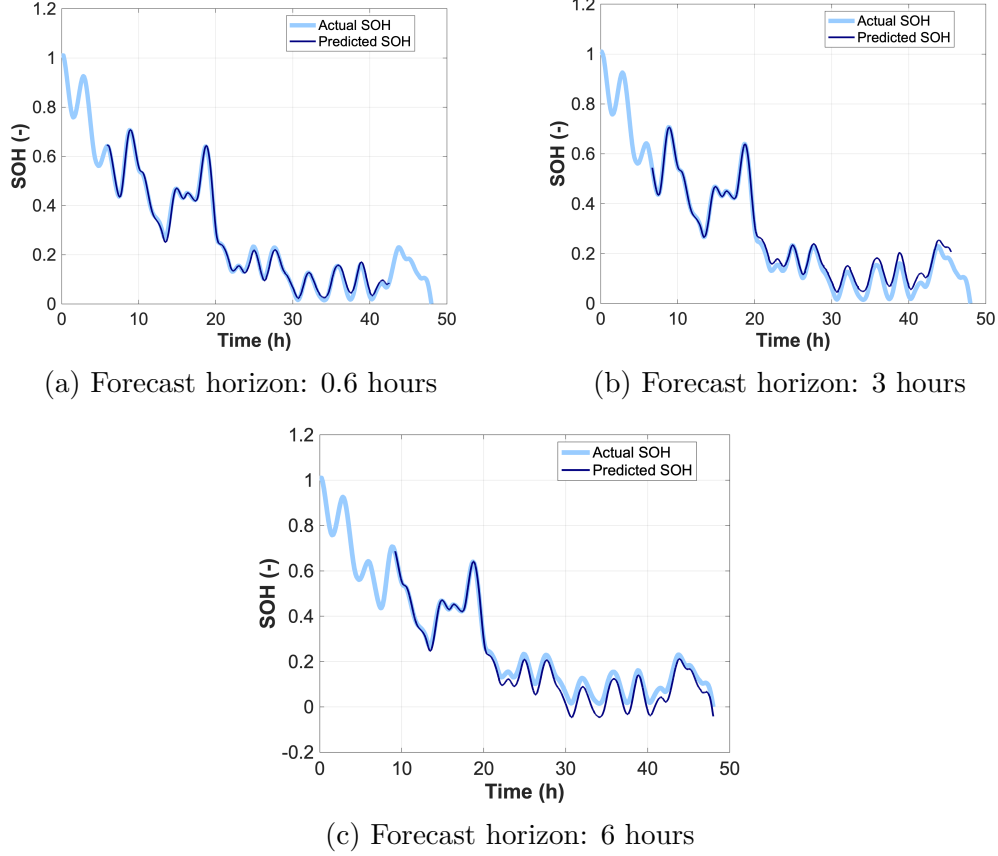


Figure 3.6: Results of GCN-LSTM network in SOH forecasting with different forecast horizons.

(GRU), and temporal convolutional network (TCN) were developed using the same input features and training procedure. All models were trained using the identical EIS-derived impedance features together with the SOH signal. In addition, the same preprocessing pipeline and temporal forecasting configuration were adopted for all models, where a 3-hour historical window was used to predict SOH evolution up to 6 hours ahead. Therefore, the comparison isolates the effect of the network architecture itself. The resulting performance metrics are summarized in Table 3.6. Detailed descriptions of the model architectures, hyperparameter optimization procedures, selected hyperparameters, and SOH prediction trajectories for the benchmark models are provided in Appendix B. To evaluate whether the performance improvements of the proposed GCN-LSTM were statistically significant, one RMSE value was calculated for each

of 50 paired prediction trajectories generated by every model. The trajectory-level RMSE values of the GCN-LSTM were then compared with those of each benchmark using a one-sided paired Wilcoxon signed-rank test. The null hypothesis assumed no systematic reduction in RMSE, whereas the alternative hypothesis stated that the GCN-LSTM produced lower trajectory-level RMSE than the corresponding benchmark. A significance level of $\alpha = 0.05$ was adopted; therefore, a p -value below 0.05 indicates that the improvement is statistically significant, while a p -value below 0.01 is commonly interpreted as stronger evidence against the null hypothesis [149]. The obtained p -values of 0.010, 0.017, and 0.031 for the comparisons with the multivariate LSTM, GRU, and TCN, respectively, are all below 0.05. These results indicate that the GCN-LSTM achieves statistically significantly lower trajectory-level RMSE than all three benchmark models, supporting its improved SOH forecasting performance.

Table 3.6: Performance comparison of SOH forecasting models. The reported p -values are obtained from one-sided paired Wilcoxon signed-rank tests comparing the trajectory-level RMSE of each benchmark model with that of the GCN-LSTM.

Model	RMSE	MAE	MAPE (%)	R ²	p -value (vs. GCN-LSTM)
Multivariate LSTM	0.093	0.074	10.2	0.92	0.010
GRU	0.090	0.072	9.9	0.93	0.017
TCN	0.088	0.069	9.3	0.93	0.031
GCN-LSTM	0.085	0.067	8.9	0.94	—

To further evaluate the generalization capability of the proposed GCN-LSTM model across diverse Redox cycling conditions, a cross-dataset validation was performed, where the training and testing datasets were systematically varied. The model was trained on seven datasets and tested on the remaining one, and this process was repeated eight times so that each dataset served once as the test set. As summarized in Table 3.7, the model achieves consistent performance across all datasets (RMSE = 0.070-0.085, R² = 0.94-

0.95), confirming generalization under varying Redox cycle designs and temperatures. The slightly higher error for R2 is attributed to its stronger SOH fluctuations.

Table 3.7: Cross-dataset validation results of GCN-LSTM trained on seven datasets and tested on the remaining one.

Test Dataset	RMSE	MAE	MAPE (%)	R^2
N1	0.070	0.055	7.5	0.95
N2	0.072	0.058	7.8	0.95
N3	0.080	0.063	8.4	0.94
N4	0.071	0.057	7.8	0.95
N5	0.078	0.061	8.2	0.94
N6	0.071	0.055	7.6	0.95
R1	0.082	0.065	8.6	0.94
R2	0.085	0.067	8.9	0.94

For deployment, offline dataset preparation is separated from online inference. The MOGPR oversampling used to regularize the dataset is an offline step and is not required during operation. In practice, the model uses streaming SOH signals, while EIS curves are collected periodically. For GCN-LSTM, a DRT feature vector is recomputed when an EIS curve is taken, whereas in CNN-LSTM, convolutional features are refreshed from each new EIS curve. Between EIS updates, the most recent feature vector is reused and only the neural model forward pass is executed.

Table 3.8 reports model inference time per sample ($t_{\text{inference}}$), CNN and DRT feature extraction time (t_{feature}), EIS acquisition time (t_{EIS}), and EIS update interval (t_{interval}). The collection of one EIS curve takes approximately 2 minutes, determined by frequency range and instrument settings. The EIS update interval is not regular and depends on the Redox cycling profile, since each Redox cycle ends with a return to OCV for EIS measurement. The reported 278 ms for DRT feature extraction on a Mac 8-Core CPU is negligible relative to the 2-minute EIS acquisition. CNN feature extraction (3.0 ms) is embedded

within the forward pass but reported separately for transparency. The process is shown in Figure 3.7.

Table 3.8: Online computational profile on NVIDIA Tesla T4 GPU for model inference ($t_{\text{inference}}$). For CNN-LSTM, convolutional feature extraction is reported separately as t_{feature} , although it is embedded within the model forward pass. For GCN-LSTM, DRT feature extraction is performed using a MATLAB implementation on a Mac 8-Core CPU, requiring 278 ms per EIS curve.

Model	t_{model} (ms)	t_{feature} (ms)	t_{interval} (min)	t_{EIS} (min)
LSTM	1.3	—	—	—
CNN-LSTM	4.5	3.0	10-35	2
GCN-LSTM	2.6	278	10-35	2

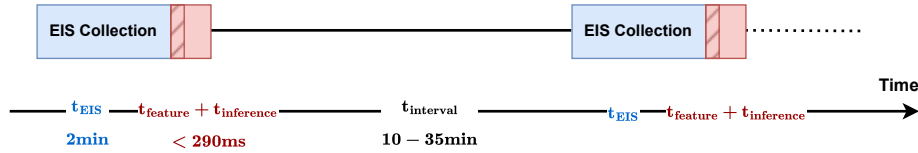


Figure 3.7: Timeline of online inference workflow. Each EIS measurement requires 2 minutes (t_{EIS}), followed by feature extraction and model inference ($t_{\text{feature}} + t_{\text{inference}} < 290$ ms). The interval between successive EIS updates (t_{interval}) depends on the Redox cycling profile (10-35 minutes). Since feature extraction and inference latency are negligible compared to EIS acquisition, the limiting factor for real-time deployment is the EIS measurement itself.

These results confirm that inference latency is negligible (< 5 ms for all models). Even when feature extraction is included, the computational cost remains far smaller than EIS acquisition time. The limiting factor in practice is the EIS acquisition time and update interval, not model execution, demonstrating that the proposed approach is compatible with real-time SOFC monitoring.

3.5 Summary of Chapter

This chapter demonstrates the effectiveness of a physics-informed TGCN for real-time SOH estimation of SOFCs. The proposed forecasting methodology

was comprehensively validated using eight datasets collected under varying Redox cycling severities and operating conditions. A comprehensive SOFC degradation dataset of over 2.1 million samples has been made publicly available to support further research and validation by the community. The main conclusions and key findings of this study are summarized as follows:

- The physics-informed TGCN integrates impedance data with voltage and current density measurements, improving real-time SOH forecasting accuracy for SOFCs under Redox cycling. Incorporating impedance features substantially enhances forecasting performance across multiple metrics. CNN-LSTM and GCN-LSTM models achieve RMSE reductions of 13% and 37%, respectively, compared to a baseline model that used only SOH/voltage history. The superior performance of the GCN-LSTM model is attributed to its physics-informed graph-structured architecture, which effectively captures the underlying degradation physics in SOFCs.
- The GCN-LSTM model was further validated using cross-dataset testing, demonstrating strong generalization capability across eight datasets collected under different Redox cycling severities and operating temperatures (RMSE = 0.070-0.085, $R^2 = 0.94$ -0.95).
- Computational analysis confirms that inference latency is negligible for all models (<5 ms on GPU), while the feature-extraction step adds only minor overhead compared to the 2-minute EIS acquisition time, confirming the feasibility of real-time deployment.

Chapter 4

Remaining Useful Life Forecasting with Uncertainty¹

4.1 Introduction

Building upon the physics-informed SOH estimation framework developed in Chapter 3, this chapter extends the analysis from short-term health monitoring to long-horizon prognostics. While accurate SOH estimation provides valuable insight into the current condition of SOFC systems, practical deployment requires forecasting the evolution of SOH over time and estimating the RUL until EOL. This task is inherently more challenging, as it involves predicting degradation trajectories over extended horizons under varying operating conditions.

A key difficulty in long-horizon forecasting arises from the dynamic nature of the prediction window. As the SOFC ages, the remaining time to failure decreases, resulting in shrinking forecast horizons that complicate model training and degrade prediction stability. In addition, conventional prognostic models typically provide deterministic predictions, which limits their usefulness in real-world applications where uncertainty plays a critical role in decision-making and risk assessment.

To address these challenges, this chapter proposes a probabilistic prognostics framework that forecasts the full SOH trajectory up to end-of-life while explicitly quantifying uncertainty. The approach leverages Bayesian sequence

¹This chapter is partially based on the following publication: [2, 3].

models, including attention-based Informer and LSTM architectures, to capture long-range temporal dependencies in degradation data. Both aleatoric and epistemic uncertainties are modeled to generate calibrated confidence intervals for SOH predictions and the corresponding RUL estimates.

Furthermore, a dynamic forecasting strategy is introduced to handle variable prediction horizons using an exponentially weighted zero-padding technique. This approach enables the use of a single model to accommodate sequences of different effective lengths while guiding predictions toward physically consistent end-of-life (EOL) behavior. In addition, consistent with the approach presented in the previous chapter, impedance-derived features are incorporated alongside SOH information to enhance the representation of degradation dynamics.

The proposed framework is validated using the same experimental datasets introduced and processed in the previous chapter. The results demonstrate that the Bayesian Informer model achieves high prediction accuracy while providing well-calibrated uncertainty estimates, enabling reliable RUL forecasting. This chapter therefore advances the prognostic capability of SOFC health monitoring by extending SOH estimation to long-horizon, uncertainty-aware lifetime prediction, forming a critical step toward robust predictive maintenance and health-aware control strategies.

4.2 RUL Forecasting Models

This study designs two Bayesian neural architectures: a Bayesian LSTM (Fig. 5.3) and a Bayesian Informer (Fig. 4.2). The Bayesian LSTM used in this study is shown in Fig. 5.3 and takes a multivariate input sequence composed of the SOH, Ohmic resistance, and DRT peaks measured from the cells. The input is processed through a two-layer stacked LSTM architecture, with a dropout layer in between the layers to provide stochastic regularization for uncertainty quantification through MC dropout. The outputs of the second LSTM layer are then passed to a fully connected (FC) layer that di-

rectly produces the forecast SOH values for all future time-steps in one shot. This architecture allows modeling of complex temporal dependencies present in multivariate SOFC degradation data while also enabling the estimation of both aleatoric and epistemic uncertainties. The Bayesian Informer, on the

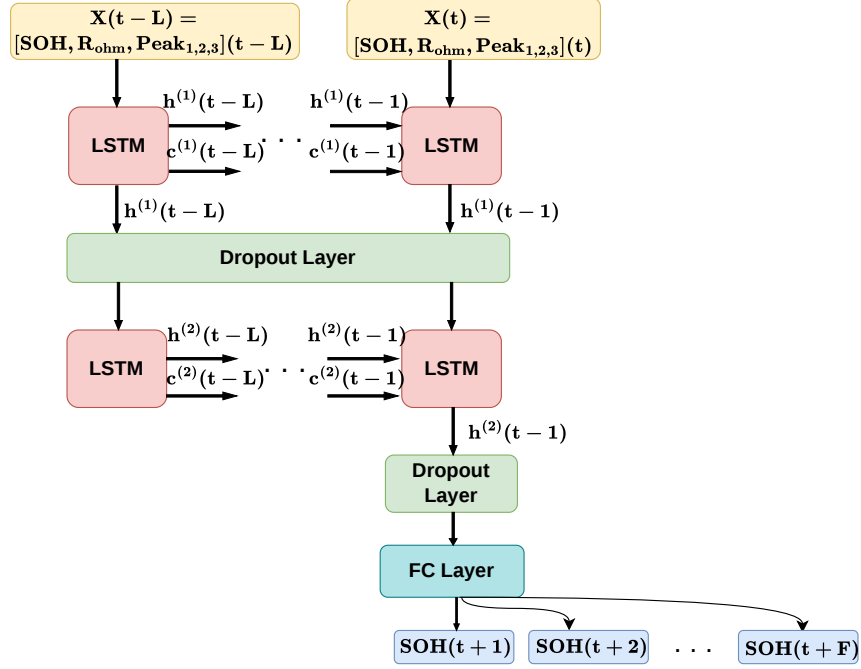


Figure 4.1: Bayesian LSTM architecture. Stacked LSTMs with dropout layers predict future SOH trajectories from historical SOH and impedance inputs.

other hand, adopts an attention-based Transformer like structure to capture long-range dependencies more effectively [150]. It begins by embedding the multivariate input (SOH, R_{ohm} , and DRT peaks) together with temporal positional encodings into a high-dimensional feature space. The encoder of the Informer applies multiple layers of *ProbSparse self-attention* interleaved with feed-forward networks, progressively distilling the input sequence into a compact and highly informative representation.

In a standard Transformer, the self-attention mechanism computes dependencies across all input tokens, with computational cost $\mathcal{O}(L^2)$ for a sequence of length L . Informer replaces this with *ProbSparse attention*, which selects only the most informative query–key pairs based on their energy contributions,

reducing complexity toward $\mathcal{O}(L \log L)$. This sparse approximation is particularly advantageous in long-horizon forecasting, where redundant dependencies dominate and efficiency is critical. Formally, given a query Q , keys K , and values V , the self-attention operation is:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V \quad (4.1)$$

where d_k is the key dimension. ProbSparse attention modifies this by retaining only the queries with the largest attention scores, thereby focusing on dominant dependencies while discarding near-uniform ones.

The decoder combines this distilled representation with historical context to generate the forecasted SOH trajectory and the associated aleatoric uncertainty estimates. Bayesian inference is facilitated by applying dropout to both the attention and feed-forward layers. By performing multiple stochastic forward passes during inference, an empirical distribution over predictions is quantified. The complete Bayesian Informer architecture is illustrated in Fig. 4.2.

Compared to the vanilla Transformer [151], the Informer is specifically designed for long-sequence time-series forecasting, achieving reduced computational complexity through ProbSparse attention and encoder distillation while preserving long-range dependency modeling. These characteristics are particularly suitable for SOFC degradation, where health evolution unfolds gradually over hundreds of hours and long-horizon dependencies are critical.

Other Transformer variants typically retain quadratic attention complexity or rely on autoregressive decoding, which may introduce error accumulation over extended horizons. In contrast, the Informer employs a one-shot generative decoder, improving stability in long-range forecasting [150]. Diffusion-based generative models, while powerful for probabilistic sequence generation, generally involve iterative sampling procedures that significantly increase inference cost and are less practical for real-time prognostics. For these reasons, the Informer provides an effective balance between long-horizon accuracy and computational tractability for SOFC RUL forecasting.

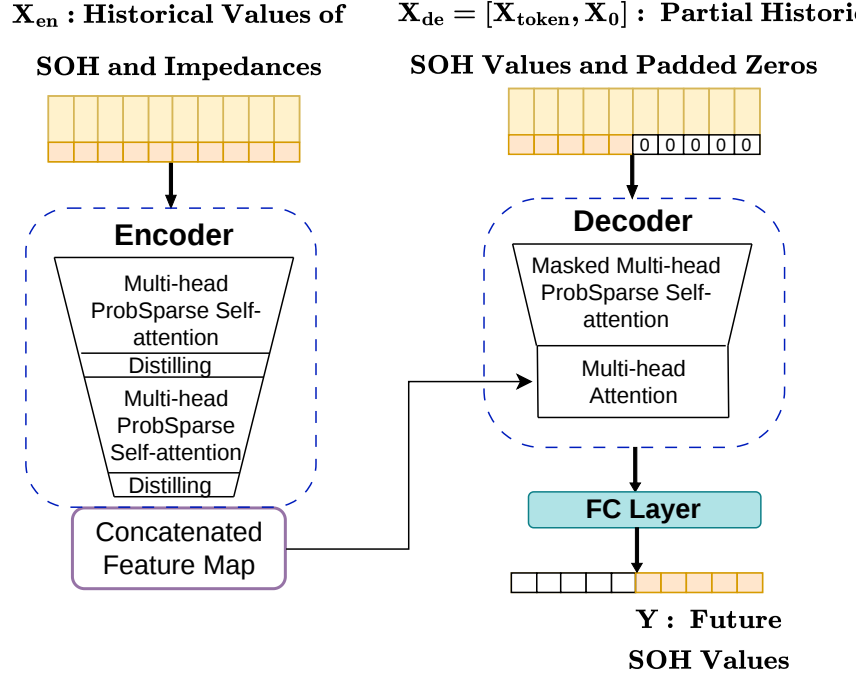


Figure 4.2: Bayesian Informer architecture. An encoder–decoder with ProbSparse attention models long-range dependencies to forecast SOH sequences.

4.2.1 Dynamic Forecast Horizon

Because the remaining horizon to EOL shrinks as a cell operates, a unified model that can forecast RUL from any point is required. To achieve this, a reference horizon equal to the longest training sequence was defined, and all shorter sequences were zero-padded to standardize batch dimensions. Initially, padded steps were masked from the loss, but this prevented the model from learning the physical constraint that $SOH \rightarrow 0$ at EOL. To address this, padded points were reintroduced with non-zero weights. However, assigning uniform weights caused premature collapse of SOH predictions toward zero, since the model learned to overfit the artificial padding. To better align with the physical degradation process, an *exponentially decaying weighting scheme* was applied to the padded region. This approach gradually reduces the influence of the artificial padding the farther it extends beyond the actual data, while still guiding the model to approach $SOH = 0$ at EOL. Formally, let t

denote the time step in a sequence and T the true sequence length. The weighting function is defined as:

$$w(t) = \begin{cases} 1, & 1 \leq t \leq T \\ \lambda^{(t-T)}, & T < t \leq T_{\max} \end{cases} \quad (4.2)$$

where $T_{\max}$ is the maximum sequence length after padding, and $0 < \lambda < 1$ is a decay parameter (set to 0.3 in this study). A sensitivity analysis comparing masking, uniform weighting, and exponential weighting (with multiple λ values) is provided in Section 4.3.4, demonstrating that $\lambda = 0.3$ offers a robust trade-off between forecast accuracy, interval calibration, and RUL estimation. The formulation in Eq.(4.2) ensures that all valid steps are equally weighted, while padded steps are down-weighted exponentially according to their distance from the true end of sequence. Conceptually, this can be viewed as a form of *curriculum learning*, where the model is first guided strongly by observed data and then softly guided toward the EOL constraint through padded regions.

4.2.2 Uncertainty Quantification

Quantifying uncertainty is critical in RUL forecasting to ensure reliable and actionable predictions for SOFCs. Accurate uncertainty estimation enables controllers and supervisory systems to adapt operating strategies according to confidence levels in predicted degradation. Two primary types of uncertainty are considered: aleatoric, which reflects inherent variability and noise in the data, and epistemic, which arises from limited model knowledge or insufficient training data. Aleatoric uncertainty is captured by modeling the predictive variance (σ_t^2) as an additional network output alongside the mean prediction (μ_t). This is achieved by modifying the loss function to include a Gaussian negative log-likelihood (NLL) term:

$$Loss = \underbrace{\sum_t w(t) (y_t - \mu_t)^2}_{\text{mean fit}} + \underbrace{\sum_t \left[\frac{(y_t - \mu_t)^2}{\sigma_t^2} + \log \sigma_t^2 \right]}_{\text{NLL on real data}} + \underbrace{\lambda \sum_{t \in \text{pad}} \sigma_t^2}_{\text{pad-penalty}} \quad (4.3)$$

Here, the first term encourages accurate mean predictions; the second term is the NLL, penalizing large errors while scaling uncertainty; and the third term penalizes variance in padded regions. To ensure positivity and stability, variance is learned in log-space. Aleatoric uncertainty is then computed as:

$$\sigma_{\text{alea}}(t) = \exp\left(\frac{1}{2} \log \sigma_t^2\right) \quad (4.4)$$

In contrast, epistemic uncertainty reflects uncertainty due to model limitations or inadequate coverage of the input data space. To quantify it, Bayesian inference is used. Because exact inference is complex in deep neural networks, MC dropout provides a practical approximation: multiple stochastic forward passes are made through the network with dropout active at inference, simulating posterior sampling of network parameters. The variance of predictions across these passes captures epistemic uncertainty [152]. The final predictive uncertainty combines both aleatoric and epistemic terms, with the total variance given by:

$$\sigma_{\text{total}}^2 = \sigma_{\text{epi}}^2 + \sigma_{\text{alea}}^2 \quad (4.5)$$

Using the predictive mean, $\bar{\mu}_t$, and the total predictive standard deviation, $\sigma_{\text{total},t}$, the two-sided 90% confidence interval at each prediction time step is calculated as:

$$\text{CI}_{90\%}(t) = \bar{\mu}_t \pm 1.645 \sigma_{\text{total},t}, \quad (4.6)$$

This formulation assumes that the resulting predictive distribution can be reasonably approximated by a Gaussian distribution. In this work, 10 MC-dropout forward passes are performed for each input sample. Repeated stochastic forward passes provide Monte Carlo estimates of the predictive mean and variance, which converge toward their expected values as the number of samples increases according to the law of large numbers. Moreover, the central limit theorem supports an approximately Gaussian sampling distribution for the estimated predictive mean when the number of stochastic samples is sufficiently large. Therefore, 10 forward passes were selected as a compromise between estimation stability and the computational cost associated with repeated inference.

To transform forecasted SOH trajectories into RUL distribution, the *first hitting times* of the EOL threshold are computed. For each MC dropout iteration $r = 1, \dots, N$, the model generates a forecasted trajectory $\{\hat{y}_k^{(r)}\}$ with sampling interval Δt . The FHT is defined as the earliest step $k^*(r)$ such that $\hat{y}_{k^*(r)}^{(r)} \leq 0$, with the corresponding RUL sample

$$\text{RUL}^{(r)} = k^*(r) \Delta t \quad (4.7)$$

Collecting all samples $\{\text{RUL}^{(r)}\}_{r=1}^N$ yields an empirical distribution of RUL. Two estimators are considered:

Histogram-based PDF — With B bins, bin width Δb , and count C_j in bin j , the probability density is [153]:

$$\text{PDF}_j = \frac{C_j}{N \cdot \Delta b}, \quad \sum_{j=1}^B \text{PDF}_j \cdot \Delta b = 1 \quad (4.8)$$

Kernel Density Estimation (KDE) — To obtain a smooth estimate, Gaussian kernel smoothing is applied [154]:

$$\hat{f}(t) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{t - r_i}{h}\right), \quad K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad (4.9)$$

where h is the bandwidth, and it is chosen by using Silverman's rule,

$$h = \frac{\sigma}{N^{1/5}} \quad (4.10)$$

with σ representing the sample standard deviation of $\{r_i\}$.

The complete data processing and forecasting workflow is summarized in Fig. 4.3, from EIS/J-V measurements through feature extraction, uncertainty estimation, and RUL distribution generation.

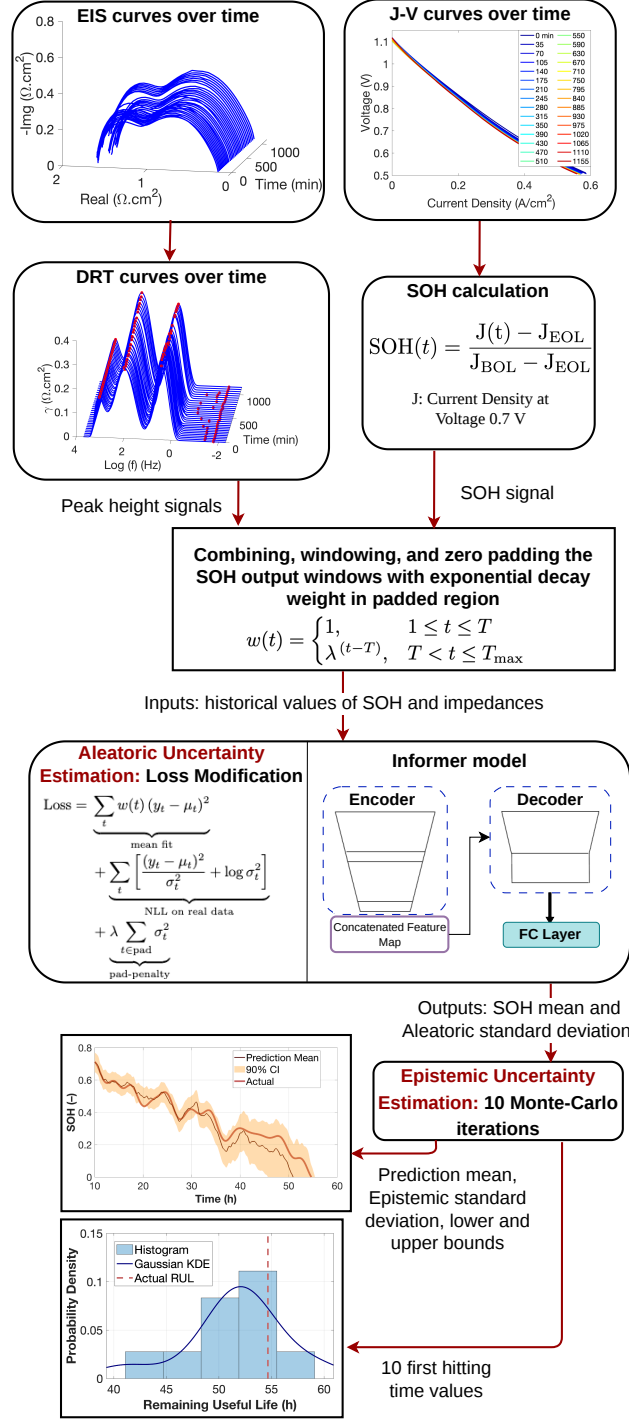


Figure 4.3: Proposed prognostics workflow. EIS/J-V data are processed into SOH and impedance features, windowed with zero-padding, and fed to the Bayesian Informer model. Aleatoric and epistemic uncertainties are quantified and RUL distributions obtained via FHT analysis. The results can inform control and maintenance decisions. 90

4.3 Results and Discussions

4.3.1 Hyper-parameter Optimization

The RUL forecasting models are trained using datasets N1 to N6 and R2, and evaluated on the unseen dataset R1. This dataset-level separation ensures that complete degradation trajectories are excluded from training, preventing information leakage across windows derived from the same cell and providing a more realistic assessment of generalization. A critical hyperparameter is the look-back window, which specifies how many past observations condition each forecast. To determine this hyperparameter in a principled way, *Takens' embedding theorem* is employed. The theorem states that the full dynamics of a deterministic system can be reconstructed from time-delayed measurements of a single observable state. The reconstructed state space is formed as

$$\mathbf{x}(t) = [x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)] \quad (4.11)$$

where m is the embedding dimension (the number of delayed coordinates) and τ is the delay time. For sufficiently large m , this embedding preserves the topology of the original dynamical system. In practice, the delay τ is selected as the first minimum of the mutual information (MI) function, ensuring that delayed samples provide complementary rather than redundant information. The embedding dimension m is chosen using the false nearest neighbors (FNN) criterion, which increases m until the fraction of spurious neighbors in the reconstructed space falls below a threshold. The corresponding look-back length is then defined as

$$L = (m - 1)\tau + 1 \quad (4.12)$$

This MI+FNN procedure was applied to all signals across the eight SOFC datasets (SOH, ohmic resistance, and DRT peak features). To standardize training, a single global look-back was set equal to the maximum L observed across signals and datasets. This yielded an optimal look-back of 3.25 hours, corresponding to 390 past samples.

For model hyperparameters, two different tuning strategies were employed.

The LSTM was optimized using a *random search*, which samples hyperparameters from predefined ranges and evaluates multiple candidates to identify a suitable configuration (Table 4.1). This approach is computationally efficient for relatively simple architectures with fewer interacting hyperparameters. By contrast, the Informer model has a larger design space due to its attention mechanisms, embedding dimensions, and learning rate schedules. For this reason, hyperparameters were tuned using the *Optuna* framework, which applies Bayesian optimization combined with pruning to adaptively explore promising regions of the search space (Table 4.2). Random search for the LSTM was conducted over 30 sampled configurations, while the Informer model was tuned using Optuna with a search budget of 50 trials. In both cases, the objective function was the validation loss defined in Eq. 4.3, evaluated on a held-out validation subset from the training data. The configuration achieving the lowest validation loss was selected. To ensure reproducibility, all experiments were conducted using fixed random seeds, and early stopping was applied based on validation performance. Training of the models was stabilized using early stopping, with a maximum of 30 epochs and a patience of 4 epochs based on the validation loss. For the Bayesian Informer, convergence was achieved at epoch 8, well before the maximum limit, indicating that the model reached its optimal point without overfitting. For the Bayesian LSTM, loss convergence occurred at epoch 13. Training and validation loss trajectories for both models are shown in Figures 4.4 and 4.5, showing stable convergence behavior and no evidence of overfitting.

4.3.2 Model Comparison

The comparative results between Bayesian LSTM and Bayesian Informer highlight the clear advantages of attention-based architectures for SOFC prognostics. As shown in Table 4.3, the Informer achieves substantially lower prediction errors, with MAE reduced by 87.5% and RMSE by 90.2% compared to LSTM. In addition, the prediction intervals produced by the Informer exhibit sharper uncertainty bounds (54.9% improvement in sharpness) while still

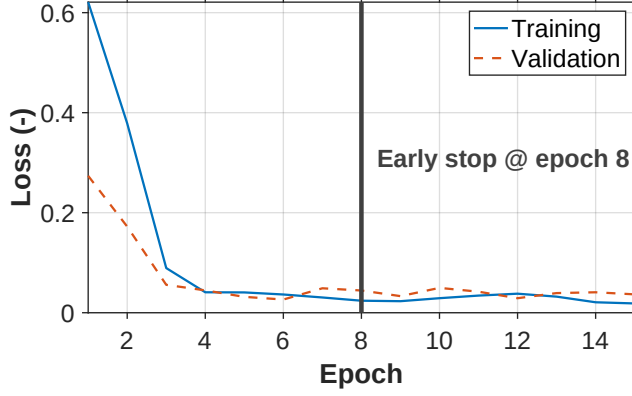


Figure 4.4: Training and validation loss curves for the Bayesian Informer model.

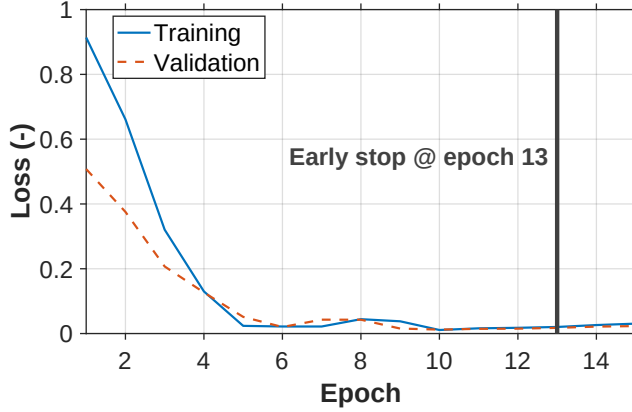


Figure 4.5: Training and validation loss curves for the Bayesian LSTM model.

achieving higher coverage (93.2% vs. 87.9%), ensuring that the true RUL remains within the confidence bounds more consistently. These results are visually confirmed in Figures 4.6 and 4.7, where the Informer predictions demonstrate tighter and more stable uncertainty envelopes around the ground truth trajectory. Although both models exhibit a slight conservative bias, with the mean prediction crossing the EOL threshold earlier than the true trajectory, the associated uncertainty intervals consistently capture the true trajectory with the expected coverage, indicating well-calibrated uncertainty estimates. The reported prediction intervals are constructed using a $\pm 1.645\sigma$ criterion corresponding to 90% confidence, which is commonly used in prognostics to balance interval sharpness and coverage. Under this definition, the Bayesian Informer achieves a coverage of 93.2%, indicating reliable uncertainty quan-

Table 4.1: Search space and optimized hyperparameters for the Bayesian LSTM model

Hyperparameter	Search Space	Optimized Value
Hidden units (1st LSTM layer)	{64, 128, 256}	128
Hidden units (2nd LSTM layer)	{32, 64, 128}	64
Dropout rate	{0.1, 0.3, 0.5}	0.3
Learning rate	$\{10^{-4}, 10^{-3}, 10^{-2}\}$	10^{-3}
Batch size	{16, 32, 64}	64
MC iterations	{5, 10, 20}	10

tification. From a prognostics perspective, such conservative behavior is often preferred, as it reduces the risk of late failure prediction.

Furthermore, the RUL distributions derived from FHTs show that the Informer yields means much closer to the true RUL of 54.6 hours, with 52.0 hours for the long-horizon case and 54.8 hours for the short-horizon case, compared to the less accurate estimates of 48.5 and 56.8 hours produced by the LSTM model. This consistency highlights the Informer’s ability to provide well-calibrated and accurate probabilistic RUL forecasts. The superior performance of the Informer model can be attributed to its attention mechanism, which effectively captures long-range temporal dependencies in degradation signals. Unlike recurrent architectures such as LSTM, which suffer from vanishing gradients and limited memory over extended sequences, the Informer leverages self-attention to selectively focus on informative time steps, leading to more accurate long-horizon forecasts. This architectural advantage makes the Informer particularly well-suited for SOFC prognostics, where degradation evolves gradually over hundreds of hours and long-term dependencies must be preserved. However, as summarized in Table 4.4, this improvement comes with a computational trade-off: the Informer requires roughly $20\times$ more parameters and substantially higher training and inference times compared to the LSTM baseline. It should be noted that the reported parameter counts correspond to the optimized architectures obtained through hyperparameter

Table 4.2: Search space and optimized hyperparameters for the Bayesian Informer model

Hyperparameter	Search Space	Optimized Value
Model dimension	{128, 256, 512}	256
Number of attention heads	{2, 4, 8}	4
Number of encoder layers	{1, 2, 3}	2
Number of decoder layers	{1, 2}	1
Feed-forward dimension	{512, 1024, 2048}	1024
Dropout rate	{0.1, 0.3, 0.5}	0.1
Learning rate	$\{10^{-4}, 10^{-3}, 10^{-2}\}$	10^{-4}
Batch size	{16, 32, 64}	32
MC iterations	{5, 10, 20}	10

tuning for each model. Therefore, these values represent configurations selected to balance predictive performance and model complexity, rather than arbitrarily oversized networks or undersized architectures that do not fully exploit the modeling capacity of each method. Despite this, the gains in accuracy, sharpness, and coverage make the Informer a compelling state-of-the-art solution for uncertainty-aware SOFC prognostics in applications where reliability is prioritized over computational efficiency.

In addition to the Bayesian models, Table 4.3 reports the performance of several commonly used deterministic baselines, including GRU, temporal convolutional network (TCN), and a deterministic Informer. Among these, the Informer architecture achieves the lowest MAE and RMSE, highlighting the benefit of attention-based modeling for long-horizon SOH forecasting. The GRU and TCN models provide reasonable accuracy but exhibit higher errors, particularly under long-range dependencies. While the deterministic Informer slightly outperforms its Bayesian counterpart in terms of point accuracy, it does not provide uncertainty estimates. By contrast, the Bayesian Informer achieves comparable accuracy while additionally offering calibrated predictive uncertainty, which is essential for risk-aware prognostics and decision-making.

Statistical significance was evaluated using a one-sided paired Wilcoxon signed-rank test based on the RMSE values calculated for 50 paired SOH prediction trajectories. Each benchmark was compared with the Bayesian Informer, and the p-values are reported in Table 4.3, using a significance level of $\alpha = 0.05$. The Bayesian Informer achieved significantly lower trajectory-level RMSE than the GRU (p=0.009), TCN (p=0.015), and Bayesian LSTM (p<0.001). In contrast, the comparison with the deterministic Informer yielded p=0.959, providing no evidence that the Bayesian Informer has lower RMSE, which is consistent with the slightly better point-prediction accuracy of the deterministic Informer. Nevertheless, the Bayesian Informer maintains competitive accuracy while additionally providing uncertainty estimates through coverage and sharpness metrics.

Table 4.3: Performance comparison of deterministic and Bayesian forecasting models for SOH trajectory prediction.

Model	MAE (-)	RMSE (-)	Coverage* (%)	Sharpness* (-)	p-value* (vs. Bayesian Informer)
GRU	0.0148	0.0196	-	-	0.009
TCN	0.0102	0.0135	-	-	0.015
Informer	0.0079	0.0098	-	-	0.959
Bayesian LSTM	0.0697	0.1101	87.93	0.2105	< 0.001
Bayesian Informer	0.0087	0.0108	93.20	0.0950	-

* Coverage and sharpness are reported only for Bayesian models.

4.3.3 Cross-Dataset Generalization

Leave-One-Dataset-Out (LODO) Validation. To strengthen the evaluation of the Bayesian Informer model, a leave-one-dataset-out (LODO) validation strategy was adopted. In this protocol, one entire dataset (cell) is excluded from training and used exclusively for testing, while the remaining datasets are used for model training. This process is repeated for each dataset, ensuring that the model is evaluated on complete degradation trajectories that

Table 4.4: Computational cost comparison between Bayesian LSTM and Bayesian Informer on an NVIDIA Tesla T4 GPU. Inference time is reported for 10 MC dropout iterations.

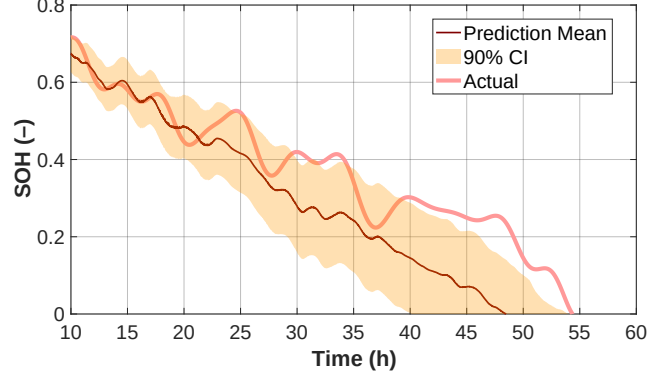
Metric	Bayesian LSTM	Bayesian Informer
Trainable parameters (M)	0.54	11.31
Training time (s/epoch)	5.52	305.32
Inference time (ms/sample)	0.32	39.53
Peak GPU memory (GB)	4.27	10.78

are not seen during training. Compared to random window-based splitting, this strategy prevents information leakage across samples from the same cell and provides a more stringent assessment of generalization across devices and cycling conditions. All performance metrics are computed on the held-out dataset in each fold and then summarized across folds in Table 4.5.

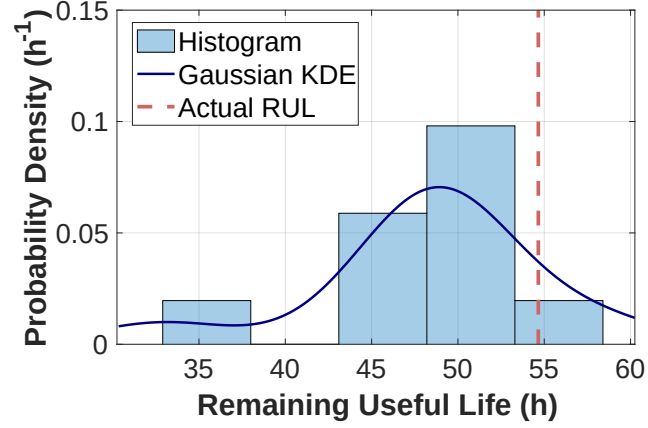
Table 4.5: LODO validation results for the Bayesian Informer model. Each fold corresponds to one dataset excluded from training and used for testing.

Test dataset	MAE (-)	RMSE (-)	Coverage (%)	Sharpness (-)
N1	0.0072	0.0091	94.1	0.082
N2	0.0073	0.0093	93.8	0.084
N3	0.0091	0.0118	92.4	0.108
N4	0.0078	0.0100	93.2	0.090
N5	0.0084	0.0107	92.9	0.098
N6	0.0066	0.0085	94.6	0.076
R1	0.0087	0.0108	93.2	0.095
R2	0.0094	0.0122	91.8	0.110
Mean	0.0081	0.0103	93.3	0.093
Std.	0.0009	0.0012	0.9	0.012

These results demonstrate generalization across different cells and cycling patterns. However, the operating conditions considered are limited to controlled experimental settings, and the model is not evaluated under operating condi-



(a) SOH trajectory with 90% CI. CI stands for confidence interval.

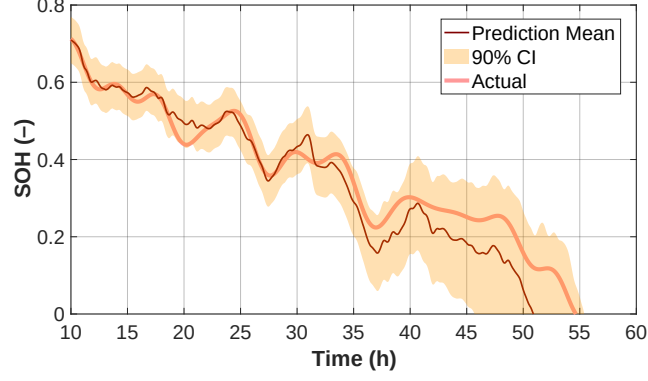


(b) Corresponding RUL distribution derived via FHT analysis, providing uncertainty-aware input for control and maintenance scheduling.

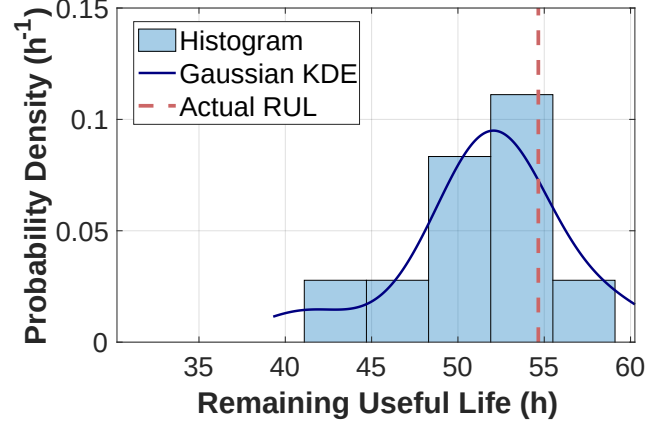
Figure 4.6: Bayesian LSTM model results

tions outside the training range, such as substantially different temperatures or poorly controlled operating conditions. In practical applications, the proposed framework is intended to be trained or updated using data representative of the target operating conditions.

Evaluation with Alternative Train–Test Splits. To further assess the generalization capability of the selected Bayesian Informer model, an additional validation experiment was conducted in which the training set consisted only of datasets N1–N6, while both randomized datasets (R1 and R2) were used exclusively for testing. This ensures that the model is evaluated on entirely unseen degradation trajectories exhibiting non-periodic redox cycling patterns.



(a) SOH trajectory with 90% CI.



(b) Corresponding RUL distribution estimated using FHT, which provides uncertainty-aware input for control and maintenance scheduling.

Figure 4.7: Bayesian Informer model results

Under this protocol, as reported in Table 4.6, the Bayesian Informer model maintains strong predictive performance across both R1 and R2, with comparable MAE, RMSE, coverage, and sharpness metrics to those reported previously. These results confirm that the proposed framework generalizes well to randomized, non-stationary degradation conditions not observed during training.

To further evaluate the model under a more stringent data-scarcity scenario, this experiment was extended by additionally removing dataset N6 from the training set and including it in the test set alongside R1 and R2. Under this setting, the Bayesian Informer is trained only on datasets N1–N5 and evaluated on three entirely unseen degradation trajectories, including both structured

Table 4.6: Bayesian Informer performance when trained on N1–N6 and tested on unseen datasets R1 and R2.

Test dataset	MAE (–)	RMSE (–)	Coverage (%)	Sharpness (–)
R1	0.0098	0.0121	92.2	0.128
R2	0.0112	0.0146	91.0	0.162

(N6) and randomized (R1–R2) cycling conditions. As expected, the overall prediction accuracy decreases modestly due to the reduced diversity in the training data (Table 4.7); however, the model continues to provide stable forecasts with reasonable uncertainty calibration, demonstrating robustness under reduced training coverage.

Table 4.7: Bayesian Informer performance when trained on N1–N5 and tested on unseen datasets R1, R2, and N6.

Test dataset	MAE (–)	RMSE (–)	Coverage (%)	Sharpness (–)
R1	0.0106	0.0134	91.5	0.142
R2	0.0125	0.0162	90.3	0.176
N6	0.0094	0.0119	92.1	0.118

Although prediction accuracy decreases modestly as the number of training datasets is reduced, the model maintains stable trajectory forecasting and well-calibrated uncertainty bounds, indicating that the framework does not inherently require the full data volume reported here.

In practical deployments, SOFC stacks often operate for 30,000–40,000 hours. Continuous monitoring over such lifetimes naturally generates substantial time-series data, even at moderate sampling frequencies, and data can be aggregated across multiple stacks operating under similar conditions. Moreover, when larger datasets are available, models trained on such data can serve as pretrained representations that are subsequently adapted to new SOFC systems or degradation modes using transfer learning techniques, requiring only limited additional data. This provides a practical pathway for extending the

proposed framework with reduced data collection effort.

4.3.4 Sensitivity Analysis of Padding Strategy and Decay Factor

To assess the impact of the exponentially weighted zero-padding scheme introduced in Eq. 4.2, a sensitivity analysis was conducted by comparing several padding strategies within the Bayesian Informer framework. The following variants were evaluated:

- **No weighting (masking):** padded time steps were excluded from the loss function.
- **Uniform weighting:** padded time steps were included in the loss with a constant weight.
- **Exponential weighting:** padded time steps were weighted using Eq. 4.2 with different decay factors $\lambda \in \{0.1, 0.3, 0.5\}$.

All variants were trained and evaluated under the same protocol, using identical training data, model architecture, and hyperparameters. Performance was assessed on the same unseen test dataset using trajectory-level accuracy and uncertainty metrics, including MAE, RMSE, coverage, and sharpness. Table 4.8 summarizes the results. The no-weighting (masking) strategy yields accurate short-term forecasts but does not enforce convergence toward EOL. Uniform weighting enforces convergence but introduces premature collapse of the SOH trajectory, degrading forecast accuracy and uncertainty calibration. In contrast, exponential weighting provides balanced behavior by softly enforcing the EOL constraint while preserving forecast fidelity. Among the tested values, $\lambda = 0.3$ consistently offers a robust trade-off between accuracy and interval calibration and was therefore selected in this study.

4.3.5 Deployment Considerations and Applications

Figure 4.8 summarizes the practical workflow of the proposed prognostics framework and highlights the elements required for real-world deployment.

Table 4.8: Sensitivity analysis of padding strategies and decay factor λ for the Bayesian Informer model.

Padding strategy	λ	MAE (-)	RMSE (-)	Coverage (%)	Sharpness (-)
No weighting (masking)	–	0.0335	0.0520	90.5	0.157
Uniform weighting	–	0.0840	0.1320	85.6	0.240
Exponential weighting	0.1	0.0102	0.0134	92.1	0.098
Exponential weighting	0.3	0.0087	0.0108	93.2	0.095
Exponential weighting	0.5	0.0094	0.0121	94.0	0.112

The upper section of the diagram represents the components implemented and validated in this work, including data collection, preprocessing, model training, and probabilistic SOH trajectory prediction with RUL estimation. From a control engineering practice perspective, several considerations must be addressed when applying this framework to real-world SOFC systems. These include the availability and frequency of monitored signals (such as voltage, current, or impedance-derived features), the definition of an application-specific EOL criterion, and computational resources for model training and inference. In practical deployments, measurements may be irregular, noisy, or limited in number; however, the probabilistic formulation of the proposed model allows uncertainty to be quantified and propagated through the prediction process. As illustrated in the diagram, these inputs enable estimation of SOH trajectories with confidence intervals and RUL distributions, which can then be used for operational decision-making.

The predicted SOH trajectories and probabilistic RUL estimates enable several practical applications in SOFC-based energy systems. First, the predicted SOH evolution provides quantitative information about degradation progression. By knowing the current SOH and its projected trajectory, supervisory controllers can adjust operational variables such as fuel and air mass flow rates and current density to reduce degradation rates and prolong the lifetime of SOFC cells or stacks. In addition, SOFC systems are often integrated with battery energy storage systems (BESS) in hybrid architectures. In such

configurations, knowledge of stack SOH can guide energy management strategies, allowing degraded stacks to be derated while the battery compensates for transient power demands. Second, the predictive uncertainty provided by the probabilistic model can support robust control strategies. Specifically, the estimated uncertainty bounds and noise variance allow controllers to adopt uncertainty-aware safety margins, tighten operational constraints under degradation, and perform risk-sensitive adjustment of operating setpoints to ensure safe and reliable operation despite model and measurement uncertainties. Third, the predicted RUL distribution supports risk-aware maintenance scheduling. These probabilistic estimates enable the determination of inspection intervals, next maintenance times, and expected replacement times, allowing operators to plan maintenance actions before critical degradation thresholds are reached.

Despite these promising capabilities, several limitations remain and provide directions for future research. Although multiple validation strategies were used to assess generalization within this experimental scope, the framework does not claim universal applicability across different SOFC designs, manufacturers, or uncontrolled field environments. Furthermore, the present study focuses primarily on prognostics and does not explicitly perform root-cause diagnostics. While impedance-derived features provide mechanistic insights that can be associated with specific electrochemical processes, integrating dedicated diagnostic algorithms remains an important future extension. Such diagnostic capabilities could further enable fault-tolerant control strategies, where identified degradation modes or affected components trigger system reconfiguration or load derating actions. Finally, the current experiments focus on redox-induced degradation, which represents an accelerated aging scenario. Future work should investigate additional degradation mechanisms and real operational datasets to further validate and extend the proposed prognostics framework.

4.4 Summary of Chapter

A novel prognostics framework is introduced for SOFCs that forecasts the full future SOH trajectory using Bayesian LSTM and Bayesian Informer models. Uncertainty is decomposed into aleatoric (learned via a Gaussian negative log-likelihood head) and epistemic (via Monte Carlo dropout), resulting in calibrated intervals for both SOH trajectories and RUL. To handle the shrinking, unknown forecast horizon with a single model, an exponentially weighted zero-padding technique is proposed, which preserves sequence length and encourages convergence to $\text{SOH} = 0$ at EOL. With multivariate inputs that integrate impedance features alongside SOH, the approach is validated on large SOFC redox aging datasets containing over 2.1 million samples, collected under diverse temperatures and cycling test profiles in this study. The validation results show that the Bayesian Informer achieves high forecasting accuracy, with MAE of 0.0087, RMSE of 0.0108, coverage of 93.20%, and sharpness of 0.0950. The proposed method outperforms the Bayesian LSTM, achieving about 90% lower forecast error, 6% higher coverage, and about 55% sharper bands under the same protocol. Moreover, RUL distributions derived from FHT analysis show that the Bayesian Informer model provides prediction means much closer to the actual RUL than Bayesian LSTM, demonstrating its ability to deliver accurate and well-calibrated probabilistic prognostics. Beyond forecasting accuracy, integrating impedance-derived features strengthens the practical relevance of the framework for health-aware control and decision support. Because impedance reflects specific electrochemical and diffusional phenomena, changes in these features can be linked to the anode, cathode, electrolyte, or interfaces, thereby supporting control reconfiguration and targeted maintenance. In particular, peaks 1 and 2, associated with oxygen ion transfer and anode charge transfer, showed the expected increase during redox cycling, consistent with the anode being most affected under fuel depletion.

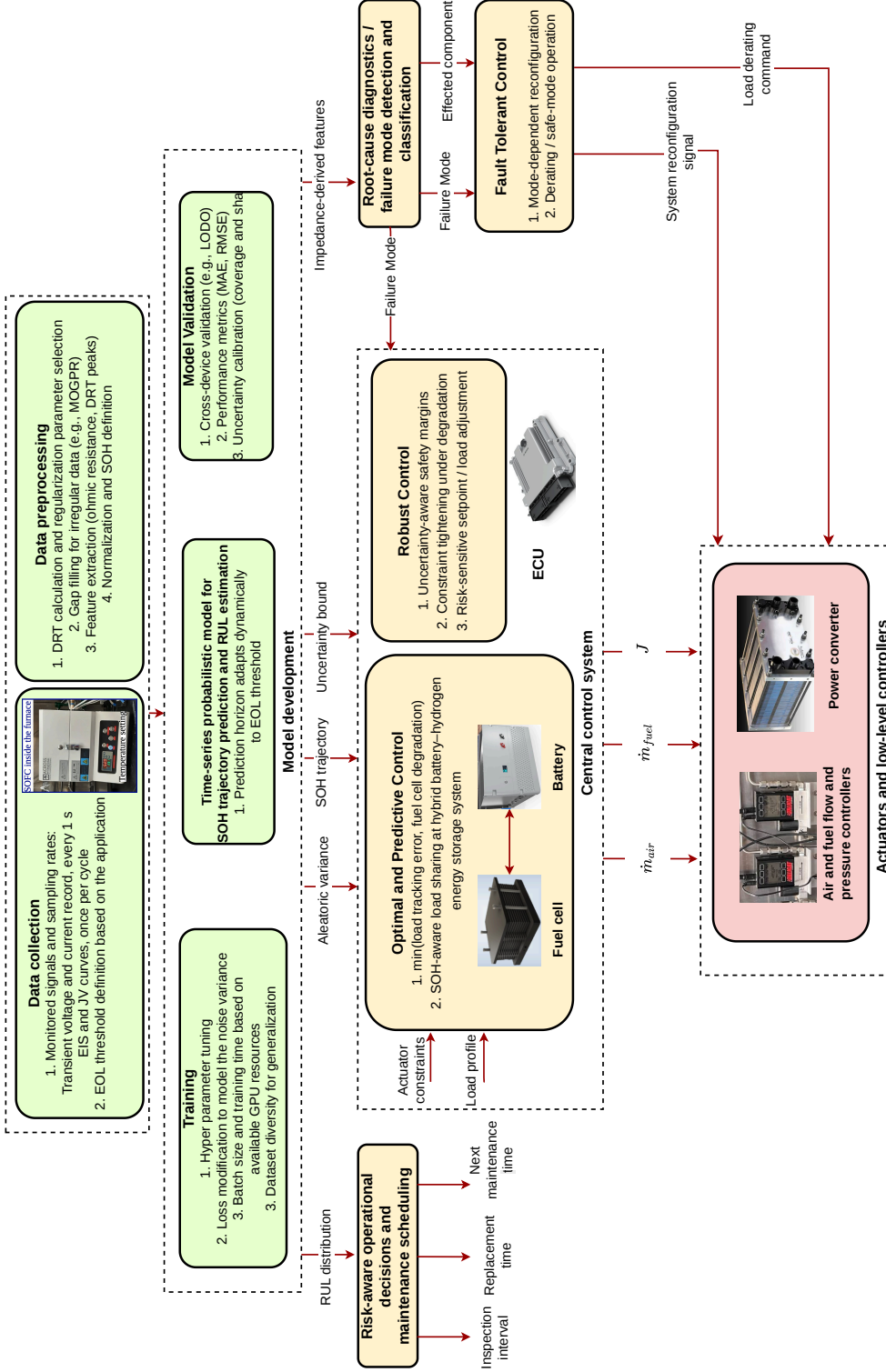


Figure 4.8: Practical deployment workflow of the proposed SOFC prognostics and control framework. Green blocks represent the stages implemented and validated in this work, including data collection, preprocessing, model training, validation, and probabilistic SOH-RUL prediction. Yellow blocks illustrate downstream control-oriented applications enabled by the prognostics outputs, such as optimal and robust control, diagnostics, and maintenance planning. Red blocks denote the low-level controllers responsible for executing control actions in the physical system.

Chapter 5

Degradation Mode Classification¹

5.1 Introduction

While the long-horizon forecasting framework developed in Chapter 4 enables accurate prediction of SOH trajectories and RUL, such prognostic models are inherently influenced by the underlying degradation mechanisms. In practical SOFC operation, degradation behavior varies significantly depending on the dominant degradation mode, which directly affects both the evolution of performance and the reliability of RUL predictions. Therefore, accurate identification of the root cause of degradation is a critical complementary task to prognostics, enabling more informed decision-making for predictive maintenance and control.

To address this need, this chapter focuses on the development of a diagnostic framework for degradation mode classification in SOFC systems. In particular, EIS measurements are utilized to capture rich information about internal electrochemical processes. These measurements are transformed into frequency-resolved diagnostic signatures using DRT analysis, providing physically interpretable features that reflect the mechanisms driving degradation. Using these impedance-derived features, a supervised learning framework is constructed to classify multiple degradation modes, including redox cycling, fuel starvation, internal short circuit, and thermal shock. To comprehensively

¹This chapter is partially based on the following publication: [4].

evaluate diagnostic performance, several machine learning models are investigated, including both feature-based ensemble methods and temporal deep learning architectures capable of capturing the evolution of degradation over time. This enables a systematic comparison between classical classifiers and sequence-based models in identifying degradation mechanisms.

The proposed approach is validated using experimentally generated degradation datasets collected under controlled degradation conditions. The results demonstrate that temporal deep learning models, particularly the attention-based Informer, achieve superior classification performance and effectively distinguish between mechanistically similar degradation modes. In addition, the trade-off between prediction accuracy and computational efficiency is analyzed across different model families.

Overall, this chapter establishes an impedance-driven diagnostic framework that enables robust identification of SOFC degradation modes. By linking observable degradation patterns to their underlying mechanisms, this work provides a critical foundation for mechanism-aware prognostics and motivates the integrated diagnostic-prognostic framework developed in the following chapter.

5.2 Data Processing and Class Labeling

Figure 5.1 presents representative DRT evolution profiles for the four experimentally induced degradation mechanisms. As shown in Figure 5.1, the characteristic frequencies and overall shapes of the DRT peaks remain largely unchanged during operation, while variations in peak amplitudes capture the temporal evolution of the associated impedance contributions. The ohmic resistance, R_{ohm} , is extracted from the high-frequency intercept of the Nyquist plot. To address nonuniform temporal sampling in the R_{ohm} and DRT peak time-series, a multi-output Gaussian process regression (MOGPR) approach is employed for interpolation [142].

For redox cycling, pronounced growth is observed in the high- and mid-frequency

DRT peaks (Peak 1 and Peak 2). Peak 1 is associated with oxygen ion transport and electrolyte–electrode interfacial polarization, while Peak 2 corresponds to anode activation polarization. These trends are consistent with microstructural degradation in the anode functional layer, including nickel particle coarsening and pore structure evolution, which reduce the effective Ni-YSZ triple-phase boundary (TPB) area and increase polarization resistance. In fuel starvation experiments, degradation manifests primarily as an increase in Peak 2 along with a concurrent rise in Ohmic resistance. The amplification of Peak 2 indicates worsening anode activation polarization, whereas the increase in Ohmic resistance suggests electrolyte degradation or loss of electrical continuity. Physically, these changes are attributed to anode delamination, microstructural depletion, and reduced electronic and ionic conductivity caused by sustained hydrogen deficiency and local re-oxidation effects. For short circuit degradation, the dominant signature is a monotonic increase in ohmic resistance. This behavior points to damage in the electrolyte or current collection pathways rather than kinetic limitations. Localized overheating during repeated shorting events can induce micro-cracks, contact degradation, or interlayer delamination, leading to increased resistive losses. Finally, thermal shock degradation is characterized by a simultaneous increase in Peak 3 and the Ohmic resistance. Peak 3 is associated with cathode activation polarization, indicating degradation of cathodic electrochemical processes. These trends are consistent with severe and rapid temperature gradients inducing thermal stresses, which promote crack formation in the YSZ electrolyte and cathode layers. Notably, the cathode constitutes the outermost layer of the tubular cell and is directly exposed to ambient air during the furnace removal step. This asymmetric cooling environment leads to accelerated heat extraction on the cathode side, making it susceptible to delamination and contact loss.

Each experiment yields a multivariate time-series composed of five input channels: four DRT peak amplitudes and the ohmic resistance. Prior to model training, these signals are resampled to ensure uniform temporal spacing, stan-

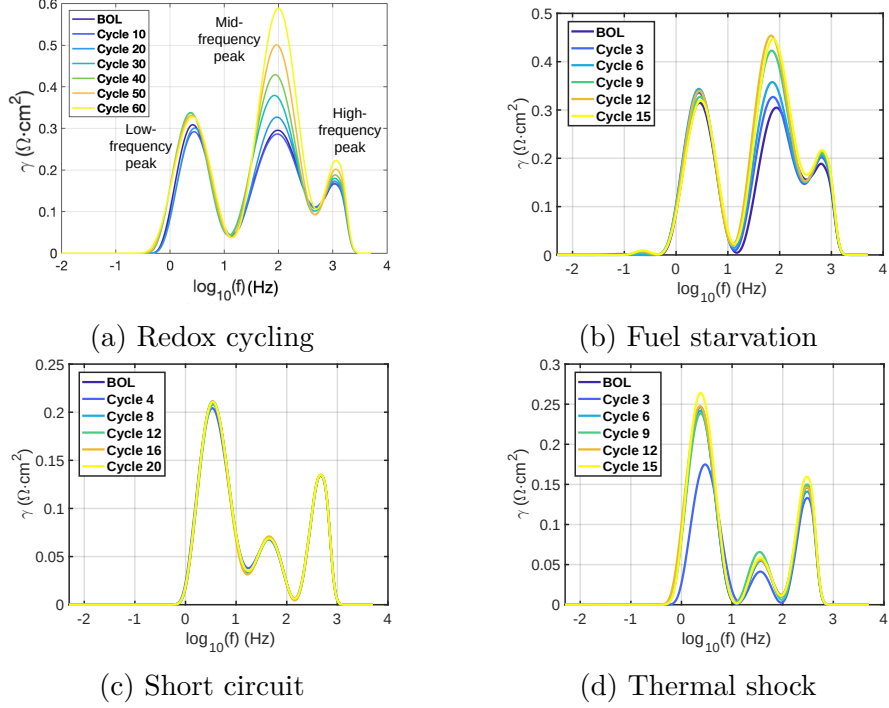


Figure 5.1: Representative DRT evolution under the four degradation mechanisms

standardized to remove scale differences across features, and segmented into fixed-length overlapping windows using a look-back horizon of 500 samples. Each window is assigned a class label corresponding to the degradation mechanism active during the associated experiment, ensuring consistency between the temporal data and the underlying degradation mode.

The four degradation mechanisms are encoded using one-hot vectors to facilitate multi-class classification, with redox cycling represented as $[1, 0, 0, 0]$, fuel starvation as $[0, 1, 0, 0]$, short circuit as $[0, 0, 1, 0]$, and thermal shock as $[0, 0, 0, 1]$. The resulting windowed and labeled dataset serves as the input to all classification models evaluated in this study.

5.3 Shallow Models for Degradation Mode Classification

In contrast to deep neural architectures, shallow learning models are usually based on decision trees, linear functions, or kernel methods and have few or

no hidden layers. When interpretability and training efficiency are desired or when the size of the available dataset is restricted, these models have been widely used for classification and regression tasks. In this section, engineered statistical features taken from sensor signals were used to classify degradation modes in SOFC systems using shallow models. They were appropriate for our feature-based classification framework due to their robustness against overfitting and comparatively low computational cost. In the following, the details of three representative shallow classifiers, RF, Extreme Gradient Boosting (XGBoost, XGB), and an ensemble model combining RF and XGB, are presented, along with the feature engineering procedure and the data oversampling strategy employed to address class imbalance.

5.3.1 Feature Engineering

The proposed workflow follows a feature-based time-series classification framework. Raw sensor signals are first segmented using a sliding-window approach, where each window contains a fixed number of consecutive samples. The window is then shifted by a constant step size, producing overlapping segments along the time axis. Each resulting segment is assigned a degradation mode label corresponding to the operating condition present during that time interval.

Rather than directly using the raw time-series data, each windowed segment is transformed into a compact feature vector that summarizes both time- and frequency-domain characteristics. These extracted features provide statistical and dynamical information that substantially reduce data dimensionality while improving model interpretability. Classical machine-learning algorithms, including RF and XGB, are subsequently trained using these feature representations. The overall pipeline can be summarized as

$$x_{1:T} \xrightarrow{\text{Sliding Window}} w_i \xrightarrow{\text{Feature Extraction}} \phi(w_i) \xrightarrow{\text{Classifier}} \hat{y}_i.$$

This framework combines computational efficiency with strong generalization capability when applied to limited experimental datasets. Statistical disper-

sion, energy, temporal trends, and spectral characteristics were all captured in a feature vector derived from each sliding-window segment. The extracted features consist of:

Interquartile Range (iqr) The interquartile range, or iqr, measures how widely the signal values are spread around the middle of their distribution. It is calculated as the difference between the 75th percentile (Q3) and the 25th percentile (Q1):

$$\text{iqr}(x) = Q_3(x) - Q_1(x) \quad (5.1)$$

This feature tells us how much the signal fluctuates, but it ignores extreme outliers. If the signal shows strong oscillations or instability, as often happens in Fuel Starvation or Redox Cycling faults, the IQR becomes larger. In contrast, stable operating conditions produce smaller IQR values.

Energy The energy of a signal describes its overall strength or intensity. It is obtained by summing the squares of all signal values within a window:

$$\text{energy}(x) = \sum_{t=1}^T x_t^2 \quad (5.2)$$

A high-energy value means the signal has large amplitudes or strong variations. For example, during Redox Cycling fault, the current density can rise sharply, leading to much higher energy. This feature therefore reflects how powerful or turbulent the signal becomes.

Slope The slope shows whether the signal is increasing, decreasing, or staying roughly constant over the duration of the window. It is computed by fitting a simple straight line to the signal values and taking the line's slope:

$$x_t \approx a + bt, \quad b = \frac{\sum_t (t - \bar{t})(x_t - \bar{x})}{\sum_t (t - \bar{t})^2} \quad (5.3)$$

If the slope is positive, the signal is trending upward. A negative slope means a downward trend, while a near-zero slope indicates steady conditions. This feature captures gradual changes that may indicate a developing degradation mode.

Lag-1 Autocorrelation The lag-1 autocorrelation measures how similar the signal is to a shifted version of itself, one step later in time. It quantifies how smooth or random the signal behavior is:

$$\text{lag1_autocorr}(x) = \frac{\sum_{t=2}^T (x_t - \bar{x})(x_{t-1} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2} \quad (5.4)$$

Values close to 1 mean the signal changes slowly and smoothly, typical of steady-state operation. Values near 0 or negative indicate rapid or irregular changes, as seen in unstable or noisy fault conditions. This feature helps separate faults that cause fast fluctuations from those that evolve gradually.

FFT Band Energies The FFT band energies describe how the signal’s power is distributed across different frequency ranges. After applying the Fast Fourier Transform (FFT), the total energy in each frequency band is computed:

$$E_{\text{band}_j} = \sum_{k \in \text{band}_j} |X_k|^2 \quad (5.5)$$

Low-frequency bands capture slow changes, such as thermal drift, while high-frequency bands reflect sudden oscillations or noise bursts. By dividing the frequency spectrum into a few bands (for example, low, medium, and high) and calculating the energy of each, the model learns to recognize which frequencies dominate during each degradation mode.

5.3.2 Handling Class Imbalance via Oversampling

The dataset considered in this study comprises four SOFC degradation modes: redox cycling, fuel starvation, short circuit, and thermal shock. Although all four degradation types are represented in the experimental campaigns, their occurrence frequencies are highly imbalanced. In particular, redox cycling is a dominant degradation mode, resulting in a substantially larger number of samples compared to the other degradations. In contrast, short circuit and thermal shock experiments consist of only a few short degradation runs, while fuel starvation events are also relatively less frequent.

This pronounced class imbalance introduces a significant bias in data-driven

classifiers such as RF and XGB, which tend to favor the majority class (redox cycling). As a consequence, these models may achieve high overall accuracy while exhibiting poor recall and F1-scores for minority degradation modes, indicating limited ability to learn meaningful decision boundaries for rare degradation modes. To mitigate this issue, several oversampling techniques were employed to rebalance the class distribution by generating additional samples for minority classes. This rebalancing step enables the classifiers to better capture discriminative patterns across all degradation modes and improves degradation-specific classification performance.

Oversampling is a method that try to solve class imbalance in datasets by increasing the number of samples in the minority class. It help achieve a better balance in data distribution for machine learning models [155]. There are two common techniques include random oversampling, which duplicates existing minority samples, and SMOTE (Synthetic Minority Over-sampling Technique), which creates new synthetic samples by interpolating between existing minority samples [156]. Oversampling improves model performance by preventing it from being biased toward the majority class but carries the risk of overfitting, especially with random oversampling [157–159]. Studies have shown that oversampling can significantly improve the accuracy, precision, recall, and F1-scores of predictive models in applications ranging from electricity theft detection to academic performance prediction and health diagnostics [160, 161]. Let:

- N_{maj} : number of majority class samples
- N_{min} : number of minority class samples
- $r = N_{\text{maj}}/N_{\text{min}}$: the imbalance ratio

To achieve a desired ratio r_0 (e.g., 1 for full balance), the number of additional synthetic minority samples needed is approximately:

$$N_{\text{new}} = N_{\text{maj}} \times r_0 - N_{\text{min}} \quad (5.6)$$

After oversampling:

$$N'_{\min} = N_{\min} + N_{\text{new}}, \quad N'_{\text{maj}} = N_{\text{maj}} \quad (5.7)$$

Random Oversampling (ROS) Random oversampling is the simplest balancing strategy, in which minority-class samples are randomly replicated with replacement until a target class ratio is achieved. While this method maintains the original statistical distribution of the minority class, it may increase the risk of overfitting, as repeated instances can lead the classifier to memorize duplicated samples rather than learn generalizable patterns. Because no new data points are generated, the feature values of the replicated samples remain unchanged. ROS (with replacement) is defined as:

$$D_{\text{new}} = N_{\min} \cup \{X_i \mid X_i \in D_{\min}, i \sim U(1, N_{\min})\}^{N_{\text{new}}} \quad (5.8)$$

Empirical tests [162] reveal that utilizing shallow learners like logistic regression, SVM, and random forest, ROS can greatly enhance recall and F1-score for minority classes.

Synthetic Minority Oversampling Technique (SMOTE) SMOTE is an enhanced oversampling method that generates synthetic samples rather than duplicating existing minority instances. For each minority sample x_i , one of its k -nearest neighbors $x_{i,\text{NN}}$ (also belonging to the minority class) is randomly selected, and a new synthetic point is generated as

$$x_{\text{new}} = x_i + \lambda (x_{i,\text{NN}} - x_i), \quad \lambda \sim U(0, 1). \quad (5.9)$$

This interpolation along the line segment connecting two minority samples produces new data points within the minority feature space, which helps reduce overfitting and improves coverage of the decision boundary.

Dozens of SMOTE variations have been suggested in recent research[162, 163], among them Borderline-SMOTE, ADASYN, KMeans-SMOTE, and Targeted Oversampling, which dynamically adjust to the data distribution or boundary structure.

Borderline-SMOTE Borderline-SMOTE extends the standard SMOTE approach by generating synthetic samples primarily near the class decision boundary, where misclassification is most likely to occur. Minority samples located close to majority instances contain the most informative and discriminative patterns for learning accurate decision boundaries. Algorithm Steps:

- Find the k -nearest neighbors for every minority sample x_i over every class.
- Count how many neighbors belong to the majority class, $n_{\text{maj}}(x_i)$.
- If $n_{\text{maj}}(x_i) > \tau$ (e.g., half of k (in KNN)), mark x_i as “danger” (i.e., close to the boundary).
- Generate synthetic data by applying SMOTE only to “danger” samples:

$$X_{\text{new}} = X_i + \lambda(X_{i_{\text{NN}}} - X_i), \quad X_{i_{\text{NN}}} \in \text{Danger Region} \quad (5.10)$$

Through oversampling, this targets unclear regions and improves boundary learning while preventing well-isolated samples from creating noise amplification. Adaptive Borderline-SMOTE [164] and Gaussian Borderline-SMOTE [165], which dynamically change the neighbor count or introduce stochastically disturbances to increase diversity, are recent advancements.

Targeted Oversampling Targeted Oversampling, an improved SMOTE variant sometimes known as Adaptive or Focused Oversampling, intentionally assigns artificial samples unevenly across minority points [166–168]. Rather than treating every minority instance equally (as in SMOTE), targeted oversampling identifies difficult-to-learn or underrepresented areas, typically those:

- close to the decision boundary,
- in sparsely populated regions of the minority space, or
- exhibiting high regional classification complexity.

Then, it generates more synthetic samples particularly in those areas to facilitate class separation. This method is occasionally implemented through:

- Less weight should be given to sparse minority groups when density is the focus.
- Over-sample points close to the majority samples (high local impurity).
- Combining both demands dynamically: hybrid directed sampling.

Recent studies [162, 169] show that such uneven oversampling reduces redundancy and noise, while enhancing minority recall and F1-score, particularly for highly imbalanced datasets (imbalance ratio $> 1:20$). The algorithm steps are:

- $D_{\min} = \{x_1, x_2, \dots, x_{N_{\min}}\}$: minority samples.
- For each sample x_i , calculate a difficulty score $\delta_i \geq 0$, directly proportional to how “hard” that instance is to classify (e.g., number of nearby majority neighbors).

$$\delta_i = \frac{n_{\text{maj}}(x_i)}{k}, \quad (5.11)$$

where $n_{\text{maj}}(x_i)$ is the number of majority samples among the k -nearest neighbors of x_i .

- Next, compute the normalized sampling probability p_i for each minority sample:

$$p_i = \frac{\delta_i}{\sum_{j=1}^{N_{\min}} \delta_j}. \quad (5.12)$$

- Then, the number of synthetic samples assigned to each minority instance is:

$$N_i = \text{round}(p_i \times N_{\text{new}}), \quad (5.13)$$

where N_{new} is the total number of synthetic points to generate.

- Finally, generate N_i new synthetic samples around x_i using a standard interpolation formula (as in SMOTE):

$$x_{\text{new}} = x_i + \lambda(x_{i_{\text{NN}}} - x_i), \quad \lambda \sim U(0, 1), \quad x_{i_{\text{NN}}} \in N_k(x_i, \text{minority}). \quad (5.14)$$

Thus, difficult or sparse minority regions get denser synthetic sampling, while well-represented regions receive fewer new samples.

Targeted Oversampling provides adaptive control over where to generate synthetic data. By focusing on locally difficult regions, it enhances minority-class recall and reduces overfitting compared to uniform oversampling. Contemporary empirical studies [162, 169, 170] validates that Targeted or Cluster-aware oversampling combined with shallow classifiers (RF, XGB, SVM) significantly boosts F1-score and G-Mean on benchmark imbalanced datasets.

In this study, a tailored oversampling strategy was implemented to address class imbalance. The short circuit class, which contained very few isolated samples, was expanded using ROS. In contrast, the thermal shock class exhibited substantial overlap with other classes and was therefore augmented using Borderline-SMOTE, which emphasizes generating synthetic samples in the vicinity of decision boundaries.

5.3.3 Random Forest Classifier

RF is one of the most important algorithms in machine learning that is used in various fields and shows significant findings and improvements [171, 172]. This algorithm is mainly used in classification and regression tasks by creating multiple decision trees and combining their results to increase accuracy and generalize to new data [172]. There are also recent works like broad granular RF algorithm that improves classification performance by granular computing to manage uncertain data more effectively [173]. This new approach outperformed traditional RF algorithms on some UCI datasets. The RF algorithm has proven to be a highly versatile and reliable tool across diverse domains. It has delivered robust results in bridge construction risk assessment [174], and improved early diabetes detection [175], while also enhancing urban mobility systems through cloud-based integration for faster and more efficient bus ticketing [176]. In healthcare, it has been used to address data imbalance in preventing stunting among toddlers [177]. Moreover, studies show that RF consistently outperforms classifiers like SVM and k-Nearest Neigh-

bors (KNN), achieving superior accuracy in applications such as urban tree classification and Internet of Things (IoT) network security through innovative feature selection strategies [178].

RF models are trained using the bootstrap aggregation (bagging) technique. Given a training dataset D and a total of B decision trees, the algorithm constructs B bootstrap datasets by repeatedly sampling from D with replacement. Each tree is trained independently on one bootstrap dataset, and the final model output is obtained by aggregating the predictions of all trees through averaging for regression or majority voting for classification [179].

The original training dataset is defined as

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}, \quad (5.15)$$

where x_i denotes the input feature vector and y_i represents the corresponding target label.

For each tree $b = 1, \dots, B$, the following procedure is applied:

- A bootstrap dataset D_b is generated by randomly selecting n samples from D with replacement. As a result, some samples may appear multiple times, while others may be omitted.
- A decision tree T_b is trained using the bootstrap dataset D_b . At each split node, a subset of input variables is randomly selected (typically $m = \sqrt{p}$, where p is the total number of features), and the best split is chosen from this subset based on an impurity criterion.

For regression tasks, the RF prediction is obtained by averaging the outputs of all individual trees:

$$F(x) = \frac{1}{B} \sum_{b=1}^B T_b(x), \quad (5.16)$$

where $T_b(x)$ denotes the prediction of the b -th tree. In classification problems, the final class label is determined by majority voting among the predictions of all trees.

5.3.4 XGBoost Classifier

Gradient Boosted Decision Trees (GBDT), including XGBoost [180], are constructed iteratively by adding weak learners one by one, typically shallow decision trees. Each newly added tree aims to improve the model by minimizing the residual error (loss) of the preceding ensemble. The model prediction is given by

$$\hat{y}_i = \sum_{t=1}^T f_t(\mathbf{x}_i), \quad (5.17)$$

where each f_t represents an individual decision tree.

XGBoost optimizes an objective function defined as the sum of a differentiable loss function (e.g., squared loss) and a regularization term Ω that penalizes model complexity [180]. This combination enables the algorithm to balance fitting accuracy and model complexity, thereby reducing the risk of overfitting while maintaining strong predictive performance [181]. Designed explicitly for high accuracy and scalability on large datasets, XGBoost employs efficient learning strategies based on parallel computation and regularization-aware optimization [182]. These advantages are regarded as some of the key innovations that distinguish XGBoost from other advanced machine learning algorithms [183].

XGBoost achieves high predictive accuracy by utilizing both first- and second-order derivatives during the optimization process, allowing more effective correction of prediction errors [180]. Overfitting is further mitigated through built-in regularization mechanisms, including L1 and L2 penalties, which help control model complexity [181]. In addition, XGBoost benefits from a parallel and memory-efficient architecture, enabling fast training and making it suitable for large-scale datasets [180]. The framework supports regression, classification, ranking, and other learning tasks based on the boosting paradigm. Moreover, XGBoost performs well on imbalanced datasets and provides feature importance scores, offering transparency by quantifying the contribution of each feature to the final prediction [183].

5.3.5 RF-XGB Ensemble Model

An ensemble model that combines RF and XGB leverages the complementary strengths of both methods [184, 185]. Specifically, RF reduces variance through bagging and random feature selection, while XGB reduces bias via gradient boosting and regularized optimization. In this ensemble framework, the predictions of the two models are combined using weighted averaging for regression tasks or voting/stacking for classification, resulting in a balanced bias–variance trade-off in which RF provides stability and XGB enhances gradient-based accuracy.

Let the base learners be defined as [186, 187]:

$$\hat{y}_{\text{RF}} = f_{\text{RF}}(\mathbf{x}), \quad \hat{y}_{\text{XGB}} = f_{\text{XGB}}(\mathbf{x}), \quad (5.18)$$

where each model is trained independently using the same feature vector $\mathbf{x}$.

The ensemble prediction is then expressed as:

$$\hat{y}_{\text{ENS}} = \alpha \hat{y}_{\text{RF}} + (1 - \alpha) \hat{y}_{\text{XGB}}, \quad 0 \leq \alpha \leq 1, \quad (5.19)$$

where α and $(1 - \alpha)$ denote the weights assigned to the RF and XGB models, respectively.

5.4 Deep Models for Degradation Mode Classification

To better exploit the sequential nature of the DRT-peak and R_{ohm} time-series, this section investigates deep learning models that operate directly on windowed multivariate sequences. A physically informed augmentation and screening pipeline is first introduced to improve data diversity while preserving realistic degradation dynamics. Building on this enhanced training set, three representative deep temporal architectures are evaluated: a LSTM network as a recurrent baseline for sequence classification, a TCN as a convolutional alternative with dilated receptive fields, and an attention-based Informer model.

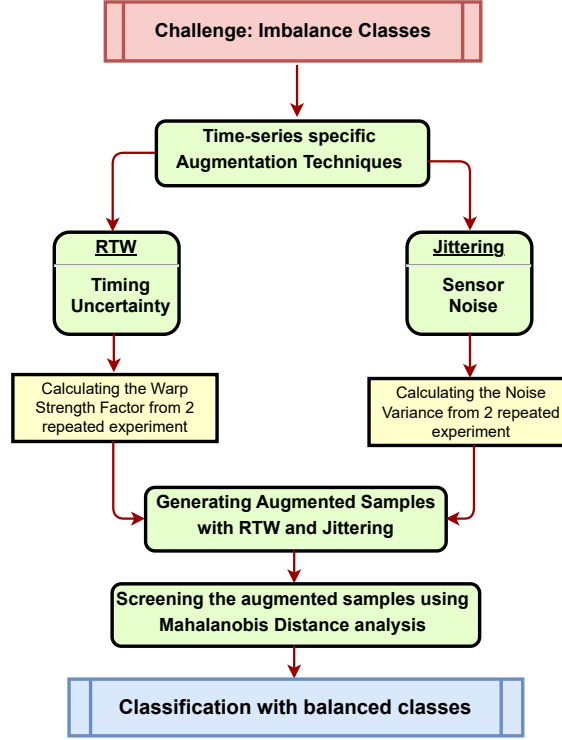


Figure 5.2: Overview of the proposed time-series-specific data augmentation and screening pipeline for addressing class imbalance. The framework combines RTW and amplitude jittering, with data-driven calibration from repeated experiments. Augmented samples are subsequently screened using Mahalanobis distance analysis to ensure physical plausibility.

5.4.1 Physically Informed Time-Series Augmentation and Screening

For deep sequence models such as LSTM, TCN, and Informer, conventional oversampling strategies are inadequate because they ignore the temporal dependencies inherent in time-series data. To address this limitation, a two-stage physically informed time-series augmentation strategy is adopted, followed by a statistical screening step to ensure physical plausibility of the generated samples. An overview of the proposed augmentation and filtering pipeline is illustrated in Figure 5.2.

Random Time Warping (RTW) Let the original time-series segment be denoted as

$$\mathbf{X}_{\text{real}} = \{a_1, a_2, \dots, a_n\}, \quad a_i = X(t_i), \quad (5.20)$$

where the normalized time grid is defined as

$$t_i = \frac{i-1}{n-1}, \quad i = 1, \dots, n. \quad (5.21)$$

RTW introduces smooth, monotonic perturbations along the time axis to emulate timing uncertainty and repeatability errors commonly observed in SOFC experiments. A warped time grid is constructed as

$$\tau_i = t_i + \alpha \epsilon_i, \quad (5.22)$$

where $\epsilon_i \sim \mathcal{N}(0, 1)$ represents random perturbations and α denotes the warp strength controlling the magnitude of temporal distortion. To preserve temporal causality, a monotonicity constraint,

$$\tau_1 < \tau_2 < \dots < \tau_n, \quad (5.23)$$

is enforced. The augmented signal is obtained by resampling the original signal over the warped timeline,

$$\mathbf{X}_{\text{RTW}} = \{b_1, b_2, \dots, b_n\}, \quad b_i = a(\tau_i). \quad (5.24)$$

This operation locally stretches or compresses the time axis while preserving the global degradation trajectory, thereby producing realistic temporal variations without altering the underlying physical mechanisms.

Amplitude Jittering Following RTW, amplitude jittering is applied to model measurement noise and small experimental fluctuations present in EIS/DRT-derived features. Each signal sample is perturbed as

$$b_i = a_i + \sigma_{\text{jitter}} \epsilon_i, \quad (5.25)$$

where σ_{jitter} denotes the jitter strength and $\epsilon_i \sim \mathcal{N}(0, 1)$. This augmentation step captures sensor noise, impedance measurement drift, and peak extraction variability while maintaining the overall structure of the signal.

Calibration of Augmentation Strengths Both augmentation parameters are calibrated directly from experimental data obtained under identical operating conditions to ensure physical consistency. The calibration procedure relies on robust statistics derived from repeated experiments.

For RTW, relative timing errors between two repeated experiments are computed as

$$r(j) = \frac{t_B(j) - t_A(j)}{\frac{1}{2}(|t_A(j)| + |t_B(j)|) + \varepsilon}, \quad (5.26)$$

where $t_A(j)$ and $t_B(j)$ denote corresponding time indices in the two experiments, and ε is a small constant introduced to prevent numerical instability. The coefficient of variation based on the median absolute deviation (MAD) is then calculated as

$$\widehat{CV}_{\text{time}} = 1.4826 \cdot \text{median}_j |r(j) - \text{median}_j r(j)|. \quad (5.27)$$

The warp strength is obtained by bounding this quantity within predefined limits,

$$\alpha = \text{clip}\left(\widehat{CV}_{\text{time}}, \alpha_{\min}, \alpha_{\max}\right). \quad (5.28)$$

For amplitude jittering, channel-wise measurement variability is estimated from pointwise differences between repeated measurements. For each channel k , the difference sequence is defined as

$$d_k(j) = A_k(j) - B_k(j), \quad j = 1, \dots, M, \quad (5.29)$$

where A_k and B_k represent repeated realizations of the same signal channel. The Gaussian-equivalent standard deviation is computed using a MAD-based estimator,

$$\hat{\sigma}_k^{(\text{raw})} = 1.4826 \cdot \text{median}_j |d_k(j) - \text{median}_j d_k(j)|. \quad (5.30)$$

This raw estimate is converted to the standardized feature space as

$$\hat{\sigma}_k^{(\text{std})} = \frac{\hat{\sigma}_k^{(\text{raw})}}{s_k}, \quad (5.31)$$

where s_k denotes the scaling factor of channel k .

In this study, the jitter standard deviation is calibrated independently for each

signal channel to capture channel-specific measurement noise, while a single global warp strength is used to represent timing uncertainty common across all channels. The resulting standardized jitter values are

$$\sigma_{\text{Peak}_1} = 0.009890, \sigma_{\text{Peak}_2} = 0.028251, \sigma_{\text{Peak}_3} = 0.028529, \sigma_{\text{Ohmic}} = 0.008536,$$

and the calibrated global warp strength is

$$\alpha = 0.052.$$

Mahalanobis-Distance-Based Screening Because data-driven augmentation may occasionally generate physically implausible sequences, a Mahalanobis-distance-based screening step is applied to filter out unrealistic samples. Each windowed signal is summarized into a compact D-dimensional statistical feature vector, where $D = 24$ in this study, consisting of the mean, standard deviation, percentiles, and range computed across all signal channels.

For each degradation class, a multivariate Gaussian distribution is fitted to the summary vectors of real samples, characterized by mean $\boldsymbol{\mu}_c$ and covariance $\mathbf{C}_c$. The squared Mahalanobis distance of an augmented sample $\mathbf{s}$ is computed as

$$d^2 = (\mathbf{s} - \boldsymbol{\mu}_c)^\top \mathbf{C}_c^{-1} (\mathbf{s} - \boldsymbol{\mu}_c). \quad (5.32)$$

Only augmented samples whose distances fall within the 95th percentile of the real-data distance distribution are retained. This screening step ensures that the augmented dataset remains statistically consistent with true SOFC behavior while still enriching minority classes.

Overall, the combination of physically informed time-series augmentation and rigorous statistical screening produces a balanced, high-quality training set that significantly improves classification performance without compromising physical realism.

5.4.2 LSTM Classifier

LSTM networks have been widely adopted as effective classifiers for time-series datasets due to their ability to model sequential patterns [188–190]. The architecture of the LSTM-based classifier is illustrated in Figure 5.3. The model

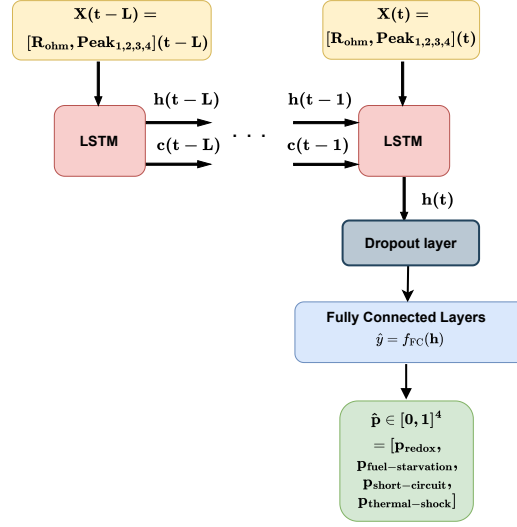


Figure 5.3: Architecture of the LSTM classifier for degradation mode identification using windowed time-series inputs $\mathbf{X}_{t-L:t} = [R_{ohm}, P_1, P_2, P_3, P_4]_{t-L:t}$.

operates on a multivariate time-series input window denoted as $\mathbf{x}_{1:L} \in \mathbb{R}^{L \times F}$, where each sample consists of L consecutive time steps and F feature channels per step. In this study, the input features comprise the amplitudes of the extracted DRT peaks together with the ohmic resistance (R_{ohm}), which collectively characterize the electrochemical state of the SOFC over time. The parameter L represents the temporal look-back window length, while F denotes the number of input features per time step.

Prior to network ingestion, each feature channel is independently standardized to zero mean and unit variance in order to reduce scale disparities and improve numerical stability during training. The normalized input sequence is then processed by a single LSTM layer with 64 hidden units, which is designed to capture both short-term dynamics and long-range temporal dependencies in the degradation signatures.

At each time step t , the LSTM updates its internal states according to the

following gating equations [191]:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f), \quad (5.33)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i), \quad (5.34)$$

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c), \quad (5.35)$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t, \quad (5.36)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o), \quad (5.37)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t), \quad (5.38)$$

where $\mathbf{f}_t$, $\mathbf{i}_t$, and $\mathbf{o}_t$ denote the forget, input, and output gates, respectively; $\tilde{\mathbf{c}}_t$ is the candidate cell state; $\mathbf{c}_t$ is the cell memory; and $\mathbf{h}_t$ is the hidden state at time step t . The operator $\odot$ denotes element-wise multiplication, $\sigma(\cdot)$ is the sigmoid activation function, and $\tanh(\cdot)$ is the hyperbolic tangent function. The matrices $\mathbf{W}_*$ and $\mathbf{U}_*$ and vectors $\mathbf{b}_*$ represent trainable parameters of the LSTM.

After processing the full input sequence, the final hidden state $\mathbf{h}_L$ provides a compact representation of the temporal evolution of the input features across the entire window. This latent representation is passed to a fully connected (FC) layer that maps $\mathbf{h}_L$ to a four-dimensional logits vector:

$$\mathbf{z} = \mathbf{W}_{\text{fc}} \mathbf{h} + \mathbf{b}_{\text{fc}}, \quad (5.39)$$

where each element of $\mathbf{z}$ corresponds to one SOFC degradation mode. A softmax activation function is then applied to obtain normalized class probabilities:

$$\hat{\mathbf{p}} = \text{softmax}(\mathbf{z}), \quad (5.40)$$

with $\hat{\mathbf{p}} \in [0, 1]^4$ and $\sum_k \hat{p}_k = 1$. The predicted degradation mode is determined as

$$\hat{y} = \arg \max_k \hat{p}_k. \quad (5.41)$$

The network is trained using the categorical cross-entropy loss function, optimized via the Adam optimizer. To mitigate overfitting and improve gen-

eralization, early stopping based on the validation loss is employed during training.

5.4.3 TCN Classifier

TCNs have emerged as an effective alternative to recurrent architectures for time-series classification by leveraging causal and dilated convolutions to model long-range temporal dependencies [192, 193]. Unlike recurrent neural networks, TCNs process entire sequences in parallel and avoid sequential hidden-state updates, resulting in improved training stability and computational efficiency. Owing to these advantages, TCNs have been successfully applied to a wide range of sequence modeling tasks, including action segmentation [192], fault diagnosis [194], and temporal pattern classification [195].

From an architectural perspective, TCNs belong to the class of one-dimensional convolutional neural networks (CNNs) designed specifically for sequential data. In contrast to standard temporal CNNs with limited receptive fields, TCNs introduce two key innovations: *causal convolutions*, which ensure that predictions at time t depend only on current and past inputs, and *dilated convolutions*, which exponentially expand the receptive field without a proportional increase in model depth or parameter count [192]. These properties enable TCNs to capture long-term temporal structures that are critical for degradation and degradation mode analysis in SOFC systems.

The TCN classifier used in this study is illustrated in Figure 5.4. The network receives a multivariate input sequence $\mathbf{x}_{1:L} \in \mathbb{R}^{L \times F}$, where L denotes the temporal window length and F represents the number of features per time step, corresponding to the DRT peak amplitudes and the ohmic resistance (R_{ohm}). Prior to convolutional processing, each feature channel is independently standardized to ensure numerical stability and balanced gradient propagation.

At the core of the model are stacked dilated causal convolutional blocks. For a one-dimensional input signal $\mathbf{x}$, a dilated convolution operation with dilation

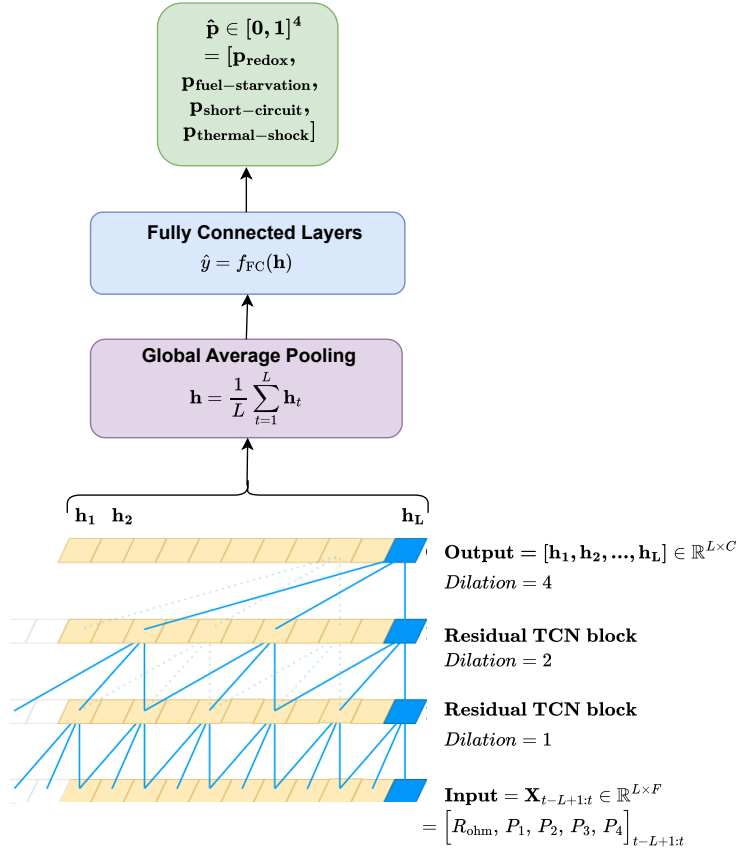


Figure 5.4: Architecture of the TCN classifier consisting of residual dilated causal convolution blocks, followed by global temporal pooling and fully connected layers for degradation mode prediction.

factor d and kernel size k is defined as

$$y(t) = \sum_{i=0}^{k-1} w_i x(t - di), \quad (5.42)$$

where i indexes the convolution kernel taps and w_i are the convolutional filter weights. By increasing the dilation factor exponentially across layers (e.g., $d = 1, 2, 4$), the receptive field grows rapidly, allowing the network to integrate information over long temporal horizons while maintaining a shallow architecture. Each TCN block consists of two successive dilated causal convolution layers, followed by batch normalization, ReLU activation, and spatial dropout for regularization. Residual connections are incorporated to facilitate gradient flow and stabilize deep network training. Let $\mathbf{h}^{(l)}(t)$ denote the activation (fea-

ture representation) at layer l and time t . The residual block can be expressed as

$$\mathbf{h}^{(l)}(t) = \text{ReLU}(\mathbf{h}^{(l-1)}(t) + \mathcal{F}(\mathbf{h}^{(l-1)}(t))), \quad (5.43)$$

where $\mathcal{F}(\cdot)$ represents the stacked dilated convolutional transformations within the block. When the number of feature channels changes, a 1×1 convolution is applied to the residual path to ensure dimensional consistency.

After passing through the sequence of dilated convolutional blocks, the resulting feature maps encode multi-scale temporal patterns spanning the entire input window. A global average pooling operation is then applied along the temporal dimension to aggregate the learned representations into a fixed-length vector:

$$\mathbf{z} = \frac{1}{L} \sum_{t=1}^L \mathbf{h}(t). \quad (5.44)$$

This pooled representation is subsequently passed through a FC layer with ReLU activation, followed by a softmax output layer that produces class probability estimates

$$\hat{\mathbf{p}} = \text{softmax}(\mathbf{W}_{\text{fc}}\mathbf{z} + \mathbf{b}_{\text{fc}}), \quad (5.45)$$

where $\hat{\mathbf{p}} \in [0, 1]^C$ and C denotes the number of degradation mode classes. The predicted class label is obtained as

$$\hat{y} = \arg \max_k \hat{p}_k. \quad (5.46)$$

5.4.4 Attention-based Informer Classifier

Originally, the Informer model was proposed for long-sequence time-series forecasting using an encoder-decoder structure. In this work, we repurpose the Informer as a *time-series classifier* by retaining only the encoder component and replacing the forecasting decoder with a lightweight multilayer perceptron (MLP). The encoder is used to efficiently compress historical multivariate measurements into a compact latent representation, which is subsequently mapped to degradation mode class probabilities.

Given an input sequence of length L , self-attention operates on three projected

representations: queries $\mathbf{Q}$, keys $\mathbf{K}$, and values $\mathbf{V}$. The canonical scaled dot-product attention is defined as

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right) \mathbf{V}, \quad (5.47)$$

where d_k denotes the dimensionality of the key vectors. This formulation enables each time step to attend directly to all others, providing a path length of $\mathcal{O}(1)$ between distant temporal events, which is particularly advantageous for degradation processes evolving over long horizons. However, computing full attention incurs quadratic complexity $\mathcal{O}(L^2)$ in both time and memory, which becomes prohibitive for long time-series inputs.

The Informer addresses this limitation by exploiting the observation that attention weights exhibit a long-tail distribution, where only a small subset of query–key pairs contributes significantly to the output [150]. Informer introduces the ProbSparse self-attention mechanism, which selectively retains only the most informative queries. Only the top- u queries with the highest sparsity scores are retained. This reduces the computational complexity of self-attention to approximately $\mathcal{O}(L \log L)$ while preserving dominant temporal dependencies.

The Informer encoder consists of stacked ProbSparse attention blocks, each followed by a self-attention distilling operation that progressively reduces the temporal resolution. Let $\mathbf{X}^{(j)} \in \mathbb{R}^{L_j \times d}$ denote the encoder representation at layer j . The distilling operation is defined as

$$\mathbf{X}^{(j+1)} = \text{MaxPool}(\text{ELU}(\text{Conv1d}([\mathbf{X}^{(j)}]_{\text{AB}}))), \quad (5.48)$$

where $[\cdot]_{\text{AB}}$ denotes the output of the ProbSparse attention block, Conv1d is a one-dimensional convolution along the temporal axis, and MaxPool downsamples the sequence with stride 2. This operation halves the sequence length at successive layers, sharply reducing memory usage while emphasizing dominant temporal patterns. Multiple encoder stacks operating on progressively shorter input slices are concatenated to form the final encoder output, yielding a robust multi-scale representation of the input sequence.

In contrast to the original forecasting formulation, the decoder is omitted in this work. The final encoder output is aggregated along the temporal dimension using global average pooling, producing a fixed-length feature vector $\mathbf{z} \in \mathbb{R}^d$. This representation is passed to a FC classifier:

$$\mathbf{o} = \mathbf{W}_{\text{fc}}\mathbf{z} + \mathbf{b}_{\text{fc}}, \quad (5.49)$$

where $\mathbf{o} \in \mathbb{R}^C$ denotes the logits for C degradation classes. A softmax activation converts the logits into class probabilities,

$$\hat{p}_c = \frac{\exp(o_c)}{\sum_{k=1}^C \exp(o_k)}, \quad (5.50)$$

and the predicted degradation mode is obtained as

$$\hat{y} = \arg \max_c \hat{p}_c. \quad (5.51)$$

The network is trained using the categorical cross-entropy loss with the Adam optimizer, and early stopping based on validation loss is employed to mitigate overfitting. Figure 5.5 presents the Informer-based classifier architecture.

5.5 Results and Discussion

5.5.1 Quantitative Evolution of DRT-Derived Features

To further support the mechanistic interpretation of the DRT-derived features, a quantitative feature evolution analysis was performed by computing the percentage change of each feature between the BOL and EOL conditions for all degradation experiments. The resulting mean and standard deviation values across the three repetitions of each failure mode are summarized in Table 5.1. Distinct degradation signatures are observed for each failure mode. Redox cycling is characterized by substantial increases in peak 2 (87.7%) and peak 1 (26.7%), indicating degradation of anode activation processes and oxygen-ion transport pathways. Fuel starvation exhibits a pronounced increase in ohmic resistance (71.0%) together with a moderate increase in peak 2 (24.3%), consistent with conductivity loss and degradation of anode electrochemical activity

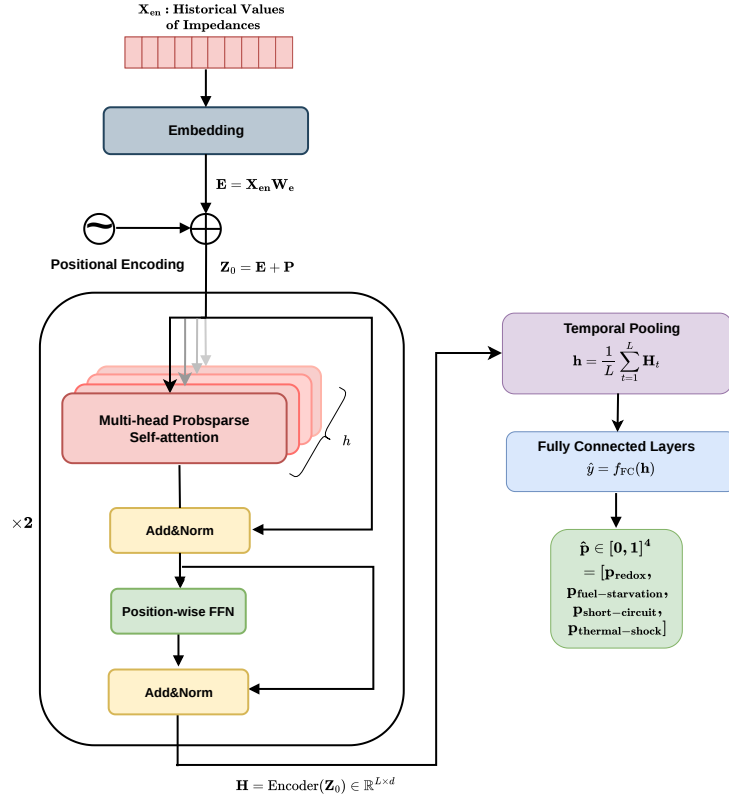


Figure 5.5: Architecture of the Informer classifier, including temporal embedding, multi-head ProbSparse self-attention encoding, temporal pooling, and a fully connected prediction head for degradation mode classification.

caused by prolonged hydrogen deficiency and local re-oxidation effects. Short circuit degradation is dominated by an increase in ohmic resistance (72.3%), while all DRT peaks remain relatively unchanged, suggesting that the primary degradation pathway is associated with resistive losses rather than electrochemical polarization processes. Thermal shock produces simultaneous increases in peak 3 (52.0%) and ohmic resistance (47.6%), indicating degradation of cathode-related processes together with increased resistive losses caused by thermally induced cracking and delamination.

These results quantitatively confirm that each failure mode produces a distinct evolution pattern in the DRT-derived features and provide additional electrochemical justification for the mechanistic interpretability of the proposed diagnostic framework. Moreover, the consistent feature-evolution trends observed

across the three repeated experiments for each failure mode support the repeatability of the degradation signatures and indicate that cells subjected to the same failure-inducing condition exhibit similar electrochemical behavior.

Table 5.1: Percentage change of DRT-derived features and ohmic resistance from BOL to EOL. Values represent mean $\pm$ standard deviation across three repeated experiments for each failure mode.

Failure Mode	R_{ohm} (%)	Peak 1 (%)	Peak 2 (%)	Peak 3 (%)
Redox Cycling	0.46 ± 6.20	26.70 ± 9.64	87.70 ± 2.48	5.96 ± 2.90
Fuel Starvation	71.04 ± 3.27	5.83 ± 0.35	24.27 ± 6.11	6.84 ± 8.02
Short Circuit	72.27 ± 3.75	1.47 ± 0.80	1.03 ± 2.06	1.59 ± 4.35
Thermal Shock	47.63 ± 3.55	4.22 ± 5.60	4.33 ± 1.87	52.03 ± 7.17

5.5.2 Evaluation Strategy: LODO

To assess model generalization performance, a Leave-One-Dataset-Out (LODO) cross-validation strategy was adopted. Each experimental file corresponds to a complete and independent test run. In each validation iteration, the model was trained using all files except one and evaluated on the held-out file. This procedure was repeated until every file had been used once as the test set.

This validation strategy offers two key advantages. First, it effectively prevents data leakage. Sliding-window samples extracted from the same experimental file are strongly correlated; randomly splitting such samples into training and testing sets would lead to overly optimistic performance estimates. LODO enforces a strict separation between training and testing data by ensuring that no samples from the same experiment appear in both sets. Second, LODO closely reflects real-world deployment scenarios, in which a model trained on historical experiments is required to identify degradation modes in a new experiment with previously unseen operating conditions. As a result, LODO provides a realistic evaluation of model robustness and generalization capability.

5.5.3 Evaluation Metrics

Seven evaluation metrics were used to assess classification performance: Macro- and Weighted-averaged variants of Precision, Recall, and F1-score, along with

overall Accuracy. For each class k , Precision, Recall, and F1-score are defined as

$$\text{Precision}_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FP}_k}, \quad (5.52)$$

$$\text{Recall}_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FN}_k}, \quad (5.53)$$

$$\text{F1}_k = 2 \times \frac{\text{Precision}_k \times \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k}. \quad (5.54)$$

Precision measures the proportion of correctly predicted samples among all samples predicted as class k , whereas Recall quantifies the proportion of true samples of class k that are correctly identified. The F1-score is the harmonic mean of Precision and Recall and penalizes large imbalances between the two. To summarize performance across multiple classes, Macro- and Weighted-averaged metrics are reported. Using the F1-score as an example, the Macro-F1 is computed as the unweighted average of the class-wise F1-scores:

$$\text{Macro-F1} = \frac{1}{K} \sum_{k=1}^K \text{F1}_k, \quad (5.55)$$

where K denotes the total number of classes. This formulation assigns equal importance to all classes and is therefore well suited for imbalanced datasets. The Weighted-F1 accounts for class frequency by weighting each class F1-score by its support:

$$\text{Weighted-F1} = \sum_{k=1}^K \frac{N_k}{N} \text{F1}_k, \quad (5.56)$$

where N_k is the number of true samples of class k , and N is the total number of samples. The same Macro and Weighted averaging procedures are applied to Precision and Recall to obtain Macro-Precision, Weighted-Precision, Macro-Recall, and Weighted-Recall.

Finally, overall classification accuracy is defined as

$$\text{Accuracy} = \frac{\sum_k \text{TP}_k}{N}, \quad (5.57)$$

which represents the fraction of correctly classified samples across all classes. However, in the presence of class imbalance, accuracy alone can be misleading. Consequently, Macro-averaged metrics, particularly Macro-F1, are emphasized as the primary indicators of balanced classification performance.

5.5.4 Hyperparameters Tuning

Hyperparameters for RF, XGB, and the RF+XGB ensemble were tuned via random search. For deep models (LSTM, TCN, Informer), random search was also adopted due to its simplicity and strong empirical performance. Trials were evaluated using a validation split drawn from training files, with early stopping and based on the Macro-F1 metric. Hyperparameters were selected using a subset of four experimental files, with one file selected from each degradation mode to ensure balanced representation during tuning. The selected hyperparameters were then fixed and used for LODO cross-validation. Tables 5.2–5.6 summarize the hyperparameters, their respective search spaces, and the optimized values for each model. For each case, the total number of possible hyperparameter combinations and the number of random search trials, set to 30% of the total combinations, are also reported.

Table 5.2: Random Forest classifier hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Look-back window length	{250, 500, 750}	500
Class imbalance threshold	{0.6, 0.7, 0.8, 0.9}	0.7
Short-circuit to majority class ratio	{0.2, 0.3, 0.4}	0.3
Maximum neighbor count	{1, 2, 3, 4}	2
Number of trees in the forest	{300, 600, 1000, 1500}	1000
Maximum tree depth	{None, 10, 15, 20, 30}	15
Minimum samples required at a leaf node	{1, 2, 4, 8}	4
Fraction of features considered at each split	{0.3, 0.5, 0.7, sqrt }	0.5
Total combinations: 46,080. Random search trials: 13,824.		

Table 5.3: XGBoost classifier hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Look-back window length	{250, 500, 750}	500
Number of nearest neighbors	{1, 2, 3, 4, 5}	3
Number of boosting trees	{300, 500, 800, 1200}	500
Learning rate	{0.03, 0.05, 0.07}	0.05
Maximum tree depth	{4, 5, 6, 7}	5
Row subsampling ratio per tree	{0.7, 0.8, 0.9}	0.7
Column subsampling ratio per tree	{0.6, 0.7, 0.8}	0.7
L2 regularization coefficient	{0.5, 1.0, 2.0}	1.0
L1 regularization coefficient	{0.0, 0.1, 0.5}	0.0
Total combinations: 58,320. Random search trials: 17,496.		

Table 5.4: LSTM classifier hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
LSTM hidden units	{32, 64, 128}	64
Dropout rate	{0.1, 0.2, 0.3}	0.2
Dense layer size	{16, 32, 64}	32
Learning rate	$\{1 \times 10^{-3}, 5 \times 10^{-4}, 1 \times 10^{-4}\}$	1×10^{-4}
Batch size	{32, 64, 128}	32
Total combinations: 243. Random search trials: 73.		

Table 5.5: TCN classifier hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of filters per convolution	{16, 32, 64}	32
Kernel size	{3, 5, 7}	3
Dilation schedule	{(1,2,4), (1,2,4,8), (1,3,9)}	(1,2,4)
Spatial dropout in residual blocks	{0.10, 0.25, 0.40}	0.25
Final dropout (after global pooling)	{0.20, 0.40, 0.50}	0.20
Label smoothing	{0.00, 0.05, 0.10}	0.05
Total combinations: 729. Random search trials: 219.		

Table 5.6: Informer (encoder-only) classifier hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Model dimension	{64, 128, 256}	128
Number of attention heads	{2, 4, 8}	4
Encoder layers	{1, 2, 3}	2
Feed-forward dimension	{64, 128, 256}	256
Attention dropout	{0.05, 0.10, 0.20}	0.10
Residual/dropout (encoder)	{0.05, 0.10, 0.20}	0.10
ProbSparse factor (sparsity control)	{3, 5}	5
Total combinations: 1458. Random search trials: 437.		

Table 5.7: RF+XGB ensemble hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Look-back window length	{250, 500, 750}	500
Class imbalance threshold	{0.6, 0.7, 0.8, 0.9}	0.7
Short-circuit to majority-class ratio	{0.2, 0.3, 0.4}	0.3
Maximum neighbor count	{1, 2, 3, 4}	2
<i>RF sub-model</i>		
Number of trees	{300, 600, 1000, 1500}	1000
Maximum tree depth	{None, 10, 15, 20, 30}	15
Minimum samples per leaf	{1, 2, 4, 8}	4
Fraction of features per split	{0.3, 0.5, 0.7, sqrt}	0.5
<i>XGB sub-model</i>		
Number of boosting trees	{300, 600, 900}	600
Learning rate	{0.03, 0.05, 0.08}	0.05
Maximum tree depth	{3, 5, 7}	5
Row subsampling ratio	{0.7, 0.9, 1.0}	0.9
Column subsampling ratio	{0.6, 0.8, 1.0}	0.8
L2 regularization coefficient	{0.5, 1.0, 2.0}	1.0
L1 regularization coefficient	{0.0, 0.1, 0.5}	0.0
Total combinations: 25,194,240. Random search trials: 7,558,272.		

5.5.5 Ablation Study of the Time-Series Augmentation Strategy

To quantify the contribution of the proposed augmentation strategy, an ablation study was conducted using the Informer classifier, as summarized in Table 5.8. Four cases were evaluated: no augmentation, RTW only, jittering only, and the combined RTW+jittering approach. The results show that data augmentation primarily benefits the more challenging failure modes, namely short circuit and thermal shock. Without augmentation, the classifier achieved strong performance for redox cycling and fuel starvation (F1-scores of 1.00 and 0.97, respectively), but substantially lower performance for short circuit (0.78) and thermal shock (0.85), resulting in a Macro-F1 score of 0.90.

Applying RTW or jittering individually significantly improved the classification of short circuit and thermal shock degradation while maintaining strong performance for the remaining classes. The combined augmentation strategy achieved the best overall performance, increasing the short circuit and thermal-shock F1-scores to 0.98 and 0.99, respectively, and yielding the highest Macro-F1 score of 0.98. These results suggest that RTW and jittering provide complementary forms of variability, improving the robustness of the classifier to both temporal distortions and measurement variations while preserving the underlying degradation characteristics.

Table 5.8: Ablation study of the proposed augmentation strategy using the Informer classifier.

Augmentation	Redox Cycling-F1	Fuel Starvation-F1	Short Circuit-F1	Thermal Shock-F1	Macro-F1
None	1.00	0.97	0.78	0.85	0.90
RTW	0.99	0.96	0.97	0.98	0.98
Jittering	0.99	0.96	0.96	0.98	0.97
RTW + Jittering	0.99	0.96	0.98	0.99	0.98

5.5.6 Models Performance and Computational Comparisons

Overall Performance (Evaluation Metrics)

A comprehensive performance comparison of all six classification models of RF, XGB, RF+XGB Ensemble, LSTM, TCN, and Informer is summarized in Table 5.9. Among the shallow learning approaches, the RF classifier achieves strong and well-balanced performance, with a Macro-F1 score of 0.92. This result confirms that the combination of targeted oversampling and engineered statistical features is effective in mitigating class imbalance and enabling RF to learn discriminative decision boundaries across all degradation modes. In contrast, the XGB model exhibits noticeably lower Macro Precision (0.69) and Macro-F1 (0.71), despite maintaining relatively high Weighted metrics and Accuracy (0.89). This discrepancy indicates that XGB is more strongly influenced by majority classes. The RF+XGB ensemble recovers much of this lost performance, achieving a Macro-F1 of 0.92, identical to RF. This improvement suggests that ensembling stabilizes the predictions of XGB by leveraging the robustness of RF, effectively compensating for the bias introduced by gradient boosting under severe imbalance.

Transitioning to deep learning models, the LSTM classifier achieves a Macro-F1 of 0.92, matching the performance of RF and the ensemble. LSTM exhibits higher Macro Recall (0.96) than RF, reflecting its improved ability to capture temporal degradation patterns and detect minority degradation modes. The TCN model further improves upon LSTM performance, achieving a Macro-F1 of 0.96. This gain can be attributed to the use of dilated causal convolutions, which allow TCN to model long temporal contexts more efficiently and stably than recurrent architectures. The high Macro Recall (0.97) and Macro Precision (0.95) demonstrates that TCN effectively balances sensitivity and specificity across all degradation modes. The Informer model delivers the best overall performance across all reported metrics. By leveraging ProbSparse self-attention and efficient temporal representation learning, Informer achieves a Macro-F1 of 0.98, corresponding to an absolute improvement of approximately

Table 5.9: Performance comparison of degradation mode classifiers

Model	Macro Precision	Weighted Precision	Macro Recall	Weighted Recall	Macro F1	Weighted F1	Accuracy
RF	0.95	0.97	0.91	0.97	0.92	0.96	0.97
XGB	0.69	0.90	0.75	0.89	0.71	0.89	0.89
Ensemble	0.94	0.97	0.90	0.97	0.92	0.96	0.97
LSTM	0.90	0.95	0.96	0.95	0.92	0.95	0.95
TCN	0.95	0.96	0.97	0.95	0.96	0.95	0.95
Informer	0.98	0.98	0.98	0.98	0.98	0.98	0.98

6.5% over RF, LSTM, and the ensemble, and a 3% improvement over TCN. Informer also attains uniformly high Macro Precision and Macro Recall (both 0.98), indicating consistent and reliable discrimination across both majority and minority degradation modes.

Class-wise Performance Analysis (Confusion Matrices)

Figures 5.6 present the confusion matrices for all classifiers, enabling a class-wise analysis of model behavior. Across all models, redox cycling is consistently identified with high accuracy, confirming that its impedance signatures are highly distinctive. Short circuit is also reliably detected due to its abrupt and high-contrast electrochemical response. The primary source of misclassification across models arises between fuel starvation and thermal shock, which share overlapping impedance characteristics. This confusion is most pronounced in XGBoost. RF partially mitigates this issue but still exhibits notable confusion between these two modes, reflecting the limitations of static feature thresholds.

Temporal sequence models progressively reduce inter-class confusion, but the nature and extent of improvement differ across architectures. In the LSTM classifier, confusion between fuel starvation and thermal shock is reduced compared to RF and XGBoost. A subset of fuel starvation samples is still misclassified as thermal shock, indicating that short- to mid-term temporal dependencies captured by the LSTM are insufficient to fully separate these two degradation modes. The TCN further improves class-wise separation, particularly for fuel starvation and short circuit. Misclassification of short circuit samples

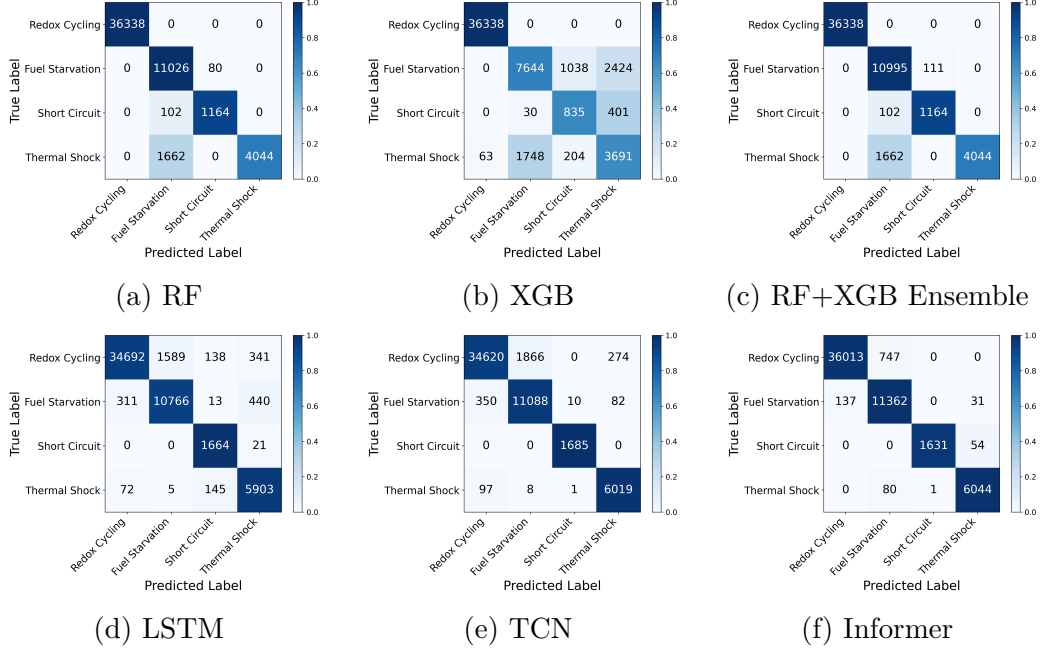


Figure 5.6: Confusion matrices for all evaluated classifiers. Cell values denote sample counts, while the color bar indicates normalized intensity in the range $[0, 1]$.

is almost entirely eliminated. Fuel starvation classification is also improved relative to the LSTM, with fewer samples leaking into the thermal shock class. The Informer demonstrates the strongest class-wise discrimination across all degradation modes. Confusion between fuel starvation and thermal shock is substantially reduced, with the majority of samples correctly classified along the diagonal. Redox cycling and short circuit are identified with near-perfect accuracy, and residual off-diagonal entries are minimal across all classes.

Computational Cost and Deployment Considerations

The computational costs are reported in Table 6.5. Unlike neural networks, RF and XGBoost do not have a well-defined notion of epochs or trainable parameters. RF training is non-iterative, while XGBoost proceeds via sequential boosting rounds. Model complexity for these models is approximated using the total number of tree nodes rather than exact parameter count. RF training corresponds to constructing the full forest per LODO fold, involving recursive tree growth and split searches that are largely CPU-based and do

Table 5.10: Computational cost comparison of degradation mode classifiers. Training time is averaged across LODO folds. All experiments were conducted on an NVIDIA T4 GPU.

Model	Training time per epoch (s)	Inference time per sample (ms)	Trainable parameters
RF	6.091*	0.292	7,118 [†]
XGB	1.128	0.019	4,834
Ensemble	16.742	0.402	10,967
LSTM	4.507	0.010	22,596
TCN	8.575	0.023	42,884
Informer	81.176	0.187	67,588

* For RF, an “epoch” denotes one complete training pass per LODO fold.

[†] For RF, the reported parameter count corresponds to the total number of tree nodes (internal nodes and leaves).

not fully benefit from GPU acceleration. During inference, RF must traverse all trees for each sample, leading to branch-heavy computations and higher latency (0.292 ms). In contrast, XGB is computationally more efficient due to its optimized boosting implementation and smaller tree structures, resulting in faster training and lower inference time (0.019 ms). The RF+XGB Ensemble incurs the highest cost among the tree-based approaches since it combines both training procedures and evaluates both models at inference.

Neural sequence models (LSTM, TCN, and Informer) rely on dense tensor operations that are efficiently parallelized on GPUs. As a result, the LSTM achieves the lowest inference latency despite having more parameters than the tree-based models. The TCN requires additional computation due to stacked dilated convolution blocks, increasing both training and inference time relative to LSTM. The Informer exhibits the highest computational cost overall because its attention-based encoder is more expensive for long input windows; however, its inference time remains within a practical range (0.187 ms), illustrating the trade-off between improved accuracy and increased computational complexity.

5.5.7 Fold-wise Classification Uncertainty Analysis

To further analyze classification robustness under LOFO validation, fold-level F1 scores were examined for all classifiers. Since each held-out file corresponds to one degradation mechanism, the fold-level F1 score indicates how reliably each model identifies the active degradation mode for that specific experimental trajectory. Therefore, this analysis helps clarify which held-out datasets contribute to the false-negative behavior observed in the confusion matrices. Table 5.11 reports the F1 scores for each held-out file. The shallow models achieve high performance on several trajectories but show catastrophic failures on specific folds. For example, RF and RF+XGB fail on SC-1 and TS-3, while XGB fails on FS-3, SC-1, and TS-3. In contrast, the temporal models provide more stable fold-wise performance. Informer achieves consistently high F1 scores across all held-out files, with all values remaining above 0.96.

To summarize the variability across folds, Table 5.12 reports the mean, standard deviation (SD), and 95% confidence interval (CI) of the fold-level F1 scores. The 95% CI was computed using the t -distribution across the 12 LOFO folds. The shallow classifiers show large fold-to-fold variability, with SD values of 0.39 for RF and RF+XGB and 0.44 for XGB, mainly due to their failures on specific held-out trajectories. In contrast, the deep temporal models show substantially lower variability. Informer achieves the highest mean F1 score of 0.99 and the narrowest confidence interval of $[0.98, 1.00]$, indicating more stable classification performance across degradation trajectories.

Paired Wilcoxon signed-rank tests were also performed to compare Informer against each baseline under identical LOFO folds. The mean fold-wise F1 difference, computed as Informer minus the corresponding baseline, and the associated one-sided Wilcoxon p -value are reported in Table 5.12. The mean difference reflects the average magnitude of improvement, while the Wilcoxon test assesses whether the paired fold-wise improvements are systematically greater than zero. Informer shows large mean F1 improvements over the shallow classifiers, especially XGB, due to the folds where these models fail. However, the Wilcoxon p -values are not always significant for RF and RF+XGB

because these models also achieve perfect or near-perfect F1 scores on several folds, resulting in many tied comparisons. In contrast, the improvement over LSTM is smaller in magnitude but more consistent across folds, leading to a lower Wilcoxon p -value. Overall, the fold-wise analysis shows that Informer provides the highest average F1 score and the most stable performance across degradation trajectories.

Since the Wilcoxon signed-rank test is based on the ranks of the paired fold-wise differences rather than their raw magnitudes, the mean F1 difference is also reported to quantify the actual average improvement of the Informer over each baseline. This is particularly informative when the number of LOFO folds is limited and several comparisons contain tied or small differences.

Table 5.11: Fold-level F1 scores obtained from LOFO validation.

Test File	RF	XGB	RF+XGB	LSTM	TCN	Informer
RC1	1.00	1.00	1.00	0.98	1.00	1.00
RC2	1.00	1.00	1.00	0.95	0.94	0.98
RC3	1.00	1.00	1.00	0.99	0.99	1.00
FS1	1.00	1.00	1.00	1.00	1.00	1.00
FS2	1.00	1.00	1.00	0.96	0.99	0.98
FS3	0.99	0.00	0.98	0.93	0.95	1.00
SC1	0.00	0.00	0.00	1.00	1.00	1.00
SC2	1.00	0.60	1.00	0.98	1.00	0.96
SC3	1.00	1.00	1.00	1.00	1.00	1.00
TS1	1.00	1.00	1.00	1.00	1.00	1.00
TS2	1.00	0.93	1.00	0.96	0.98	0.98
TS3	0.00	0.00	0.00	0.98	1.00	1.00

Table 5.12: Fold-level F1 summary across the 12 LOFO folds. Mean, standard deviation (SD), and 95% confidence interval (CI) are reported for each model. The last two columns show the mean fold-wise F1 difference and paired Wilcoxon signed-rank p -value for each baseline relative to Informer.

Model	Mean F1	SD	95% CI	Mean Difference (vs Informer)	p -value (vs Informer)
RF	0.83	0.39	[0.59, 1.00]	0.16	0.51
XGB	0.71	0.44	[0.43, 0.99]	0.28	0.04
RF+XGB	0.83	0.39	[0.59, 1.00]	0.16	0.43
LSTM	0.98	0.02	[0.96, 0.99]	0.01	0.03
TCN	0.99	0.02	[0.97, 1.00]	0.00	0.34
Informer	0.99	0.01	[0.99, 1.00]	–	–

5.5.8 Scope of Applicability and Deployment Considerations

It should be noted that the proposed framework was developed and validated using a population of nominally identical tubular oxygen-ion conducting SOFCs fabricated using the same materials and operated under similar conditions. Consequently, the reported results demonstrate classification performance within a single cell population and do not establish universal applicability across different SOFC designs. Variations in cell geometry, manufacturer, electrolyte chemistry, operating temperature, and fuel composition may alter degradation behavior, impedance characteristics, DRT signatures, and the absolute values of the extracted diagnostic features. Therefore, although the same methodology can be utilized for other SOFC types and geometries, the resulting feature distributions and model performance are not expected to be identical without additional validation and retraining. Transfer learning strategies provide a potential pathway for extending the framework to other SOFC configurations.

To assess the practicality of online deployment, the computational requirements of the proposed diagnostic framework are summarized in Table 5.13, including model inference time ($t_{\text{inference}}$), DRT feature extraction and feature engineering time (t_{feature}), EIS acquisition time (t_{EIS}), and the interval between consecutive EIS measurements (t_{interval}). A complete EIS measurement requires approximately 2 min, depending on the selected frequency range and measurement settings. The interval between successive EIS acquisitions varies according to the degradation experiment and cycling protocol, as impedance measurements are collected during periodic returns to open-circuit conditions. The DRT feature extraction process requires approximately 278 ms on a Mac 8-Core CPU, while model inference is completed in less than 0.5 ms for all evaluated classifiers. For the shallow models, the engineered feature calculation adds only 3.36 ms to the feature-processing stage, resulting in a total feature-processing time of 278+3.36 ms. Consequently, both feature extraction and classification introduce negligible computational overhead compared with the

EIS acquisition process. These results indicate that the primary constraint for real-time implementation is the acquisition of impedance measurements rather than the execution of the diagnostic algorithm itself, demonstrating the feasibility of integrating the proposed framework into online SOFC health-monitoring systems.

Table 5.13: Online computational profile of the diagnostic framework. Model inference time ($t_{\text{inference}}$) is measured on an NVIDIA Tesla T4 GPU. DRT feature extraction is performed using a MATLAB implementation on a Mac 8-Core CPU and requires 278 ms per EIS curve. For shallow models, engineered feature calculation adds 3.36 ms to the feature-processing time.

Model	$t_{\text{inference}}$ (ms)	t_{feature} (ms)	t_{interval} (min)	t_{EIS} (min)
RF	0.292	278+3.36	10-45	2
XG-Boost	0.019	278+3.36	10-45	2
Ensemble	0.402	278+3.36	10-45	2
LSTM	0.010	278	10-45	2
TCN	0.023	278	10-45	2
Informer	0.187	278	10-45	2

5.6 Summary of Chapter

This chapter presented an experimentally grounded SOFC degradation mode classification framework designed to support mechanism-aware health monitoring and predictive maintenance. A large-scale dataset consisting of 822,463 samples was generated from controlled degradation experiments performed on twelve nominally identical tubular SOFCs under four degradation mechanisms: redox cycling, fuel starvation, short circuit, and thermal shock. EIS measurements were processed using DRT analysis to extract frequency-domain peak features and ohmic resistance, which were then resampled, standardized, and segmented into labeled multivariate temporal windows to form a mechanism-aware classification dataset.

A systematic comparison across six classifiers was performed using a LODO protocol. Among shallow models, RF achieved strong balanced performance (Macro-F1 = 0.92), whereas XGB exhibited lower class-balanced performance

(Macro-F1 = 0.71). Combining RF and XGB in an ensemble restored performance to the RF level. Temporal deep learning models provided further improvements, with LSTM matching RF (Macro-F1 = 0.92) and TCN yielding higher balanced performance (Macro-F1 = 0.96). The Informer achieved the best overall results (Macro-F1 = 0.98), demonstrating the strongest class-wise separation across mechanistically similar degradation modes. Computational cost analysis highlighted a clear performance-efficiency trade-off: tree-based models were efficient and well suited for lightweight deployment (with inference latency ranging from 0.019 to 0.402 ms), LSTM provided the lowest inference latency among deep models, and Informer required the highest training cost while maintaining practical inference time (0.187 ms) for diagnostic use. Overall, these results demonstrate that experimentally grounded impedance diagnostics combined with temporal learning enable robust SOFC degradation mode identification, with attention-based modeling providing the highest class-balanced performance at the expense of increased computational complexity. The proposed dataset and benchmarking study establish a foundation for mechanism-aware SOFC health management, enabling future integration with prognostics and supervisory control strategies.

Chapter 6

Degradation Mechanism-aware Prognostics¹

6.1 Introduction

The diagnostic and prognostic frameworks developed in the previous chapters provide complementary capabilities for SOFC health monitoring. While Chapter 4 introduced an uncertainty-aware framework for long-horizon SOH and RUL forecasting, and Chapter 5 established an impedance-based diagnostic approach for identifying degradation mechanisms, these two tasks have been treated independently. However, in practical applications, degradation behavior and lifetime evolution are strongly dependent on the underlying degradation mode. As a result, mechanism-agnostic prognostic models often exhibit limited generalization and reduced prediction accuracy when operating across diverse degradation scenarios.

To address this limitation, this chapter focuses on integrating diagnostics and prognostics into a unified framework for mechanism-aware RUL prediction. In contrast to the previous chapter, where the full SOH trajectory is forecasted to estimate RUL, the approach in this chapter directly predicts RUL as a single target variable. This formulation simplifies the prognostic task while enabling the model to explicitly leverage degradation mode information for improved prediction accuracy.

An asymmetric multi-task learning (MTL) framework is proposed to jointly

¹This chapter is partially based on the following publication: [5].

perform degradation mode classification and RUL estimation within a single model. The approach is based on an attention-driven Informer architecture, which captures long-range temporal dependencies in degradation data. To effectively incorporate diagnostic information into the prognostic task, different fusion strategies are investigated. In particular, the proposed architecture transfers latent representations learned by the classifier to the RUL prediction branch through late-stage feature fusion, allowing the model to exploit the relationship between degradation mechanisms and lifetime evolution.

The proposed framework is validated using experimentally generated degradation datasets under multiple degradation modes, including redox cycling, fuel starvation, short circuit, and thermal shock. The results demonstrate that incorporating mechanism awareness significantly improves RUL prediction accuracy compared to conventional single-task models. In addition, the integrated framework achieves improved computational efficiency relative to sequential execution of separate diagnostic and prognostic models.

Overall, this chapter establishes a unified, mechanism-aware prognostic framework that combines degradation mode identification and lifetime prediction. By leveraging shared representations across tasks, the proposed approach enhances both accuracy and robustness, providing a practical and scalable solution for real-time SOFC health monitoring and predictive maintenance.

6.2 Data Processing and Class Labeling

To generate informative inputs for degradation mode classification and RUL estimation, electrochemical measurements were processed to extract physically meaningful features from impedance data. In parallel, a SOH signal was derived to quantify the temporal evolution of SOFC performance degradation.

SOH Estimation For degradation experiments performed under different operating modes, the SOH is quantified using the evolution of electrochemical power output, enabling a unified definition applicable to both constant-voltage and constant-current conditions. Specifically, the redox cycling experiments

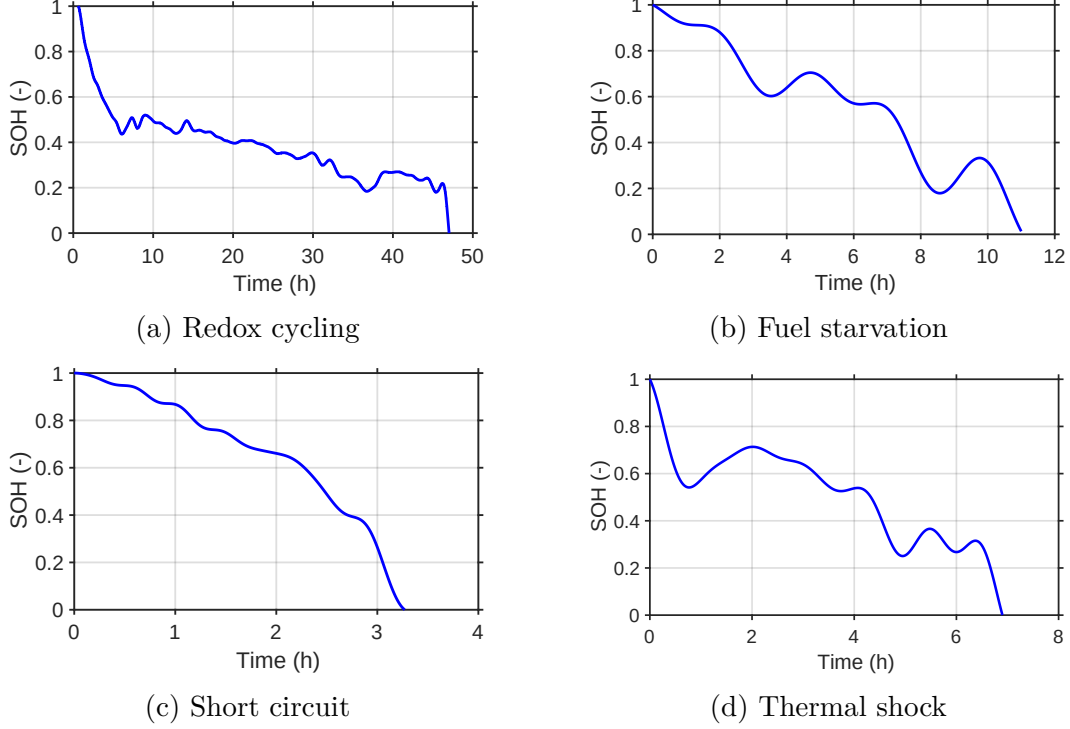


Figure 6.1: Representative SOL trajectories from BOL to EOL under the four experimentally induced degradation mechanisms.

were conducted under constant-voltage operation ($V = 0.7$ V), whereas the fuel starvation, short circuit, and thermal shock experiments were performed under constant-current operation. To ensure consistency across all degradation mechanisms, SOH is therefore defined based on power:

$$\text{SOH} = \frac{P - P_E}{P_B - P_E}, \quad (6.1)$$

where P denotes the instantaneous power output, P_B is the power at the beginning of life (BOL), and P_E corresponds to the power at end of life (EOL). In this study, the EOL condition is defined as a 10% reduction in the initial power output, such that $P_E = 0.9P_B$, consistent with commonly adopted SOFC degradation criteria. This formulation yields a normalized SOH trajectory bounded between zero and one, providing a compact and consistent indicator of performance degradation across all experimental conditions. Representative SOH trajectories from BOL to EOL for the four investigated degradation mechanisms are illustrated in Fig. 6.1.

Electrochemical Impedance and DRT Feature Extraction EIS measurements were used to extract diagnostic features characterizing internal electrochemical processes. As described in previous chapters, the measured impedance spectra were processed using the DRT method to obtain physically interpretable features associated with distinct electrochemical mechanisms. In this chapter, the same DRT-based feature extraction procedure is adopted. The resulting DRT peak features, together with the ohmic resistance, are used as compact and mechanism-informed representations of degradation behavior and serve as inputs to the proposed multi-task learning framework.

For redox cycling, pronounced growth is observed in the high- and mid-frequency DRT peaks (Peak 1 and Peak 2). Peak 1 is associated with oxygen ion transport and electrolyte-electrode interfacial polarization, while Peak 2 corresponds to anode activation polarization. These trends are consistent with microstructural degradation in the anode functional layer, including nickel particle coarsening and pore structure evolution, which reduce the effective Ni-YSZ triple-phase boundary (TPB) area and increase polarization resistance. In fuel starvation experiments, degradation manifests primarily as an increase in Peak 2 along with a concurrent rise in Ohmic resistance. The amplification of Peak 2 indicates worsening anode activation polarization, whereas the increase in Ohmic resistance suggests electrolyte degradation or loss of electrical continuity. Physically, these changes are attributed to anode delamination, microstructural depletion, and reduced electronic and ionic conductivity caused by sustained hydrogen deficiency and local re-oxidation effects. For short circuit degradation, the dominant signature is a monotonic increase in ohmic resistance. This behavior points to damage in the electrolyte or current collection pathways rather than kinetic limitations. Localized overheating during repeated shorting events can induce micro-cracks, contact degradation, or interlayer delamination, leading to increased resistive losses. Finally, thermal shock degradation is characterized by a simultaneous increase in Peak 3 and the Ohmic resistance. Peak 3 is associated with cathode activation polarization, indicating degradation of cathodic electrochemical processes. These

trends are consistent with severe and rapid temperature gradients inducing thermal stresses, which promote crack formation in the YSZ electrolyte and cathode layers. Notably, the cathode constitutes the outermost layer of the tubular cell and is directly exposed to ambient air during the furnace removal step. This asymmetric cooling environment leads to accelerated heat extraction on the cathode side, making it susceptible to delamination and contact loss.

Labeling and RUL Target Definition Each degradation experiment yields a multivariate time series comprising six input channels: four DRT peak amplitudes, the ohmic resistance (R_{ohm}), and the SOH signal. To ensure consistency across experiments, all channels were resampled to a uniform temporal grid and standardized to remove scale disparities. The data were then segmented into overlapping windows using a look-back horizon of 500 samples. Each window was assigned a categorical label corresponding to the active degradation mechanism. The four degradation modes of redox cycling, fuel starvation, short circuit, and thermal shock were encoded using one-hot vectors to serve as targets for the classification task, with redox cycling represented as $[1, 0, 0, 0]$, fuel starvation as $[0, 1, 0, 0]$, short circuit as $[0, 0, 1, 0]$, and thermal shock as $[0, 0, 0, 1]$.

In parallel, a continuous RUL target was computed for each window to support regression-based prognostics. The RUL was defined as the remaining time from the end of the window to the experimentally determined EOL:

$$\text{RUL}(t) = t_{\text{EOL}} - t, \quad (6.2)$$

where t denotes the time stamp of the final sample in the window and t_{EOL} corresponds to the time at which the 10% peak power loss criterion was reached. The resulting dataset therefore supports joint learning of degradation mode classification and RUL estimation, providing the foundation for the multi-task learning framework developed in this study.

6.3 Mechanism-aware Remaining Useful Life Prediction

In the following subsections, the proposed framework is described in a step-by-step manner. First, the single-task Informer-based RUL estimator is introduced as the baseline model. Next, a standalone degradation mode classifier is presented to identify the active degradation mechanism. The third subsection explains the early and late fusion strategies used to inject degradation mode information into the RUL estimator and to empirically determine the most effective fusion location. The baseline pipeline and the two fusion variants are summarized in Figure 6.2. Finally, the asymmetric multi-task learning (MTL) architecture is described, in which the classifier and the RUL estimator are trained jointly using a combined loss and the classifier latent representation is transferred to the RUL branch via late stage feature fusion, as illustrated in Figure 6.3.

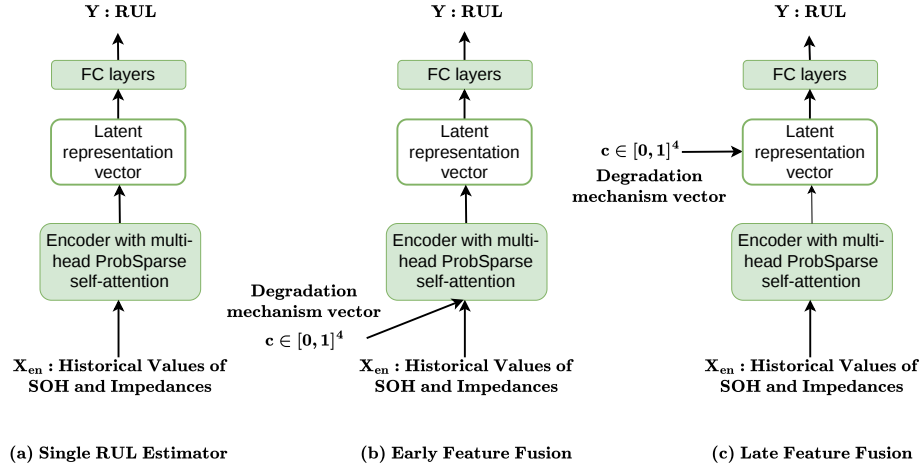


Figure 6.2: Overview of the baseline RUL estimation pipeline and the investigated fusion strategies for incorporating degradation mode information.

6.3.1 Single-task RUL Estimator

The single-task RUL estimator is constructed using the Informer architecture, an efficient attention-based sequence model designed for long-horizon forecasting. Its primary advantage lies in its ability to capture long-range temporal

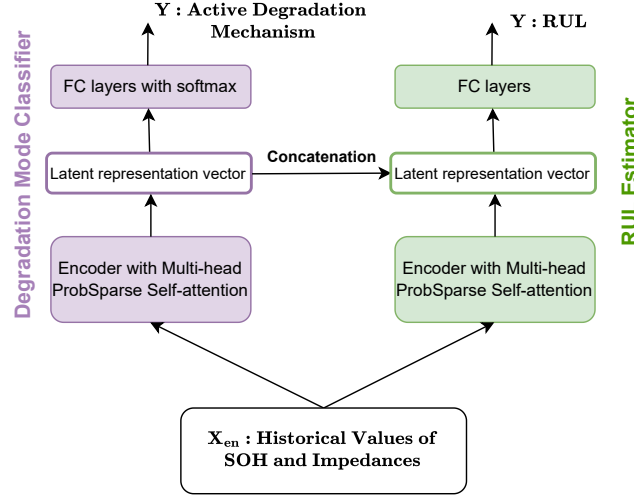


Figure 6.3: Proposed asymmetric MTL architecture for degradation mechanism-aware prognostics.

dependencies while maintaining computational manageability for extended input sequences. For each SOFC degradation trajectory, the model receives a multivariate look-back window composed of the historical values of SOH, ohmic resistance, and DRT-derived impedance features. The output is a single scalar representing the remaining distance from the window endpoint to EOL.

The Informer embeds the multivariate input sequence into a high-dimensional latent space, supplemented by temporal positional encodings to preserve ordering information. The encoder is composed of multiple stacked layers combining *ProbSparse self-attention* with feed-forward transformations, yielding a compact representation of the underlying degradation dynamics. Unlike the full Informer architecture used for sequence-to-sequence forecasting, no decoder is required in this application because the objective is to regress a single RUL value. As illustrated in Figure 6.4, the input look-back window $\mathbf{X}_{\text{en}} \in \mathbb{R}^{L \times F}$ is first projected into an embedded representation:

$$\mathbf{Z}_0 = \mathbf{X}_{\text{en}} \mathbf{W}_e + \mathbf{P}, \quad (6.3)$$

where $\mathbf{W}_e$ denotes the learnable embedding matrix and $\mathbf{P}$ represents positional encodings. The encoder is composed of $K = 2$ stacked ProbSparse attention

blocks. The embedded sequence is transformed as:

$$\mathbf{Z}_1 = \text{AttnBlock}_1(\mathbf{Z}_0), \quad \mathbf{H} = \mathbf{Z}_2 = \text{AttnBlock}_2(\mathbf{Z}_1). \quad (6.4)$$

Following the encoder, temporal aggregation is performed via pooling across the sequence dimension to obtain a fixed-length latent representation:

$$\mathbf{h} = \text{Pool}(\mathbf{H}). \quad (6.5)$$

Finally, the encoder output is passed through a set of fully connected (FC) layers to produce the one-shot RUL prediction:

$$\hat{y}_{\text{RUL}} = f_{\text{FC}}(\mathbf{h}). \quad (6.6)$$

This single-task estimator serves as a baseline for evaluating the benefits of the proposed multi-task learning framework and provides the regression component later integrated into the asymmetric MTL architecture.

6.3.2 Degradation Mode Classifier

To estimate the active degradation mechanism associated with each operating trajectory, a degradation mode classifier is developed using the same encoder-only Informer-inspired architecture as the RUL estimator. The key difference is that the classifier outputs a discrete degradation mode label rather than a continuous RUL value. This design enables learning a compact, discriminative representation of degradation signatures from the same multivariate input sequence used in RUL estimation. Given the input look-back window $\mathbf{X}_{\text{en}} \in \mathbb{R}^{L \times F}$, the embedding and positional encoding stage is identical to the RUL estimator:

$$\mathbf{Z}_0 = \mathbf{X}_{\text{en}} \mathbf{W}_e + \mathbf{P}, \quad (6.7)$$

where $\mathbf{W}_e$ is the learnable embedding matrix and $\mathbf{P}$ denotes positional encodings. The encoded sequence is then processed by two stacked attention blocks:

$$\mathbf{Z}_1 = \text{AttnBlock}_1(\mathbf{Z}_0), \quad \mathbf{H} = \mathbf{Z}_2 = \text{AttnBlock}_2(\mathbf{Z}_1). \quad (6.8)$$

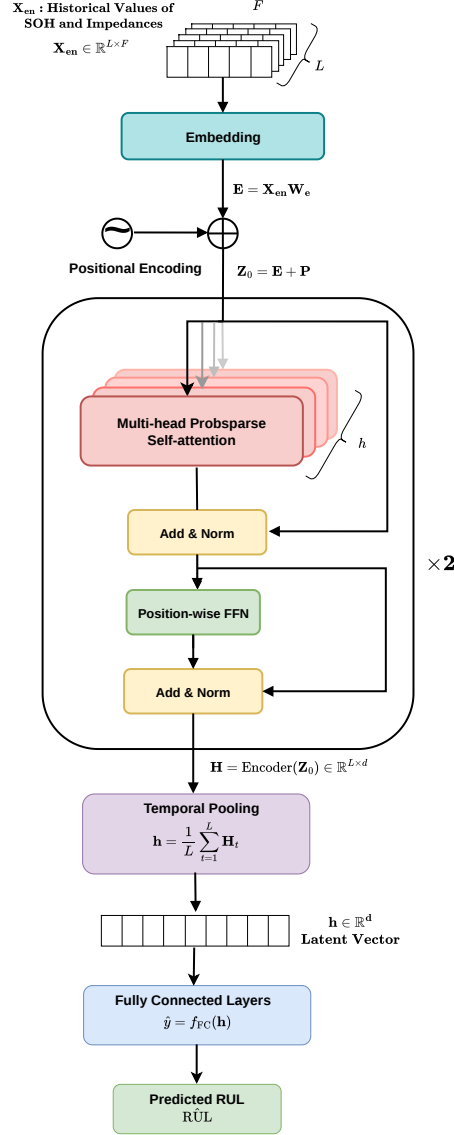


Figure 6.4: Architecture of the baseline single-task RUL estimator based on an Informer-style encoder. The multivariate input sequence, consisting of SOH and impedance-derived features, is embedded and augmented with positional encoding before being processed by stacked Probsparse self-attention blocks. Temporal pooling is applied to obtain a fixed-length latent representation, which is mapped to a scalar RUL prediction through fully connected layers.

A fixed-length representation is obtained by temporal pooling:

$$\mathbf{h}_{\text{CLS}} = \text{Pool}(\mathbf{H}), \quad \mathbf{h}_{\text{CLS}} \in \mathbb{R}^d, \quad (6.9)$$

where d is the encoder hidden dimension. Finally, a fully connected classification head maps the latent representation to class logits:

$$\hat{\mathbf{p}} = \text{softmax}(f_{\text{CLS}}(\mathbf{h}_{\text{CLS}})), \quad \hat{\mathbf{p}} \in \mathbb{R}^C, \quad (6.10)$$

where C is the number of degradation modes and $\hat{\mathbf{p}}$ represents the predicted class probability vector. The predicted degradation mode is then obtained as:

$$\hat{c} = \arg \max_{j \in \{1, \dots, C\}} \hat{\mathbf{p}}_j. \quad (6.11)$$

This classifier provides mechanism-awareness that can be injected into the RUL estimator through the feature fusion strategies evaluated in the next subsection.

6.3.3 Feature Fusion

To systematically evaluate where degradation mode information should be injected into the RUL prediction pipeline, two fusion strategies are investigated: *early fusion* and *late fusion*. The investigated early and late fusion strategies are illustrated in Figure 6.5. In this ablation study, the fusion input is the *ground-truth degradation mode label* to isolate the effect of injection location and remove potential classifier uncertainty.

Let the original encoder input be:

$$\mathbf{X}_{\text{en}} \in \mathbb{R}^{L \times F}, \quad (6.12)$$

where L is the look-back window length and F is the number of input features (SOH, ohmic resistance, and DRT-derived peaks). The degradation mode label is represented as a one-hot vector:

$$\mathbf{c} \in \{0, 1\}^C, \quad (6.13)$$

where C is the number of degradation modes.

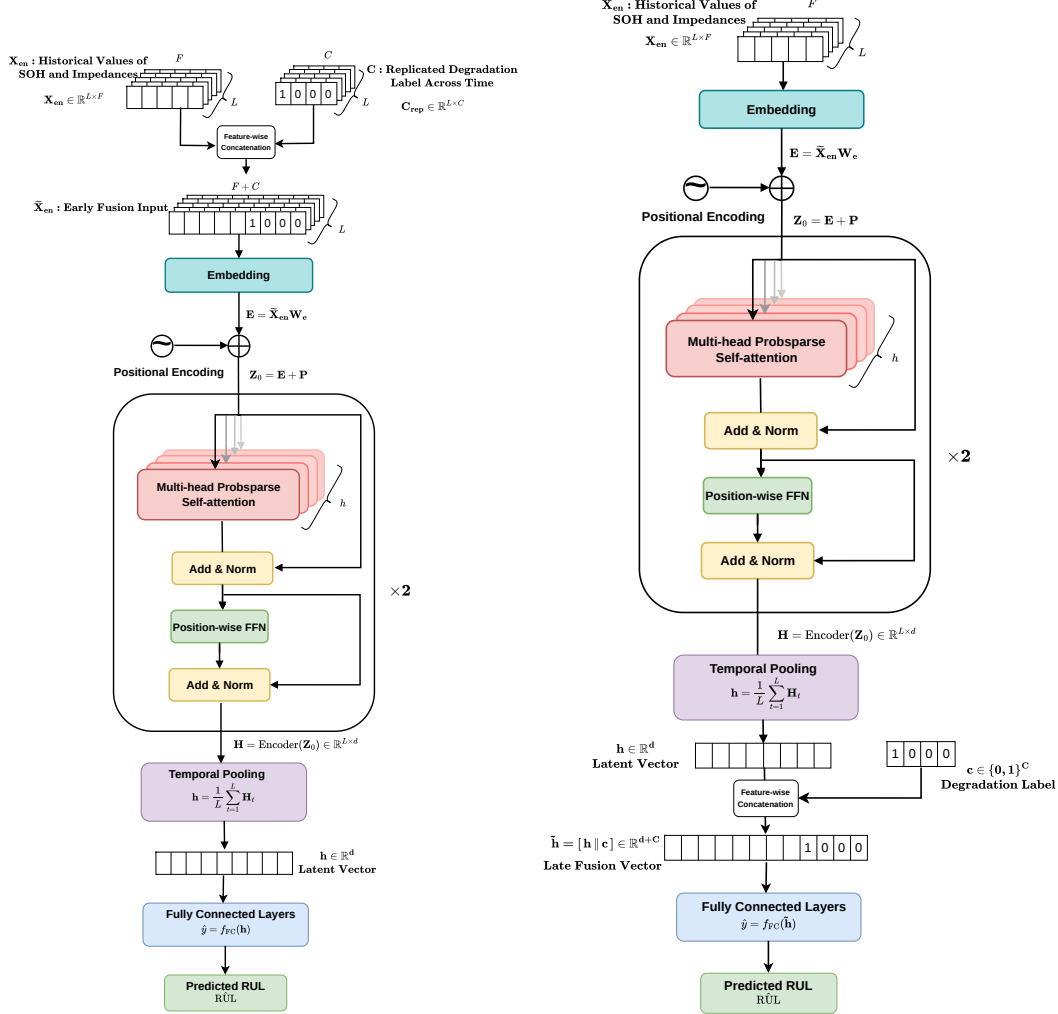


Figure 6.5: Comparison of fusion strategies for incorporating degradation mechanism information into the RUL estimator.

Early Fusion In early fusion, the one-hot label is replicated across the temporal dimension and concatenated with the input features at each time step. The replicated label tensor is:

$$\mathbf{C}_{\text{rep}} \in \{0, 1\}^{L \times C}, \quad (6.14)$$

The fused input becomes:

$$\tilde{\mathbf{X}}_{\text{en}} = [\mathbf{X}_{\text{en}} \parallel \mathbf{C}_{\text{rep}}] \in \mathbb{R}^{L \times (F+C)}, \quad (6.15)$$

where $\parallel$ denotes feature-wise concatenation. The RUL encoder then operates on $\tilde{\mathbf{X}}_{\text{en}}$:

$$\tilde{\mathbf{Z}}_0 = \tilde{\mathbf{X}}_{\text{en}} \mathbf{W}_e + \mathbf{P}, \quad \tilde{\mathbf{Z}}_2 = \text{Encoder}(\tilde{\mathbf{Z}}_0), \quad (6.16)$$

followed by pooling and regression:

$$\mathbf{h}_{\text{early}} = \text{Pool}(\tilde{\mathbf{Z}}_2) \in \mathbb{R}^d, \quad \hat{y}_{\text{RUL}} = f_{\text{FC}}(\mathbf{h}_{\text{early}}). \quad (6.17)$$

Late Fusion In late fusion, the encoder first extracts a latent representation without explicit access to degradation mode identity. The baseline latent vector is:

$$\mathbf{h} = \text{Pool}(\text{Encoder}(\mathbf{Z}_0)) \in \mathbb{R}^d. \quad (6.18)$$

The one-hot label is concatenated after pooling:

$$\tilde{\mathbf{h}} = [\mathbf{h} \parallel \mathbf{c}] \in \mathbb{R}^{d+C}, \quad (6.19)$$

and the fused vector is mapped to the RUL prediction:

$$\hat{y}_{\text{RUL}} = f_{\text{FC}}(\tilde{\mathbf{h}}). \quad (6.20)$$

Early fusion injects mechanism identity into the sequence representation learning stage, which can bias attention patterns and feature extraction toward class-specific shortcuts. In contrast, late fusion preserves a mechanism-agnostic encoder representation and uses the degradation mode label primarily for conditional calibration in the regression head.

Latent Space Separability To quantify how explicitly incorporating degradation mode information affects the structure of the learned latent space, we evaluate the *silhouette score* of encoder embeddings under different fusion strategies. The silhouette score measures the degree of intra-class compactness and inter-class separation, providing a cluster-quality metric that is independent of the downstream regression loss.

Let $\mathbf{h}_i \in \mathbb{R}^d$ denote the pooled latent embedding of the i -th sample, and let A be the set of samples belonging to the same degradation mode (cluster) as $\mathbf{h}_i$. The average intra-cluster distance is defined as

$$a(i) = \frac{1}{|A| - 1} \sum_{\substack{j \in A \\ j \neq i}} d(\mathbf{h}_i, \mathbf{h}_j), \quad (6.21)$$

where $d(\cdot, \cdot)$ denotes the Euclidean distance in the latent space.

For all other clusters $C \neq A$, the mean distance between $\mathbf{h}_i$ and samples in cluster C is computed, and the minimum of these values defines the nearest inter-cluster distance:

$$b(i) = \min_{C \neq A} \left(\frac{1}{|C|} \sum_{j \in C} d(\mathbf{h}_i, \mathbf{h}_j) \right). \quad (6.22)$$

The silhouette score for sample i is then given by

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \quad -1 \leq s(i) \leq 1. \quad (6.23)$$

A value close to 1 indicates that the sample is well matched to its own cluster and far from neighboring clusters, while values near 0 indicate overlapping clusters and negative values suggest potential misclassification.

Finally, the fold-level silhouette score is obtained by averaging over all N samples in the fold:

$$S = \frac{1}{N} \sum_{i=1}^N s(i). \quad (6.24)$$

While the silhouette score provides a global, normalized measure of latent space separability, it does not reveal how specific degradation modes are positioned relative to one another. To enable pairwise analysis and physical

interpretation of inter-mechanism relationships, class-wise centroid distances in the latent embedding space is computed.

For each degradation mode $k \in \{1, \dots, K\}$, the centroid of its latent representations is defined as

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{i \in \mathcal{C}_k} \mathbf{h}_i, \quad (6.25)$$

where $\mathcal{C}_k$ denotes the set of samples belonging to class k and $N_k = |\mathcal{C}_k|$ is the number of samples in that class. The pairwise distance between two degradation modes k and ℓ is then computed as the Euclidean distance between their centroids:

$$D_{k\ell} = d(\boldsymbol{\mu}_k, \boldsymbol{\mu}_\ell) = \|\boldsymbol{\mu}_k - \boldsymbol{\mu}_\ell\|_2. \quad (6.26)$$

6.3.4 Asymmetric MTL

Following the fusion study, the proposed asymmetric multi-task learning (MTL) framework is constructed to jointly perform RUL estimation and degradation mode classification in a single end-to-end model. Unlike the fusion ablations that rely on ground-truth labels, the MTL framework replaces the one-hot class label with a learned classifier embedding, enabling mechanism-aware RUL prediction during both training and inference.

As shown in Figure 6.6, the MTL framework contains two separate encoder branches (soft parameter sharing): one dedicated to RUL estimation and one dedicated to classification. Both branches receive the same input window $\mathbf{X}_{\text{en}} \in \mathbb{R}^{L \times F}$, but learn task-specific representations. The RUL encoder produces:

$$\mathbf{Z}_0^{\text{RUL}} = \mathbf{X}_{\text{en}} \mathbf{W}_e^{\text{RUL}} + \mathbf{P}^{\text{RUL}}, \quad \mathbf{H}_{\text{RUL}} = \text{Encoder}_{\text{RUL}}(\mathbf{Z}_0^{\text{RUL}}) \in \mathbb{R}^{L \times d}, \quad (6.27)$$

and its pooled latent vector is:

$$\mathbf{h}_{\text{RUL}} = \text{Pool}(\mathbf{H}_{\text{RUL}}) \in \mathbb{R}^d. \quad (6.28)$$

Similarly, the classifier encoder produces:

$$\mathbf{Z}_0^{\text{CLS}} = \mathbf{X}_{\text{en}} \mathbf{W}_e^{\text{CLS}} + \mathbf{P}^{\text{CLS}}, \quad \mathbf{H}_{\text{CLS}} = \text{Encoder}_{\text{CLS}}(\mathbf{Z}_0^{\text{CLS}}) \in \mathbb{R}^{L \times d}, \quad (6.29)$$

and the classifier latent vector is:

$$\mathbf{h}_{\text{CLS}} = \text{Pool}(\mathbf{H}_{\text{CLS}}) \in \mathbb{R}^d. \quad (6.30)$$

Motivated by the superior performance of late fusion, mechanism-awareness is injected at the latent representation level. Specifically, the RUL head receives both its own latent vector and the classifier latent vector through feature-wise concatenation:

$$\mathbf{h}_{\text{MTL}} = [\mathbf{h}_{\text{RUL}} \parallel \mathbf{h}_{\text{CLS}}] \in \mathbb{R}^{2d}. \quad (6.31)$$

The RUL prediction is then obtained by:

$$\hat{y}_{\text{RUL}} = f_{\text{RUL}}(\mathbf{h}_{\text{MTL}}). \quad (6.32)$$

Meanwhile, the classifier prediction relies only on its own latent vector:

$$\hat{\mathbf{p}} = \text{softmax}(f_{\text{CLS}}(\mathbf{h}_{\text{CLS}})). \quad (6.33)$$

The asymmetric MTL framework is trained end-to-end using a weighted sum of regression and classification losses:

$$\mathcal{L}_{\text{MTL}} = \lambda_{\text{RUL}} \mathcal{L}_{\text{RUL}} + \lambda_{\text{CLS}} \mathcal{L}_{\text{CLS}}, \quad (6.34)$$

where λ_{RUL} and λ_{CLS} control the trade-off between tasks. The regression loss is defined as mean squared error (MSE):

$$\mathcal{L}_{\text{RUL}} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (6.35)$$

and the classification loss is cross-entropy:

$$\mathcal{L}_{\text{CLS}} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C \mathbf{1}(c_i = j) \log(\hat{\mathbf{p}}_{i,j}). \quad (6.36)$$

This design yields an asymmetric information flow: the classifier learns discriminative mechanism embeddings that regularize and enrich the RUL estimation pathway, while the classification task remains independent of the RUL latent space. The resulting framework improves RUL prediction accuracy while remaining deployable in real operation, since it does not require access to ground-truth degradation labels at inference time.

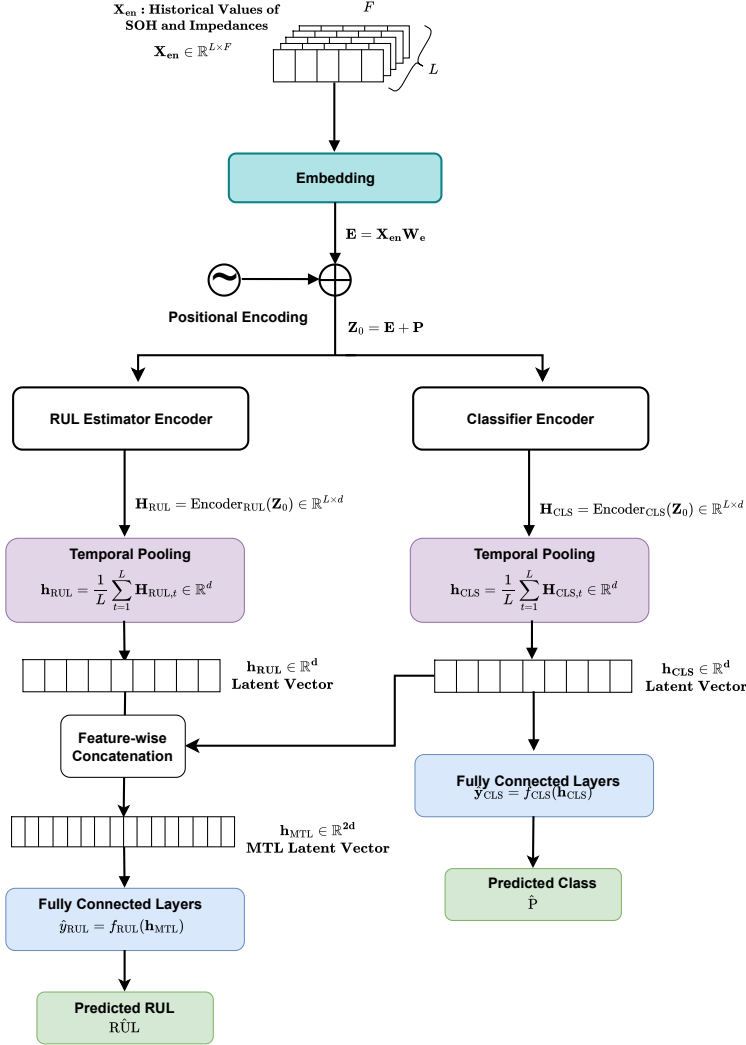


Figure 6.6: Proposed asymmetric MTL framework for joint degradation mode classification and RUL estimation. Two task-specific encoders process the same input sequence: one optimized for RUL estimation and the other for degradation mode classification. The classifier latent embedding is transferred to the RUL branch via late-stage feature concatenation.

6.4 Results and Discussion

6.4.1 Evaluation Strategy and Metrics

The evaluation strategy and classification metrics used in this chapter follow the same framework introduced in Chapter 5. In particular, the Leave-One-Dataset-Out (LODO) cross-validation scheme is adopted to assess model generalization across independent degradation trajectories, ensuring strict separation between training and testing data and preventing temporal data leakage. For degradation mode classification, the same performance metrics, Precision, Recall, F1-score, and their Macro-averaged variants, are used to evaluate diagnostic accuracy across multiple degradation modes. In addition, to assess the performance of the RUL estimation task, two regression-based metrics, MAE and RMSE, are introduced. These metrics provide complementary measures of prediction accuracy, capturing both average deviation and sensitivity to larger errors.

6.4.2 Results and Discussion

Early vs Late Degradation Label Fusion Table 6.1 compares the baseline single-task RUL estimator against early and late fusion strategies under LODO cross-validation. Both fusion approaches outperform the baseline across all degradation cases, demonstrating that incorporating explicit degradation mode information improves RUL prediction accuracy. In particular, early fusion reduces the average RMSE from 0.49 h to 0.39 h and the average MAE from 0.41 h to 0.32 h, whereas late fusion achieves the strongest overall performance with an average RMSE of 0.31 h and MAE of 0.26 h. The corresponding averaged improvement percentages are summarized in Table 6.4, where early fusion yields approximately $21.49\% \pm 2.96$ RMSE improvement and $21.62\% \pm 2.93$ MAE improvement, while late fusion achieves larger gains of $37.34\% \pm 4.19$ (RMSE) and $37.06\% \pm 4.86$ (MAE).

The difference between early and late fusion can be explained by how the fault label interacts with representation learning. In early fusion, the multivariate

input sequence is augmented by concatenating the one-hot fault label at each time step, forcing the encoder to learn a fault-conditional representation. This changes the representation learning problem by explicitly conditioning feature extraction on the degradation identity, which can be detrimental under LODO because trajectories within the same fault category still exhibit variability in duration, degradation rate, noise levels, and drift patterns. As a result, early fusion can encourage the encoder to over-specialize toward mechanism-specific cues, weakening its ability to extract robust degradation embeddings that generalize to unseen trajectories within the same fault class. In contrast, late fusion preserves a stable, fault-invariant latent space by first learning an encoder representation without fault information and only injecting the degradation mode at the regression stage. This allows the encoder to capture generic degradation dynamics shared across all modes, while the class label acts as a conditioning variable that calibrates the final RUL mapping. These results justify selecting late fusion as the preferred injection point for degradation-mode knowledge, motivating the design choice in the subsequent asymmetric MTL framework.

Table 6.1: Fold-wise RUL prediction performance comparison between the baseline RUL estimator and fusion-based models under LODO cross-validation.

Test Set	RUL Estimator RMSE (h)	RUL Estimator MAE (h)	Early Fusion RMSE (h)	Early Fusion MAE (h)	Late Fusion RMSE (h)	Late Fusion MAE (h)
Redox Cycling 1	0.94	0.80	0.76	0.64	0.63	0.53
Redox Cycling 2	0.50	0.41	0.39	0.32	0.29	0.24
Redox Cycling 3	0.74	0.62	0.61	0.51	0.51	0.43
Fuel Starvation 1	0.55	0.45	0.44	0.36	0.37	0.31
Fuel Starvation 2	0.69	0.53	0.51	0.39	0.41	0.32
Fuel Starvation 3	0.70	0.57	0.57	0.46	0.45	0.36
Short Circuit 1	0.35	0.32	0.27	0.24	0.21	0.20
Short Circuit 2	0.14	0.12	0.11	0.09	0.09	0.08
Short Circuit 3	0.26	0.20	0.20	0.16	0.15	0.11
Thermal Shock 1	0.28	0.22	0.21	0.17	0.16	0.12
Thermal Shock 2	0.43	0.37	0.32	0.28	0.29	0.25
Thermal Shock 3	0.35	0.29	0.29	0.24	0.21	0.18
Average	0.49	0.41	0.39	0.32	0.31	0.26

To further interpret the performance differences between early and late fusion, the quality of the learned latent representations is analyzed using silhouette scores and pairwise centroid distances between degradation modes. Figure 6.7

illustrates the relationship between silhouette score (computed on training embeddings) and test RMSE across LODO folds, showing a clear negative correlation: better cluster separability in the latent space consistently corresponds to lower RUL prediction error. This trend is most pronounced for late fusion, which achieves both the highest silhouette scores and the lowest RMSE values, indicating that preserving a fault-invariant encoder representation leads to more discriminative and stable embeddings.

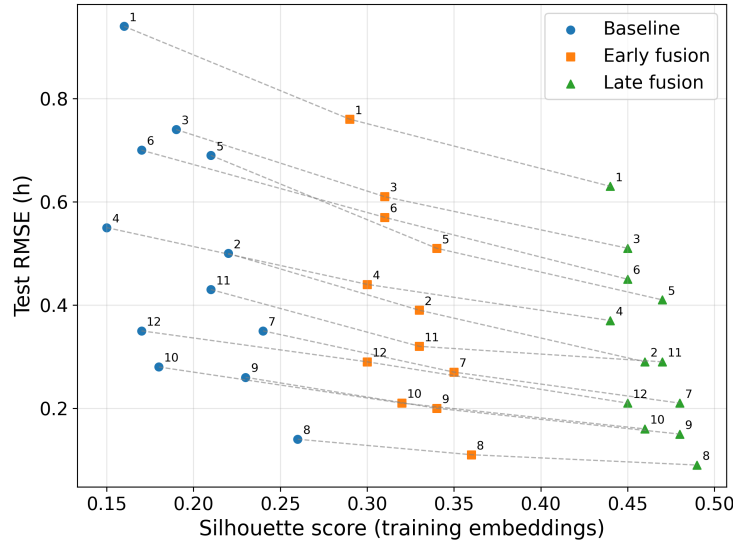


Figure 6.7: Silhouette score of training embeddings versus test RMSE for baseline, early fusion, and late fusion models across LODO folds.

Quantitatively, Table 6.2 summarizes the average pairwise centroid distances between degradation modes under different fusion strategies. Across all mode pairs, late fusion yields the largest centroid separations, followed by early fusion, while the baseline exhibits the weakest separability. The largest distance is observed between Redox Cycling and Short Circuit, reflecting their fundamentally different degradation dynamics (slow, diffusion-dominated degradation in Redox Cycling versus rapid, ohmic-driven degradation in Short circuit). In contrast, the smallest distance occurs between Thermal Shock and Fuel Starvation, consistent with both mechanisms being characterized by moderate degradation rates and dominant ohmic resistance growth. Intermediate separability is observed for Short Circuit-Fuel Starvation and for Short Circuit-

Thermal Shock, where overlapping ohmic effects reduce discriminability. The Redox Cycling-Thermal Shock and Redox Cycling-Fuel Starvation pairs exhibit better separation than these intermediate cases but remain less distinct than the extreme.

Table 6.2: Pairwise Euclidean centroid distances between degradation modes in the latent space for different fusion strategies, averaged across LODO folds. Distances reflect the relative separability learned by the encoder.

Degradation Mode Pair	Baseline	Early Fusion	Late Fusion
Redox Cycling – Short Circuit	8.985	11.811	11.960
Redox Cycling – Fuel Starvation	9.300	10.320	10.555
Redox Cycling – Thermal Shock	7.074	7.124	7.357
Short Circuit – Fuel Starvation	4.658	4.697	5.712
Short Circuit – Thermal Shock	5.440	6.217	6.579
Fuel Starvation – Thermal Shock	3.617	3.907	4.108

Asymmetric MTL Performance Table 6.3 reports the fold-wise comparison between the baseline RUL estimator and the asymmetric MTL model, where the classifier and RUL estimator are trained jointly and the classifier encoder embedding is fused into the RUL regressor. The asymmetric MTL model improves the baseline performance consistently across all folds, reducing the average RMSE from 0.49 h to 0.29 h and the average MAE from 0.41 h to 0.27 h. The improvement magnitudes are also reported explicitly per fold in Table 6.3, with average reductions of 42.8% in RMSE and 40.0% in MAE. This trend is summarized in Table 6.4, where asymmetric MTL provides the largest overall benefit compared to both early and late fusion, establishing the effectiveness of joint optimization and latent feature transfer.

The key reason asymmetric MTL can outperform late fusion is that it injects richer mechanism information than a one-hot class vector. In late fusion, the class label $\mathbf{c}$ provides only a discrete indicator of fault identity. In contrast, the classifier encoder produces a continuous latent embedding that captures not only the most likely degradation mode, but also uncertainty, inter-class similarity, and fine-grained degradation characteristics reflected in the input sequence. Even when classification is not perfect, the classifier embedding

can still provide partially informative features that help the RUL head disambiguate degradation regimes and improve calibration. This explains why asymmetric MTL can exceed late fusion performance despite relying on learned diagnostic features rather than an oracle label.

Table 6.3: Fold-wise RUL prediction performance comparison between the baseline RUL estimator and the asymmetric MTL model under LODO cross-validation. The magnitude of improvement varies across degradation modes and reflects classifier diagnostic reliability.

Test Set	RUL Estimator RMSE (h)	RUL Estimator MAE (h)	Asymmetric MTL RMSE (h)	Asymmetric MTL MAE (h)	RMSE ↓ (%)	MAE ↓ (%)
Redox Cycling 1	0.94	0.80	0.51	0.46	45.7	42.5
Redox Cycling 2	0.50	0.41	0.27	0.24	45.3	41.5
Redox Cycling 3	0.74	0.62	0.41	0.37	44.6	40.3
Fuel Starvation 1	0.55	0.45	0.32	0.27	41.8	40.0
Fuel Starvation 2	0.69	0.53	0.39	0.32	43.5	39.6
Fuel Starvation 3	0.70	0.57	0.40	0.34	42.9	40.4
Short Circuit 1	0.35	0.32	0.22	0.21	37.1	34.4
Short Circuit 2	0.14	0.12	0.09	0.08	38.6	36.7
Short Circuit 3	0.26	0.20	0.16	0.13	38.5	36.0
Thermal Shock 1	0.28	0.22	0.15	0.12	46.1	44.5
Thermal Shock 2	0.43	0.37	0.24	0.21	44.2	43.2
Thermal Shock 3	0.35	0.29	0.19	0.17	45.7	41.4
Average	0.49	0.41	0.29	0.27	42.8	40.0

Table 6.4: Comparison of RUL estimation strategies and their performance improvements relative to the baseline single-task RUL estimator.

Model Variant	Fusion Strategy Description	RMSE / MAE Reduction (%)
Single RUL Estimator (Baseline)	No degradation mode information used; RUL prediction based solely on degradation signals.	– / –
Early Fusion	True degradation class label is concatenated with the input feature vector at each time step prior to encoder processing.	21.49 ± 2.96 / 21.62 ± 2.93
Late Fusion	True degradation class label is concatenated with the pooled latent embedding after encoder processing and before the regression head.	37.34 ± 4.19 / 37.06 ± 4.86
Asymmetric MTL	Latent embedding from the classifier encoder is fused with the RUL encoder embedding; joint learning via asymmetric multi-task learning.	42.83 ± 3.02 / 40.04 ± 2.88

Computation Cost Comparison The computational cost of the standalone RUL estimator, the degradation mode classifier, and their sequential execution is compared with the proposed asymmetric MTL framework in Table 6.5. Although the MTL model contains a slightly larger number of trainable parameters, it reduces the average training cost per sample by 8.3%

(1.625 ms vs. 1.772 ms) and decreases the average inference time per sample by 7.5% (0.345 ms vs. 0.373 ms) relative to the combined RUL+Classifier baseline. These efficiency gains arise from the execution of a single integrated forward pass and the partial sharing of embedding and positional encoding operations across tasks. Overall, the results demonstrate that the asymmetric MTL strategy enhances predictive accuracy and mechanism awareness without increasing the operational computational burden, making it well suited for real-time SOFC monitoring applications.

Table 6.5: Computational cost comparison between the single-task RUL estimator, degradation mode classifier, their sequential execution (RUL+Classifier), and the proposed asymmetric MTL framework under LODO cross-validation.

Metric	RUL Estimator	Classifier	RUL+ Classifier	Asymmetric MTL	Cost ↓ (%)
Avg. Training time (per epoch per sample)(ms)	0.881	0.891	1.772	1.625	8.3
Avg. Inference time (per sample) (ms)	0.186	0.187	0.373	0.345	7.5
Number of trainable parameters	71,553	67,588	139,141	147,013	–

Classifier Performance Tables 6.6 and 6.7 report the degradation mode classification performance in the standalone setting and when the classifier is integrated into the asymmetric MTL framework. Although the overall performance remains high in both cases, a modest reduction is observed when training jointly (Macro F1-score decreases from 0.98 to 0.97). This degradation is expected in multi-task learning because the classifier encoder is no longer optimized solely for discrimination; instead, its representation is simultaneously shaped to support RUL estimation through the asymmetric fusion pathway. Despite the small reduction in standalone classification metrics, the classifier branch still provides highly informative mechanism embeddings that enhance the RUL prediction. Importantly, the fold-wise RUL improvements achieved by the asymmetric MTL framework (Table 6.3) are linked to the reliability of the classifier features, particularly the class-wise recall. When recall degrades for a specific degradation mode, the RUL estimator may receive a

less informative representation, reducing the achievable performance gain in the corresponding LODO fold. In our experiments, Short Circuit exhibits the lowest recall (0.95), which is consistent with the comparatively smaller RUL accuracy gains in the Short Circuit test folds. A similar trend is observed for Fuel Starvation, where a lower precision (0.92) results in reduced improvement compared to modes with near-perfect classification reliability.

Table 6.6: Standalone classifier performance of the degradation mode classifier under LODO cross-validation.

Degradation Mode	Precision	Recall	F1-score
Redox Cycling	1.00	0.98	0.99
Fuel Starvation	0.93	0.99	0.96
Short Circuit	1.00	0.97	0.98
Thermal Shock	0.99	0.99	0.99
Macro Avg	0.98	0.98	0.98

Table 6.7: Classification performance of the degradation mode classifier when integrated into the asymmetric MTL framework under LODO cross-validation.

Degradation Mode	Precision	Recall	F1-score
Redox Cycling	0.99	0.98	0.99
Fuel Starvation	0.92	0.98	0.95
Short Circuit	1.00	0.95	0.97
Thermal Shock	0.98	0.99	0.99
Macro Avg	0.97	0.97	0.97

6.5 Summary of Chapter

This chapter presented a mechanism-aware prognostics framework for tubular SOFCs by combining degradation mode identification and RUL prediction within a unified learning architecture. A comprehensive experimental dataset was first developed, comprising 822,463 samples collected from twelve nominally identical cells subjected to four controlled degradation mechanisms: redox cycling, fuel starvation, short circuit, and thermal shock. An Informer-style encoder was adopted as the baseline RUL estimator, and fusion-based experiments were conducted to examine how degradation information should

be incorporated to improve prognostic accuracy. Using true degradation mode labels, late fusion consistently outperformed early fusion, indicating that mechanism awareness is most effective when used to condition the final regression mapping. In addition, latent space analysis based on silhouette scores and class-pair centroid distances revealed improved separability of degradation mechanisms in late fusion and provided physical interpretability of the learned representations. In particular, redox cycling and short circuit degradation emerged as the most distinctly separated pair in the latent space, whereas fuel starvation and thermal shock formed the closest neighboring classes, reflecting their partially overlapping electrochemical signatures. Based on this insight, an asymmetric MTL design was introduced, where the classifier latent embedding is fused into the RUL estimation pathway while both tasks are optimized jointly. Under LODO evaluation, the proposed asymmetric MTL framework achieved an average RMSE reduction of approximately 43% relative to the mechanism-agnostic single-task estimator, confirming that diagnostic supervision can substantially enhance RUL prediction reliability. Moreover, the proposed framework remains computationally practical for online monitoring. Although the asymmetric MTL model includes slightly more trainable parameters than the sum of two separately trained networks, it provides approximately 8% lower training and inference cost compared with sequential execution of a standalone classifier and RUL estimator, due to shared embedding computations and a single forward pass. Overall, the results demonstrate that multi-task learning enables accurate and efficient SOFC prognostics by explicitly incorporating degradation mechanism awareness into the RUL estimation process.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis addressed the critical challenge of reliable health monitoring, diagnosis, and prognostics for solid oxide fuel cell (SOFC) systems, where complex degradation mechanisms and limited availability of comprehensive datasets hinder practical deployment. By combining experimental investigation with advanced data-driven and physics-informed learning approaches, a unified framework was developed to improve the accuracy, interpretability, and robustness of SOFC health assessment.

7.1.1 Experimental Framework and Dataset Development

A comprehensive experimental framework was established to generate diverse degradation datasets under multiple representative degradation modes, including redox cycling, fuel starvation, internal short circuit, and thermal shock. These experiments provided high-resolution, multi-modal measurements, including voltage, current, polarization curves, and electrochemical impedance spectroscopy (EIS) data. The resulting datasets form a valuable foundation for data-driven modeling, enabling systematic analysis of degradation behavior across different operating conditions and degradation mechanisms. The availability of such structured and diverse datasets addresses a key limitation in existing literature and supports the development of more generalizable

machine learning models for SOFC systems.

7.1.2 Physics-Informed and Probabilistic Prognostics

Building on the experimental datasets, a physics-informed framework was developed for accurate and interpretable state-of-health (SOH) estimation. By incorporating impedance-derived features extracted through distribution of relaxation times (DRT) analysis into a temporal graph convolutional network, the model effectively captured the interactions among underlying electrochemical processes. Compared to conventional data-driven approaches that rely primarily on voltage signals, the proposed physics-informed framework provided improved interpretability and more reliable degradation forecasting by incorporating physically meaningful electrochemical indicators into the learning process. The results demonstrated that integrating DRT-based features substantially improved prediction accuracy while also enabling better understanding of degradation evolution and dominant electrochemical processes.

This work was further extended to long-horizon prognostics through the development of a probabilistic framework for SOH trajectory forecasting and remaining useful life (RUL) estimation. Attention-based sequence models, particularly the Informer architecture, were employed to capture long-range temporal dependencies in degradation trajectories. The results demonstrated that attention-based forecasting models are well-suited for long-horizon SOH prediction, where conventional recurrent architectures often struggle to maintain accuracy over extended prediction windows. To improve reliability, both aleatoric and epistemic uncertainties were explicitly modeled using Bayesian learning approaches, enabling the generation of calibrated confidence intervals for SOH and RUL predictions. These uncertainty estimates provide important information for risk-aware maintenance scheduling and operational planning by quantifying the confidence associated with degradation forecasts. In addition, a dynamic forecasting strategy based on exponentially weighted zero-padding was introduced to address the challenge of shrinking prediction horizons as the system approaches end-of-life. Together, these contributions

advance the state-of-the-art in SOFC prognostics by enabling accurate, interpretable, uncertainty-aware, and long-horizon degradation forecasting.

7.1.3 Impedance-Based Diagnostics

In parallel with prognostics, a diagnostic framework was developed to identify the underlying degradation mechanisms responsible for performance degradation. By leveraging EIS measurements and DRT-based feature extraction, physically interpretable signatures of degradation processes were obtained. These features were used to train and evaluate multiple machine learning classifiers, including both traditional ensemble methods and temporal deep learning models.

The results demonstrated that sequence-based and attention-based architectures provide superior performance in distinguishing between different degradation modes, particularly for mechanistically similar degradation processes with partially overlapping impedance characteristics. The study also highlighted the importance of temporal modeling in improving pairwise class separability between degradation modes that exhibit similar electrochemical behavior in certain frequency regions. In addition, the challenges associated with class imbalance in degradation datasets were addressed through physics-consistent augmentation strategies based on repeated experiments under identical operating conditions, improving classifier robustness and generalization. The impedance-derived DRT features further enabled physical interpretation of class-wise classification performance by linking dominant spectral regions and relaxation processes to specific degradation mechanisms.

7.1.4 Unified Mechanism-Aware Prognostics Framework

Recognizing the strong dependency of degradation behavior on degradation mechanisms, this thesis further developed a unified framework that integrates diagnostics and prognostics within a multi-task learning (MTL) architecture. In contrast to conventional approaches that treat classification and RUL prediction as separate tasks, the proposed framework jointly learns both tasks

using a shared representation based on an attention-driven Informer model. The results demonstrated that incorporating degradation mechanism information into the prognostic framework significantly improves RUL prediction accuracy, highlighting the strong dependency of SOFC lifetime evolution on the underlying degradation mode. By transferring latent representations from the diagnostic branch to the prognostic branch, the model effectively integrates degradation mechanism awareness into long-term forecasting while maintaining a unified learning structure. In addition, the proposed framework demonstrated improved computational efficiency compared to sequential execution of separate diagnostic and prognostic models, making it more suitable for practical deployment in real-time health monitoring systems.

An ablation study further showed that degradation mechanism information is most effectively incorporated after the temporal encoding stage, indicating that the feature extraction process should remain largely degradation-mechanism agnostic while degradation-specific information is fused at higher representation levels. This finding provides important insight into the design of unified prognostic-diagnostic architectures for multi-degradation-mode systems.

7.1.5 Overall Contributions and Impact

Overall, this thesis presents a comprehensive and scalable framework for SOFC health monitoring that integrates experimental data generation, physics-informed feature extraction, uncertainty-aware prognostics, impedance-based diagnostics, and unified multi-task learning. The proposed methodologies address key limitations in existing research, including the lack of diverse datasets, limited interpretability of data-driven models, absence of uncertainty quantification, and the separation of diagnostic and prognostic tasks.

By bridging these gaps, this work enables more accurate, robust, and interpretable health assessment of SOFC systems, supporting predictive maintenance and health-aware control strategies. The integration of physics-informed learning with advanced deep learning architectures further demonstrates the

potential of combining domain knowledge with data-driven methods for complex energy systems. Ultimately, the contributions of this thesis provide a foundation for the reliable and efficient deployment of SOFC technologies in future clean energy applications.

Figure 7.1 summarizes the overall framework and practical guidelines developed in this thesis for SOFC diagnostics and prognostics. The figure presents the complete workflow from experimental degradation testing and impedance-based characterization to model selection, training, and validation. It also summarizes the major prognostic and diagnostic modeling approaches investigated throughout this thesis, including physics-informed, probabilistic, attention-based, and mechanism-aware learning frameworks.

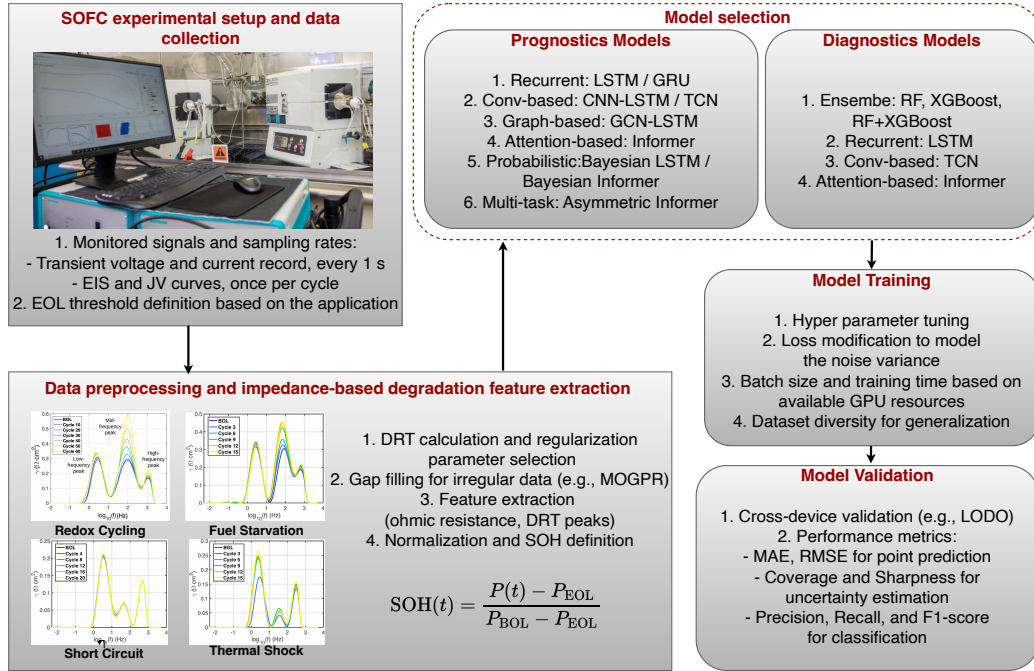


Figure 7.1: Summary of the proposed SOFC diagnostics and prognostics framework developed in this thesis, including experimental degradation testing, impedance-based feature extraction, model selection, training, and validation strategies for physics-informed and data-driven health monitoring.

7.2 Future Works

While this thesis presents a comprehensive framework for SOFC health monitoring, diagnostics, and prognostics, several avenues remain for further research to enhance the applicability, robustness, and practical deployment of the proposed methodologies. Future work should focus on extending model generalizability, enabling real-time deployment, integrating control strategies, and leveraging prognostic insights for operational decision-making. The key directions for future research are outlined as follows:

7.2.1 Health-aware and fault-tolerant control integration

An important extension of this work is the integration of prognostic and diagnostic models into real-time supervisory and closed-loop control frameworks. The predicted SOH trajectories and RUL estimates can provide actionable insights for adjusting operating conditions, such as fuel and air flow rates and current density, to mitigate degradation and extend system lifetime. Future studies should investigate the incorporation of these predictions into advanced control strategies, such as model predictive control, enabling both health-aware operation and fault-tolerant response to emerging degradation. In addition, stability and robustness analysis under closed-loop conditions is required to ensure safe and reliable deployment of such integrated control systems.

7.2.2 Integration with hybrid energy management systems

In hybrid SOFC–battery systems, prognostic outputs can support optimal energy management. Future studies should explore how SOH and RUL predictions can be used to dynamically allocate power demand, enabling SOFC derating during degraded conditions while leveraging batteries to handle transient loads.

7.2.3 Uncertainty-aware control and decision-making

The probabilistic outputs of the proposed models, including confidence intervals and uncertainty estimates, can be used to develop robust control strategies. Future work should investigate uncertainty-aware control frameworks that incorporate safety margins, constraint tightening, and risk-sensitive decision-making under degradation.

7.2.4 Risk-aware maintenance scheduling and planning

The predicted RUL distributions and associated uncertainty bounds provide a foundation for advanced maintenance strategies. Future research should focus on integrating these predictions into maintenance planning frameworks to optimize inspection intervals, maintenance scheduling, and component replacement decisions under uncertainty.

7.2.5 External validation

Although the proposed models were validated using extensive laboratory-generated datasets, real-world operating environments often exhibit higher variability, measurement noise, and unmodeled disturbances. Future work should prioritize external validation using independent datasets collected under diverse conditions, including different cell architectures (e.g., planar stacks versus tubular cells), additional degradation mechanisms (e.g., sulfur poisoning and carbon deposition), and realistic duty cycles with balance-of-plant interactions. Furthermore, since accelerated degradation experiments may introduce biases compared to natural aging, comparative studies between accelerated and field-representative degradation trajectories are needed to assess model reliability in practical applications.

7.2.6 Scaling to stack-level and fleet-level health management

Extending the proposed framework from single-cell analysis to stack-level and fleet-level applications introduces additional challenges, including cell-to-cell

variability, thermal gradients, and interacting degradation mechanisms. Future research should explore hierarchical and graph-based modeling approaches to capture these interactions, enabling fault localization, system-level degradation estimation, and coordinated health monitoring across multiple units.

7.2.7 Online transfer learning and adaptation

SOFC systems exhibit non-stationary behavior due to evolving degradation dynamics and changing operating conditions. Future research should incorporate online adaptation mechanisms, such as continual learning, streaming calibration, and drift-aware model updates, to maintain prediction accuracy over time. In addition, transfer learning approaches should be explored to enable efficient deployment across different cells and system configurations, reducing data requirements for new units. Domain adaptation techniques can further address distribution shifts between laboratory and field data.

7.2.8 Real-time implementation

Although the proposed models demonstrate strong performance on high-performance computing platforms, practical deployment requires implementation on resource-constrained hardware. Future work should evaluate latency, memory usage, and energy consumption on embedded systems, such as edge devices or industrial controllers. Model compression techniques, including quantization, pruning, and knowledge distillation, should be investigated to ensure real-time feasibility under strict computational and power constraints.

7.2.9 Minimum experimental dataset requirements

Future work should investigate the minimum number of independent experimental tests required for each degradation mode to train and validate diagnostic and prognostic models reliably. This analysis should evaluate how model accuracy, uncertainty, and generalization change as additional trajectories are introduced, enabling the definition of statistically justified dataset-size requirements for each degradation mechanism.

7.2.10 Closed-loop co-design of cells, experiments, and prognostics

Future research should establish a feedback framework in which prognostic performance informs cell design, experimental planning, and data-processing decisions. For example, model sensitivity and uncertainty could be used to guide EIS acquisition settings, DRT preprocessing parameters, and the number and location of extracted DRT peaks, ensuring that the collected features are both physically meaningful and sufficiently informative for downstream diagnostic and prognostic models.

References

- [1] Z. Salehi *et al.*, “State-of-health forecasting of solid oxide fuel cells using physics-informed temporal graph convolutional network,” *Journal of Power Sources*, vol. 665, p. 239 050, 2026, ISSN: 0378-7753. DOI: 10 . 1016/j . jpowsour . 2025 . 239050.
- [2] Z. Salehi, M. F. Hasanabadi, D. J. Smith, A. R. Hanifi, and M. Shahbakhti, “Bayesian informer for probabilistic remaining useful life forecasting of solid oxide fuel cells,” *Control Engineering Practice*, vol. 173, p. 107 011, 2026, ISSN: 0967-0661. DOI: 10 . 1016/j . conengprac . 2026 . 107011.
- [3] Z. Salehi, M. F. Hasanabadi, D. J. Smith, A. R. Hanifi, and M. Shahbakhti, “Uncertainty-aware degradation trajectory forecasting for fuel cell prognostics and health management,” in *Proceedings of the 23rd IFAC World Congress*, Accepted, Busan, Republic of Korea, 2026.
- [4] Z. Salehi, S. Sobhi, S. Hanifi, D. J. Smith, A. R. Hanifi, and M. Shahbakhti, “Diagnosis and health monitoring of solid oxide fuel cells using electrochemical impedance spectroscopy under multiple failure modes,” *Journal of Power Sources*, 2026, Under review.
- [5] Z. Salehi, D. J. Smith, A. R. Hanifi, and M. Shahbakhti, “Degradation mechanism-aware prognostics of solid oxide fuel cells via asymmetric multi-task learning,” *Renewable Energy*, 2026, Under review.
- [6] Z. Salehi *et al.*, “Transfer learning-based deep neural network model for performance prediction of hydrogen-fueled solid oxide fuel cells,” *International Journal of Hydrogen Energy*, vol. 99, pp. 102–111, 2025, ISSN: 0360-3199. DOI: 10 . 1016/j . ijhydene . 2024 . 12 . 092.
- [7] Z. Salehi *et al.*, “Performance prediction of a range of diverse solid oxide fuel cells using deep learning and principal component analysis,” *IFAC-PapersOnLine*, vol. 58, no. 28, pp. 935–940, 2024, The 4th Modeling, Estimation, and Control Conference – 2024, ISSN: 2405-8963. DOI: 10 . 1016/j . ifacol . 2025 . 01 . 116.
- [8] P. Costamagna, A. Selimovic, M. D. Borghi, and G. Agnew, “Electrochemical model of the integrated planar solid oxide fuel cell (IP-SOFC),” *Chemical Engineering Journal*, vol. 102, no. 1, pp. 61–69, Aug. 2004. DOI: 10 . 1016/j . cej . 2004 . 02 . 005.

- [9] E. R. Nielsen *et al.*, “Status on demonstration of fuel cell based micro-CHP units in europe,” *Fuel Cells*, Jul. 2019. DOI: 10.1002/fuce.201800189.
- [10] T. V. V. S. Lakshmi, P. Geethanjali, and P. S. Krishna, “Mathematical modelling of solid oxide fuel cell using matlab/simulink,” *IEEE*, Jun. 2013. DOI: 10.1109/AICERA-ICMiCR.2013.6576016.
- [11] Y. Wang and X.-D. Zhou, “Theoretical Understanding of Effects of Operating Modes on the Performance Durability of Solid Oxide Cells: A Comparison between Potentiostatic and Galvanostatic Operations,” *Journal of The Electrochemical Society*, vol. 171, no. 3, p. 034516, Apr. 2024, ISSN: 1945-7111. DOI: 10.1149/1945-7111/ad3059.
- [12] L. Del Zotto, G. Tamburrano, A. Dell’Era, A. Hatunoglu, and E. Ciro, “Determination and validation of polarization losses parameters to predict current/voltage-characteristics for SOFC button cell,” *Energy Conversion and Management*, vol. 299, p. 117877, Jan. 2024, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2023.117877.
- [13] H. Shimada, T. Suzuki, T. Yamaguchi, H. Sumi, K. Hamamoto, and Y. Fujishiro, “Challenge for lowering concentration polarization in solid oxide fuel cells,” *Journal of Power Sources*, vol. 302, pp. 53–60, Jan. 2016. DOI: 10.1016/j.jpowsour.2015.10.024.
- [14] A. K. ElKashat, A. R. Razmi, S. Vafaeenezhad, A. R. Hanifi, T. Etsell, and M. Shahbakhti, “Experimental and thermodynamic analyses of a novel anode supported solid oxide fuel cell,” *Canadian Society for Mechanical Engineering (CSME) Conference, Edmonton, Canada*, 2022. DOI: 10.7939/r3-zr8x-cv76.
- [15] M. Atif Mahmood *et al.*, “Sensitivity analysis of performance and thermal impacts of a single hydrogen fueled solid oxide fuel cell to optimize the operational and design parameters,” *Sustainable Energy Technologies and Assessments*, vol. 57, p. 103241, Jun. 2023, ISSN: 2213-1388. DOI: 10.1016/j.seta.2023.103241. (visited on 10/28/2024).
- [16] M. Gandiglio, P. Marocco, A. Nieminen, M. Santarelli, and J. Kiviahio, “Energy and environmental performance from field operation of commercial-scale SOFC systems,” *International Journal of Hydrogen Energy*, vol. 85, pp. 997–1009, Oct. 2024, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2024.08.332. (visited on 10/28/2024).
- [17] M. Tofigh *et al.*, “Control-Oriented Modeling of a Solid Oxide Fuel Cell Affected by Redox Cycling Using a Novel Deep Learning Approach,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 147, no. 2, p. 021006, Sep. 2024, ISSN: 0022-0434. DOI: 10.1115/1.4066268.

- [18] J. Peng, D. Zhao, Y. Xu, X. Wu, and X. Li, “Comprehensive Analysis of Solid Oxide Fuel Cell Performance Degradation Mechanism, Prediction, and Optimization Studies,” en, *Energies*, vol. 16, no. 2, p. 788, Jan. 2023, ISSN: 1996-1073. DOI: 10.3390/en16020788.
- [19] S. Zarabi Golkhatmi, M. I. Asghar, and P. D. Lund, “A review on solid oxide fuel cell durability: Latest progress, mechanisms, and study tools,” *Renewable and Sustainable Energy Reviews*, vol. 161, p. 112 339, Jun. 2022, ISSN: 1364-0321. DOI: 10.1016/j.rser.2022.112339.
- [20] V. Subotić and C. Hochenauer, “Analysis of solid oxide fuel and electrolysis cells operated in a real-system environment: State-of-the-health diagnostic, failure modes, degradation mitigation and performance regeneration,” *Progress in Energy and Combustion Science*, vol. 93, p. 101 011, Nov. 2022, ISSN: 0360-1285. DOI: 10.1016/j.pecs.2022.101011.
- [21] M. Gallo, P. Polverino, J. Mougin, B. Morel, and C. Pianese, “Coupling electrochemical impedance spectroscopy and model-based aging estimation for solid oxide fuel cell stacks lifetime prediction,” *Applied Energy*, vol. 279, p. 115 718, Dec. 2020, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115718.
- [22] H. Sumi, H. Shimada, Y. Yamaguchi, T. Yamaguchi, and Y. Fujishiro, “Degradation evaluation by distribution of relaxation times analysis for microtubular solid oxide fuel cells,” *Electrochimica Acta*, vol. 339, p. 135 913, Apr. 2020, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2020.135913.
- [23] Z. Gong, W. Sun, J. Cao, Y. Wu, L. Miao, and W. Liu, “A new in situ strategy to eliminate partial internal short circuit in ce0.8sm0.2o1.9-based solid oxide fuel cells,” *J. Mater. Chem. A*, vol. 5, pp. 12 873–12 878, 25 2017. DOI: 10.1039/C7TA03527F.
- [24] X.-l. Wu *et al.*, “Health state prediction and analysis of SOFC system based on the data-driven entire stage experiment,” *Applied Energy*, vol. 248, pp. 126–140, Aug. 2019, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.04.053.
- [25] N. Q. Minh and T. Takahashi, “Chapter 6 - Anode,” in N. Q. Minh and T. Takahashi, Eds., Oxford: Elsevier Science Ltd, Jan. 1995, pp. 147–164, ISBN: 9780444895684. DOI: 10.1016/B978-044489568-4/50007-4.
- [26] M. Cassidy, G. Lindsay, and K. Kendall, “The reduction of nickel zirconia cermet anodes and the effects on supported thin electrolytes,” *Journal of Power Sources*, vol. 61, no. 1, pp. 189–192, Jul. 1996, ISSN: 0378-7753. DOI: 10.1016/S0378-7753(96)02359-2.

- [27] D. Waldbillig, A. Wood, and D. G. Ivey, "Electrochemical and microstructural characterization of the redox tolerance of solid oxide fuel cell anodes," *Journal of Power Sources*, Selected papers presented at the Fuel Cells Science and Technology Meeting, vol. 145, no. 2, pp. 206–215, Aug. 2005, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2004.12.071.
- [28] J. Wang *et al.*, "Understanding thermal and redox cycling behaviors of flat-tube solid oxide fuel cells," *International Journal of Hydrogen Energy*, vol. 48, no. 57, pp. 21 886–21 897, 2023, ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2023.03.062>.
- [29] M. Ettler, H. Timmermann, J. Malzbender, A. Weber, and N. H. Menzler, "Durability of Ni anodes during reoxidation cycles," *Journal of Power Sources*, vol. 195, no. 17, pp. 5452–5467, Sep. 2010, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2010.03.049.
- [30] G. Brus, K. Miyoshi, H. Iwai, M. Saito, and H. Yoshida, "Change of an anode's microstructure morphology during the fuel starvation of an anode-supported solid oxide fuel cell," *International Journal of Hydrogen Energy*, vol. 40, no. 21, pp. 6927–6934, 2015, ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2015.03.143>.
- [31] W. Sun and W. Liu, "A novel ceria-based solid oxide fuel cell free from internal short circuit," *Journal of Power Sources*, vol. 217, pp. 114–119, 2012, ISSN: 0378-7753. DOI: <https://doi.org/10.1016/j.jpowsour.2012.05.065>.
- [32] Y. Ling *et al.*, "Review of experimental and modelling developments for ceria-based solid oxide fuel cells free from internal short circuits," *Journal of Materials Science*, vol. 55, pp. 1–23, 2020. DOI: 10.1007/s10853-019-03876-z.
- [33] Y. Wang, Y. Shi, X. Yu, and N. Cai, "Thermal shock resistance and failure probability analysis on solid oxide electrolyte direct flame fuel cells," *Journal of Power Sources*, vol. 255, pp. 377–386, 2014, ISSN: 0378-7753. DOI: <https://doi.org/10.1016/j.jpowsour.2014.01.035>.
- [34] C.-K. Lin, T.-T. Chen, Y.-P. Chyou, and L.-K. Chiang, "Thermal stress analysis of a planar sofc stack," *Journal of Power Sources*, vol. 164, no. 1, pp. 238–251, 2007, ISSN: 0378-7753. DOI: <https://doi.org/10.1016/j.jpowsour.2006.10.089>.
- [35] G. De Marco *et al.*, "Testing the voltage and power as function of current density: Polarisation curve for a sofc single cell," *JRC: Petten, The Netherlands*, 2010.
- [36] G. De Marco *et al.*, "Sofc single cell performance and endurance test modules," 2010.

- [37] N. Thambiraj, I. Waernhus, C. Suci, A. Vik, and A. C. Hoffmann, "Single-Cell Tests to Explore the Reliability of Sofc Installations Operating Offshore," en, *Energies*, vol. 13, no. 7, p. 1624, Jan. 2020, ISSN: 1996-1073. DOI: 10.3390/en13071624.
- [38] D. D. Macdonald, "Reflections on the history of electrochemical impedance spectroscopy," *Electrochimica Acta*, Electrochemical Impedance Spectroscopy, vol. 51, no. 8, pp. 1376–1388, Jan. 2006, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2005.02.107.
- [39] F. Fasmin and R. Srinivasan, "Review—Nonlinear Electrochemical Impedance Spectroscopy," en, *Journal of The Electrochemical Society*, vol. 164, no. 7, H443, May 2017, ISSN: 1945-7111. DOI: 10.1149/2.0391707jes.
- [40] D. Vladikova, "The technique of the differential impedance analysis Part I: Basics of the impedance spectroscopy," Jan. 2004.
- [41] K. M. Jeerage, R. D. Noble, and C. A. Koval, "Investigation of an aqueous lithium iodide/triiodide electrolyte for dual-chamber electrochemical actuators," *Sensors and Actuators B: Chemical*, vol. 125, no. 1, pp. 180–188, Jul. 2007, ISSN: 0925-4005. DOI: 10.1016/j.snb.2007.02.003.
- [42] V. Subotić *et al.*, "In-situ electrochemical characterization methods for industrial-sized planar solid oxide fuel cells Part I: Methodology, qualification and detection of carbon deposition," *Electrochimica Acta*, vol. 207, pp. 224–236, Jul. 2016, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2016.05.025.
- [43] A. K. Padinjarethil, S. Pollok, and A. Hagen, "Degradation studies using machine learning on novel solid oxide cell database," en, *Fuel Cells*, vol. 21, no. 6, pp. 566–576, 2021, ISSN: 1615-6854. DOI: 10.1002/fuce.202100097.
- [44] Q.-A. Huang, R. Hui, B. Wang, and J. Zhang, "A review of AC impedance modeling and validation in SOFC diagnosis," *Electrochimica Acta*, vol. 52, no. 28, pp. 8144–8164, Nov. 2007, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2007.05.071.
- [45] A. Baldinelli, A. Staffolani, G. Bidini, L. Barelli, and F. Nobili, "An extensive model for renewable energy electrochemical storage with Solid Oxide Cells based on a comprehensive analysis of impedance deconvolution," *Journal of Energy Storage*, vol. 33, p. 102052, Jan. 2021, ISSN: 2352-152X. DOI: 10.1016/j.est.2020.102052.
- [46] C. Graves, S. D. Ebbesen, and M. Mogensen, "Co-electrolysis of CO₂ and H₂O in solid oxide cells: Performance and durability," *Solid State Ionics*, Proceedings of the 17th International Conference on Solid State Ionics, vol. 192, no. 1, pp. 398–403, Jun. 2011, ISSN: 0167-2738. DOI: 10.1016/j.ssi.2010.06.014.

- [47] Y. Yang *et al.*, “Study of solid oxide electrolysis cells operated in potentiostatic mode: Effect of operating temperature on durability,” *Chemical Engineering Journal*, vol. 417, p. 129 260, Aug. 2021, ISSN: 1385-8947. DOI: 10.1016/j.cej.2021.129260.
- [48] V. Subotić, B. Stoeckl, V. Lawlor, J. Strasser, H. Schroettner, and C. Hochenauer, “Towards a practical tool for online monitoring of solid oxide fuel cell operation: An experimental study and application of advanced data analysis approaches,” *Applied Energy*, vol. 222, pp. 748–761, Jul. 2018, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.03.182.
- [49] B. Liu, Y. Zhang, B. Tu, Y. Dong, and M. Cheng, “Electrochemical impedance investigation of the redox behaviour of a Ni-YSZ anode,” *Journal of Power Sources*, vol. 165, no. 1, pp. 114–119, Feb. 2007, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2006.11.052.
- [50] B. Liu, T. Matsui, H. Muroyama, K. Tomida, T. Kabata, and K. Eguchi, “Impedance Analysis of Practical Segmented-in-Series Tubular Solid Oxide Fuel Cells,” en, *ECS Transactions*, vol. 35, no. 1, p. 637, Apr. 2011, ISSN: 1938-5862. DOI: 10.1149/1.3570042.
- [51] V. Subotić *et al.*, “Performance assessment of industrial-sized solid oxide cells operated in a reversible mode: Detailed numerical and experimental study,” *International Journal of Hydrogen Energy*, vol. 45, no. 53, pp. 29 166–29 185, Oct. 2020, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2020.07.165.
- [52] V. Subotić, P. Harter, M. Kusnezoff, T. W. Napporn, H. Schroettner, and C. Hochenauer, “Identification of carbon deposition and its removal in solid oxide fuel cells by applying a non-conventional diagnostic tool,” en, *Sustainable Energy & Fuels*, vol. 5, no. 7, pp. 2065–2076, Apr. 2021, ISSN: 2398-4902. DOI: 10.1039/D0SE01914C.
- [53] T.-H. Lee *et al.*, “Degradation analysis of anode-supported intermediate temperature-solid oxide fuel cells under various failure modes,” *Journal of Power Sources*, vol. 276, pp. 120–132, Feb. 2015, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2014.11.077.
- [54] M. J. Jørgensen, S. Primdahl, and M. Mogensen, “Characterisation of composite SOFC cathodes using electrochemical impedance spectroscopy,” *Electrochimica Acta*, vol. 44, no. 24, pp. 4195–4201, Jul. 1999, ISSN: 0013-4686. DOI: 10.1016/S0013-4686(99)00134-6.
- [55] A. Kromp, A. Leonide, A. Weber, and E. Ivers-Tiffée, “Electrochemical Analysis of Reformate-Fuelled Anode Supported SOFC,” en, *Journal of The Electrochemical Society*, vol. 158, no. 8, B980, Jun. 2011, ISSN: 1945-7111. DOI: 10.1149/1.3597177.

- [56] V. Vibhu, S. Yildiz, I. C. Vinke, R.-A. Eichel, J.-M. Bassat, and L. G. J. de Haart, “High performance LSC infiltrated LSCF oxygen electrode for high temperature steam electrolysis application,” *Journal of The Electrochemical Society*, vol. 166, no. 2, F102–F108, 2019. DOI: 10.1149/2.0741902jes.
- [57] G. Nusev, B. Morel, J. Mougin, Đ. Juričić, and P. Boškoski, “Condition monitoring of solid oxide fuel cells by fast electrochemical impedance spectroscopy: A case example of detecting deficiencies in fuel supply,” *Journal of Power Sources*, vol. 489, p. 229 491, Mar. 2021, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2021.229491.
- [58] L. Ljung, “System Identification,” en, in ser. Applied and Numerical Harmonic Analysis, A. Procházka, J. Uhlíř, P. W. J. Rayner, and N. G. Kingsbury, Eds., Boston, MA: Birkhäuser, 1998, pp. 163–173, ISBN: 9781461217688. DOI: 10.1007/978-1-4612-1768-8_11.
- [59] P. Fortin, M. R. Gerhardt, Ø. Ulleberg, F. Zenith, and T. Holm, “Multi-Sine EIS for Early Detection of PEMFC Failure Modes,” *Frontiers in Energy Research*, vol. 10, 2022, ISSN: 2296-598X.
- [60] A. Leonide, S. Hansmann, A. Weber, and E. Ivers-Tiffée, “Performance simulation of current/voltage-characteristics for SOFC single cell by means of detailed impedance analysis,” *Journal of Power Sources*, Proceedings of 2010 European Solid Oxide Fuel Cell Forum, vol. 196, no. 17, pp. 7343–7346, Sep. 2011, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2010.10.052.
- [61] A. Kromp, S. Dierickx, A. Leonide, A. Weber, and E. Ivers-Tiffée, “Electrochemical Analysis of Sulfur-Poisoning in Anode Supported SOFCs Fuelled with a Model Reformate,” en, *Journal of The Electrochemical Society*, vol. 159, no. 5, B597, Mar. 2012, ISSN: 1945-7111. DOI: 10.1149/2.015206jes.
- [62] H. Sumi *et al.*, “AC impedance characteristics for anode-supported microtubular solid oxide fuel cells,” *Electrochimica Acta*, vol. 67, pp. 159–165, Apr. 2012, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2012.02.021.
- [63] J. Nielsen and M. Mogensen, “SOFC LSM:YSZ cathode degradation induced by moisture: An impedance spectroscopy study,” *Solid State Ionics*, vol. 189, no. 1, pp. 74–81, May 2011, ISSN: 0167-2738. DOI: 10.1016/j.ssi.2011.02.019.
- [64] M. Ghamarinia, A. Babaei, C. Zamani, and H. Aslannejad, “Application of the distribution of relaxation time method in electrochemical analysis of the air electrodes in the SOFC/SOEC devices: A review,” *Chemical Engineering Journal Advances*, vol. 15, p. 100 503, Aug. 2023, ISSN: 2666-8211. DOI: 10.1016/j.cej.2023.100503.

- [65] A. Nechache, M. Cassir, and A. Ringuedé, “Solid oxide electrolysis cell analysis by means of electrochemical impedance spectroscopy: A review,” *Journal of Power Sources*, vol. 258, pp. 164–181, Jul. 2014, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2014.01.110.
- [66] S. Dierickx, A. Weber, and E. Ivers-Tiffée, “How the distribution of relaxation times enhances complex equivalent circuit models for fuel cells,” *Electrochimica Acta*, vol. 355, p. 136 764, Sep. 2020, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2020.136764.
- [67] J. Weese, “A reliable and fast method for the solution of Fredholm integral equations of the first kind based on Tikhonov regularization,” *Computer Physics Communications*, vol. 69, no. 1, pp. 99–111, Feb. 1992, ISSN: 0010-4655. DOI: 10.1016/0010-4655(92)90132-I.
- [68] M. Saccoccio, T. H. Wan, C. Chen, and F. Ciucci, “Optimal Regularization in Distribution of Relaxation Times applied to Electrochemical Impedance Spectroscopy: Ridge and Lasso Regression Methods - A Theoretical and Experimental Study,” *Electrochimica Acta*, vol. 147, pp. 470–482, Nov. 2014, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2014.09.058.
- [69] F. Ciucci and C. Chen, “Analysis of Electrochemical Impedance Spectroscopy Data Using the Distribution of Relaxation Times: A Bayesian and Hierarchical Bayesian Approach,” *Electrochimica Acta*, vol. 167, pp. 439–454, Jun. 2015, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2015.03.123.
- [70] M. B. Effat and F. Ciucci, “Bayesian and Hierarchical Bayesian Based Regularization for Deconvolving the Distribution of Relaxation Times from Electrochemical Impedance Spectroscopy Data,” *Electrochimica Acta*, vol. 247, pp. 1117–1129, Sep. 2017, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2017.07.050.
- [71] G. DiGiuseppe, A. Hunter, and F. Zhu, “Combined equivalent circuits and distribution of relaxation times analysis and interfacial effects of (La_{0.60}Sr_{0.40})_{0.95}Co_{0.20}Fe_{0.80}O_{3-x} cathodes,” *Electrochimica Acta*, vol. 350, p. 136 252, Aug. 2020, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2020.136252.
- [72] Y. Zhao *et al.*, “Performance and distribution of relaxation times analysis of Ruddlesden-Popper oxide Sr₃Fe_{1.3}Co_{0.2}Mo_{0.5}O₇₋ as a potential cathode for protonic solid oxide fuel cells,” *Electrochimica Acta*, vol. 352, p. 136 444, Aug. 2020, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2020.136444.
- [73] N. Shi *et al.*, “Performance and DRT analysis of P-SOFCs fabricated using new phase inversion combined tape casting technology,” en, *Journal of Materials Chemistry A*, vol. 5, no. 37, pp. 19 664–19 671, Sep. 2017, ISSN: 2050-7496. DOI: 10.1039/C7TA04967F.

- [74] X. Wang *et al.*, “Charge-Transfer Modeling and Polarization DRT Analysis of Proton Ceramics Fuel Cells Based on Mixed Conductive Electrolyte with the Modified Anode–Electrolyte Interface,” *ACS Applied Materials & Interfaces*, vol. 10, no. 41, pp. 35 047–35 059, Oct. 2018, ISSN: 1944-8244. DOI: 10.1021/acsami.8b10429.
- [75] W. Cai *et al.*, “Characterization and polarization DRT analysis of a stable and highly active proton-conducting cathode,” *Ceramics International*, vol. 44, no. 12, pp. 14 297–14 302, Aug. 2018, ISSN: 0272-8842. DOI: 10.1016/j.ceramint.2018.05.035.
- [76] A. Hauch, A. Hagen, J. Hjelm, and T. Ramos, “Sulfur Poisoning of SOFC Anodes: Effect of Overpotential on Long-Term Degradation,” en, *Journal of The Electrochemical Society*, vol. 161, no. 6, F734, May 2014, ISSN: 1945-7111. DOI: 10.1149/2.080406jes.
- [77] A. Soltanizade, A. Babaei, A. Ataie, and S. V. Seyed-Vakili, “Temperature dependency of activity of nano-catalysts on La_{0.6}Sr_{0.4}Co_{0.2}Fe_{0.8}O₃–cathode of solid oxide fuel cells,” en, *Journal of Applied Electrochemistry*, vol. 49, no. 11, pp. 1113–1122, Nov. 2019, ISSN: 1572-8838. DOI: 10.1007/s10800-019-01355-6.
- [78] D. Clematis, A. Barbucci, S. Presto, M. Viviani, and M. P. Carpanese, “Electrocatalytic activity of perovskite-based cathodes for solid oxide fuel cells,” *International Journal of Hydrogen Energy*, vol. 44, no. 12, pp. 6212–6222, Mar. 2019, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2019.01.128.
- [79] H. Sumi, T. Yamaguchi, K. Hamamoto, T. Suzuki, and Y. Fujishiro, “High performance of La_{0.6}Sr_{0.4}Co_{0.2}Fe_{0.8}O₃–Ce_{0.9}Gd_{0.1}O_{1.95} nanoparticulate cathode for intermediate temperature microtubular solid oxide fuel cells,” *Journal of Power Sources*, vol. 226, pp. 354–358, Mar. 2013, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2012.11.015.
- [80] P. Caliandro, A. Nakajo, S. Diethelm, and J. Van herle, “Model-assisted identification of solid oxide cell elementary processes by electrochemical impedance spectroscopy measurements,” *Journal of Power Sources*, vol. 436, p. 226 838, Oct. 2019, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2019.226838.
- [81] D. Montinaro, A. R. Contino, A. Dellai, and M. Rolland, “Determination of the impedance contributions in anode supported solid oxide fuel cells with (La,Sr)(Co,Fe)O₃– cathode,” *International Journal of Hydrogen Energy*, vol. 39, no. 36, pp. 21 638–21 646, Dec. 2014, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2014.09.081.
- [82] M. Kornely, N. H. Menzler, A. Weber, and E. Ivers-Tiffée, “Degradation of a High Performance SOFC Cathode by Cr-Poisoning at OCV-Conditions,” en, *Fuel Cells*, vol. 13, no. 4, pp. 506–510, 2013, ISSN: 1615-6854. DOI: 10.1002/fuce.201200182.

- [83] Z. Willemin *et al.*, “Local Activation and Degradation of Electrochemical Processes in a SOFC,” en, *ECS Transactions*, vol. 57, no. 1, p. 561, Oct. 2013, ISSN: 1938-5862. DOI: 10.1149/05701.0561ecst.
- [84] A. Weber, J. Szász, S. Dierickx, C. Endler-Schuck, and E. Ivers-Tiffée, “Accelerated Lifetime Tests for SOFCs,” en, *ECS Transactions*, vol. 68, no. 1, p. 1953, Jun. 2015, ISSN: 1938-5862. DOI: 10.1149/06801.1953ecst.
- [85] J. R. Wilson, D. T. Schwartz, and S. B. Adler, “Nonlinear electrochemical impedance spectroscopy for solid oxide fuel cell cathode materials,” *Electrochimica Acta*, Electrochemical Impedance Spectroscopy, vol. 51, no. 8, pp. 1389–1402, Jan. 2006, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2005.02.109.
- [86] S. B. Adler, “Applications of Nonlinear Electrochemical Impedance Spectroscopy (NLEIS),” en, *ECS Transactions*, vol. 58, no. 3, p. 101, Aug. 2013, ISSN: 1938-5862. DOI: 10.1149/05803.0101ecst.
- [87] J. R. Wilson, M. Sase, T. Kawada, and S. B. Adler, “Measurement of Oxygen Exchange Kinetics on Thin-Film $\text{La}_{0.6}\text{Sr}_{0.4}\text{CoO}_3$ — Using Nonlinear Electrochemical Impedance Spectroscopy,” en, *Electrochemical and Solid-State Letters*, vol. 10, no. 5, B81, Feb. 2007, ISSN: 1944-8775. DOI: 10.1149/1.2710178.
- [88] C. R. Kreller and S. Adler, “Insight into Rate Limiting Mechanisms of Porous $\text{La}_{1-x}\text{Sr}_x\text{CoO}_3$ Electrodes through Nonlinear Electrochemical Impedance Spectroscopy (NLEIS),” en, *ECS Transactions*, vol. 25, no. 2, p. 2743, Sep. 2009, ISSN: 1938-5862. DOI: 10.1149/1.3205835.
- [89] V. Subotić, S. Pofahl, V. Lawlor, N. H. Menzler, T. Thaller, and C. Hochenauer, “Online monitoring tools for SoH diagnostic and prognostic of remaining lifetime of reversible solid oxide cell (rSOC) systems,” *Energy Procedia*, Innovative Solutions for Energy Transitions, vol. 158, pp. 2329–2334, Feb. 2019, ISSN: 1876-6102. DOI: 10.1016/j.egypro.2019.01.271.
- [90] J. Mougín *et al.*, “Monitoring and Diagnostics of SOFC Stacks and Systems,” en, *ECS Transactions*, vol. 91, no. 1, p. 731, Jul. 2019, ISSN: 1938-5862. DOI: 10.1149/09101.0731ecst.
- [91] B. Tu *et al.*, “Thermodynamic analysis and experimental study of electrode reactions and open circuit voltages for methane-fuelled SOFC,” *International Journal of Hydrogen Energy*, vol. 45, no. 58, pp. 34069–34079, Nov. 2020, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2020.09.088.

- [92] M. M. Welander, M. S. Zachariasen, C. D. Hunt, S. W. Sofie, and R. A. Walker, “Operando Studies of Redox Resilience in ALT Enhanced NiO-YSZ SOFC Anodes,” en, *Journal of The Electrochemical Society*, vol. 165, no. 3, F152, Feb. 2018, ISSN: 1945-7111. DOI: 10.1149/2.0531803jes.
- [93] M. Preininger, B. Stoeckl, V. Subotić, and C. Hochenauer, “Characterization and performance study of commercially available solid oxide cell stacks for an autonomous system,” *Energy Conversion and Management*, vol. 203, p. 112 215, Jan. 2020, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2019.112215.
- [94] M. Höber *et al.*, “Holistic Approach to Design, Test, and Optimize Stand-Alone SOFC-Reformer Systems,” en, *Processes*, vol. 9, no. 2, p. 348, Feb. 2021, ISSN: 2227-9717. DOI: 10.3390/pr9020348.
- [95] V. Subotić, C. Schluckner, B. Stöckl, V. Lawlor, H. Schroettner, and C. Hochenauer, “Strategy for Carbon Gasification from Porous Ni-YSZ Anodes of Industrial-Sized ASC-SOFCs and Effects of Carbon Growth,” en, *Journal of The Electrochemical Society*, vol. 163, no. 14, F1515, Nov. 2016, ISSN: 1945-7111. DOI: 10.1149/2.0521614jes.
- [96] V. Subotić, C. Schluckner, and C. Hochenauer, “An experimental and numerical study of performance of large planar ESC-SOFCs and experimental investigation of carbon depositions,” *Journal of the Energy Institute*, vol. 89, no. 1, pp. 121–137, Feb. 2016, ISSN: 1743-9671. DOI: 10.1016/j.joei.2015.01.004.
- [97] J. Sehested, J. A. P. Gelten, and S. Helveg, “Sintering of nickel catalysts: Effects of time, atmosphere, temperature, nickel-carrier interactions, and dopants,” *Applied Catalysis A: General*, vol. 309, no. 2, pp. 237–246, Aug. 2006, ISSN: 0926-860X. DOI: 10.1016/j.apcata.2006.05.017.
- [98] W. Ming *et al.*, “A systematic review of machine learning methods applied to fuel cells in performance evaluation, durability prediction, and application monitoring,” *International Journal of Hydrogen Energy*, vol. 48, no. 13, pp. 5197–5228, Feb. 2023, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2022.10.261.
- [99] M. Ananyev, A. Gavriluk, D. Bronin, R. Steinberger-Wilckens, and J. Mertens, *SOFC Degradation Quantification Using Image Analysis*. Jun. 2011.
- [100] D. Simwonis, F. Tietz, and D. Stöver, “Nickel coarsening in annealed Ni/8YSZ anode substrates for solid oxide fuel cells,” *Solid State Ionics*, Solid Oxide Fuel Cells dedicated to Prof. H. Tagawa, vol. 132, no. 3, pp. 241–251, Jul. 2000, ISSN: 0167-2738. DOI: 10.1016/S0167-2738(00)00650-0.

- [101] D. Yan *et al.*, “Degradation analysis and durability improvement for SOFC 1-cell stack,” *Applied Energy*, vol. 175, pp. 414–420, Aug. 2016, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2016.04.094.
- [102] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, and Đ. Juričić, “State of health estimation and remaining useful life prediction of solid oxide fuel cell stack,” *Energy Conversion and Management*, vol. 148, pp. 993–1002, Sep. 2017, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2017.06.041.
- [103] R. S. Gemmen, M. C. Williams, and K. Gerdes, “Degradation measurement and analysis for cells and stacks,” *Journal of Power Sources*, vol. 184, no. 1, pp. 251–259, Sep. 2008, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2008.06.047.
- [104] J. Mason, I. Celik, S. Lee, H. Abernathy, and G. Hackett, “Performance degradation predictions based on microstructural evolution due to grain coarsening effects in solid oxide fuel cell electrodes,” *Journal of The Electrochemical Society*, vol. 165, no. 2, F64–F74, 2018. DOI: 10.1149/2.0721802jes.
- [105] Q. Fu, Z. Li, W. Wei, F. Liu, X. Xu, and Z. Liu, “Performance degradation prediction of direct internal reforming solid oxide fuel cell due to Ni-particle coarsening in composite anode,” *Energy Conversion and Management*, vol. 233, p. 113 902, Apr. 2021, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2021.113902.
- [106] P. Zhu *et al.*, “Ni coarsening and performance attenuation prediction of biomass syngas fueled SOFC by combining multi-physics field modeling and artificial neural network,” *Applied Energy*, vol. 322, p. 119 508, Sep. 2022, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2022.119508.
- [107] V. Zaccaria, D. Tucker, and A. Traverso, “A distributed real-time model of degradation in a solid oxide fuel cell, part I: Model characterization,” *Journal of Power Sources*, vol. 311, pp. 175–181, Apr. 2016, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2016.02.040.
- [108] A. Nakajo, F. Mueller, D. McLarty, J. Brouwer, J. V. Herle, and D. Favrat, “The effects of dynamic dispatch on the degradation and lifetime of solid oxide fuel cell systems,” *ECS Transactions*, vol. 35, no. 1, pp. 285–296, Apr. 2011. DOI: 10.1149/1.3570004.
- [109] M. Z. Khan, M. T. Mehran, R.-H. Song, J.-W. Lee, S.-B. Lee, and T.-H. Lim, “A simplified approach to predict performance degradation of a solid oxide fuel cell anode,” *Journal of Power Sources*, vol. 391, pp. 94–105, Jul. 2018, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2018.04.080.

- [110] X. Fang and Z. Lin, “Numerical study on the mechanical stress and mechanical failure of planar solid oxide fuel cell,” *Applied Energy*, vol. 229, pp. 63–68, Nov. 2018, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.07.077.
- [111] H. Sezer, J. H. Mason, I. B. Celik, and T. Yang, “Three-dimensional modeling of performance degradation of planar SOFC with phosphine exposure,” *International Journal of Hydrogen Energy*, vol. 46, no. 9, pp. 6803–6816, Feb. 2021, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2020.11.176.
- [112] E. Hardjo, D. S. Monder, and K. Karan, “Numerical modeling of nickel-impregnated porous YSZ-supported anodes and comparison to conventional composite ni-YSZ electrodes,” *ECS Transactions*, vol. 35, no. 1, pp. 1823–1832, Apr. 2011. DOI: 10.1149/1.3570171.
- [113] A. Nakajo, F. Mueller, J. Brouwer, J. Van herle, and D. Favrat, “Mechanical reliability and durability of SOFC stacks. Part I : Modelling of the effect of operating conditions and design alternatives on the reliability,” *International Journal of Hydrogen Energy*, vol. 37, no. 11, pp. 9249–9268, Jun. 2012, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2012.03.043.
- [114] M. Rao *et al.*, “Data-Driven State Prediction and Analysis of SOFC System Based on Deep Learning Method,” *en, Energies*, vol. 15, no. 9, p. 3099, Jan. 2022, ISSN: 1996-1073. DOI: 10.3390/en15093099.
- [115] X. Wu and Q. Ye, “Fault diagnosis and prognostic of solid oxide fuel cells,” *Journal of Power Sources*, vol. 321, pp. 47–56, Jul. 2016, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2016.04.080.
- [116] X. Wu, Q. Ye, and J. Wang, “A hybrid prognostic model applied to SOFC prognostics,” *International Journal of Hydrogen Energy*, vol. 42, no. 39, pp. 25 008–25 020, Sep. 2017, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2017.08.114.
- [117] M. Li *et al.*, “Data-Driven Voltage Prognostic for Solid Oxide Fuel Cell System Based on Deep Learning,” *en, Energies*, vol. 15, no. 17, p. 6294, Jan. 2022, ISSN: 1996-1073. DOI: 10.3390/en15176294.
- [118] D. Marra, M. Sorrentino, C. Pianese, and B. Iwanschitz, “A neural network estimator of Solid Oxide Fuel Cell performance for on-field diagnostics and prognostics applications,” *Journal of Power Sources*, vol. 241, pp. 320–329, Nov. 2013, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2013.04.114.
- [119] S. Sonderskov, J. Ilsoe, D. Rasmussen, D. Blom-Hansen, and S. Munk-Nielsen, “State of Health Estimation and Prediction of Fuel Cell Stacks in Backup Power Systems,” *Engrxiv*, 2019. DOI: 10.31224/osf.io/6b75s.

- [120] S.-J. Cheng, W.-K. Li, T.-J. Chang, and C.-H. Hsu, "Data-Driven Prognostics of the SOFC System Based on Dynamic Neural Network Models," en, *Energies*, vol. 14, no. 18, p. 5841, Jan. 2021, ISSN: 1996-1073. DOI: 10.3390/en14185841.
- [121] Y. Xu *et al.*, "Real-Time State of Health Estimation for Solid Oxide Fuel Cells Based on Unscented Kalman Filter," en, *Energies*, vol. 15, no. 7, p. 2534, Jan. 2022, ISSN: 1996-1073. DOI: 10.3390/en15072534.
- [122] K. Yang *et al.*, "Machine-learning-assisted prediction of long-term performance degradation on solid oxide fuel cell cathodes induced by chromium poisoning," *Journal of Materials Chemistry A*, vol. 10, no. 44, pp. 23 683–23 690, 2022. DOI: 10.1039/d2ta03944c.
- [123] C. B. Muñoz, D. Pumiglia, F. Santoni, S. J. McPhail, G. Comodi, and M. Carlini, "Performance degradation prediction of a low-temperature SOFC via impedance spectroscopy and CFD modelling," *ECS Transactions*, vol. 68, no. 1, pp. 2227–2235, Jun. 2015. DOI: 10.1149/06801.2227ecst.
- [124] J. Gazzarri and O. Kesler, "Non-destructive delamination detection in solid oxide fuel cells," *Journal of Power Sources*, vol. 167, no. 2, pp. 430–441, May 2007. DOI: 10.1016/j.jpowsour.2007.02.042.
- [125] X. Wu, L. Xu, J. Wang, D. Yang, F. Li, and X. Li, "A prognostic-based dynamic optimization strategy for a degraded solid oxide fuel cell," *Sustainable Energy Technologies and Assessments*, vol. 39, p. 100 682, Jun. 2020, ISSN: 2213-1388. DOI: 10.1016/j.seta.2020.100682.
- [126] C. Sheng *et al.*, "A Comparative Study of the Kalman Filter and the LSTM Network for the Remaining Useful Life Prediction of SOFC," en, *Energies*, vol. 16, no. 9, p. 3628, Jan. 2023, ISSN: 1996-1073. DOI: 10.3390/en16093628.
- [127] Y. Shi *et al.*, "Hybrid neural network and dynamic decay model for life prediction of solid oxide fuel cell combined heat and power systems," *Journal of Power Sources*, vol. 645, p. 237 165, 2025, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2025.237165.
- [128] L. Žnidarič, Ž. Gradišar, and Đ. Juričić, "Predicting the remaining useful life of solid oxide fuel cell systems using adaptive trend models of health indicators," *Energies*, vol. 17, no. 11, 2024, ISSN: 1996-1073. DOI: 10.3390/en17112729.
- [129] T. Vairo, D. Cademartori, D. Clematis, M. P. Carpanese, and B. Fabiano, "Solid oxide fuel cells for shipping: A machine learning model for early detection of hazardous system deviations," *Process Safety and Environmental Protection*, vol. 172, pp. 184–194, Apr. 2023, ISSN: 0957-5820. DOI: 10.1016/j.psep.2023.02.022.

- [130] Y. Zu, J. Jians, and X. Li, "A hybrid prediction method for SOFC degradation prediction," in *2021 China Automation Congress (CAC)*, ISSN: 2688-0938, Oct. 2021, pp. 6532–6536. DOI: 10.1109/CAC53003.2021.9728595.
- [131] P. Costamagna, A. De Giorgi, L. Magistri, G. Moser, L. Pellaco, and A. Trucco, "A classification approach for model-based fault diagnosis in power generation systems based on solid oxide fuel cells," *IEEE Transactions on Energy Conversion*, vol. 31, no. 2, pp. 676–687, 2016. DOI: 10.1109/TEC.2015.2492938.
- [132] P. Polverino, C. Pianese, M. Sorrentino, and D. Marra, "Model-based development of a fault signature matrix to improve solid oxide fuel cell systems on-site diagnosis," *Journal of Power Sources*, vol. 280, pp. 320–338, 2015, ISSN: 0378-7753. DOI: <https://doi.org/10.1016/j.jpowsour.2015.01.037>.
- [133] G. T. Le, L. Mastropasqua, J. Brouwer, and S. B. Adler, "Simulation-informed machine learning diagnostics of solid oxide fuel cell stack with electrochemical impedance spectroscopy," *Journal of The Electrochemical Society*, vol. 169, no. 3, p. 034530, Mar. 2022. DOI: 10.1149/1945-7111/ac59f4.
- [134] K. Zhu *et al.*, "Research on fault diagnosis of solid oxide fuel cell based on improved random forest algorithm," in *Proceedings of the 43rd Chinese Control Conference (CCC)*, IEEE, 2024, pp. 4810–4815. DOI: 10.23919/CCC63176.2024.10662020.
- [135] S. Vafaeenezhad *et al.*, "Stability of infiltrated cathodes using Pr₂NiO₄+delta precursor for low-temperature fuel cell applications," *International Journal of Hydrogen Energy*, vol. 47, no. 82, pp. 35 081–35 095, Sep. 2022, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2022.08.108.
- [136] A. R. Razmi, S. Sharifi, S. Vafaeenezhad, A. R. Hanifi, and M. Shahbakhti, "Modeling and microstructural study of anode-supported solid oxide fuel cells: Experimental and thermodynamic analyses," *International Journal of Hydrogen Energy*, vol. 54, pp. 613–634, Feb. 2024, ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2023.08.296.
- [137] N. Q. Minh and M. B. Mogensen, "Reversible solid oxide fuel cell technology for green fuel and power production," *ChemInform*, vol. 46, 2015. [Online]. Available: <https://api.semanticscholar.org/CorpusID:101397724>.
- [138] *Report on the status of the solid oxide fuel cell program*, <https://www.energy.gov/>, Report to Congress August 2019.

- [139] T. H. Wan, M. Saccoccio, C. Chen, and F. Ciucci, "Influence of the Discretization Methods on the Distribution of Relaxation Times Deconvolution: Implementing Radial Basis Functions with DRTtools," *Electrochimica Acta*, vol. 184, pp. 483–499, Dec. 2015, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2015.09.097.
- [140] A. Weber, S. Dierickx, A. Kromp, and E. Ivers-Tiffée, "Sulfur poisoning of anode-supported sofc's under reformat operation," *Fuel Cells*, vol. 13, no. 4, pp. 487–493, 2013. DOI: <https://doi.org/10.1002/fuce.201200180>.
- [141] L. Pipia, J. Muñoz-Marí, E. Amin, S. Belda, G. Camps-Valls, and J. Verrelst, "Fusing optical and SAR time series for LAI gap filling with multioutput Gaussian processes," *Remote Sensing of Environment*, vol. 235, p. 111452, 2019, ISSN: 0034-4257. DOI: 10.1016/j.rse.2019.111452.
- [142] A. Mateo-Sanchis, J. Muñoz-Marí, M. Campos-Taberner, J. García-Haro, and G. Camps-Valls, "Gap Filling of Biophysical Parameter Time Series with Multi-Output Gaussian Processes," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, ISSN: 2153-7003, Jul. 2018, pp. 4039–4042. DOI: 10.1109/IGARSS.2018.8519254.
- [143] A. Verdone, S. Scardapane, and M. Panella, "Explainable Spatio-Temporal Graph Neural Networks for multi-site photovoltaic energy production," *Applied Energy*, vol. 353, p. 122151, Jan. 2024, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2023.122151.
- [144] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The Graph Neural Network Model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, Jan. 2009, ISSN: 1941-0093. DOI: 10.1109/TNN.2008.2005605.
- [145] W. L. Hamilton, "The Graph Neural Network Model," en, in W. L. Hamilton, Ed., Cham: Springer International Publishing, 2020, pp. 51–70, ISBN: 9783031015885. DOI: 10.1007/978-3-031-01588-5_5.
- [146] L. Zhao *et al.*, "T-GCN: A Temporal Graph ConvolutionalNetwork for Traffic Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020, ISSN: 1524-9050, 1558-0016. DOI: 10.1109/TITS.2019.2935152.
- [147] Z. Feng, R. Wang, T. Wang, M. Song, S. Wu, and S. He, "A comprehensive survey of dynamic graph neural networks: Models, frameworks, benchmarks, experiments and challenges," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–20, 2025. DOI: 10.1109/TKDE.2025.3621291.

- [148] Y. Yu, X. Si, C. Hu, and J. Zhang, “A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures,” *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019, ISSN: 0899-7667. DOI: 10.1162/neco_a_01199.
- [149] F. Wilcoxon, “Individual comparisons by ranking methods,” in *Breakthroughs in Statistics: Methodology and Distribution*, S. Kotz and N. L. Johnson, Eds. New York, NY: Springer New York, 1992, pp. 196–202, ISBN: 978-1-4612-4380-9. DOI: 10.1007/978-1-4612-4380-9_16.
- [150] H. Zhou *et al.*, *Informer: Beyond efficient transformer for long sequence time-series forecasting*, arXiv preprint, 2020. arXiv: 2012.07436 [cs.LG].
- [151] A. Vaswani *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, I. Guyon *et al.*, Eds., vol. 30, Curran Associates, Inc., 2017.
- [152] A. Mohammadi, M. Krysander, and D. Jung, “Consistency-based diagnosis using data-driven residuals and limited training data,” *Control Engineering Practice*, vol. 159, p. 106 283, 2025, ISSN: 0967-0661. DOI: 10.1016/j.conengprac.2025.106283.
- [153] E. M. Aiello, C. Toffanin, L. Magni, and G. De Nicolao, “Model-based identification of eating behavioral patterns in populations with type 1 diabetes,” *Control Engineering Practice*, vol. 123, p. 105 128, 2022, ISSN: 0967-0661. DOI: 10.1016/j.conengprac.2022.105128.
- [154] C. Chen, J. Shi, M. Shen, L. Feng, and G. Tao, “A predictive maintenance strategy using deep learning quantile regression and kernel density estimation for failure prediction,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023. DOI: 10.1109/TIM.2023.3240208.
- [155] Y. Yang, H. A. Khorshidi, and U. Aickelin, “A review on over-sampling techniques in classification of multi-class imbalanced datasets: Insights for medical problems,” *Frontiers in Digital Health*, vol. Volume 6 - 2024, 2024, ISSN: 2673-253X. DOI: 10.3389/fdgth.2024.1430245.
- [156] M. Mujahid, E. Kina, F. Rustam, *et al.*, “Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering,” *Journal of Big Data*, vol. 11, p. 87, 2024. DOI: 10.1186/s40537-024-00943-4.
- [157] R. A. Danquah, *Handling imbalanced data: A case study for binary class problems*, 2020. arXiv: 2010.04326 [stat.ML].
- [158] H. Hairani, T. Widiyaningtyas, and D. D. Prasetya, “Addressing class imbalance of health data: A systematic literature review on modified synthetic minority oversampling technique (SMOTE) strategies,” *International Journal on Informatics Visualization*, vol. 8, no. 3, pp. 1310–1318, 2024. DOI: 10.62527/joiv.8.3.2283.

- [159] Z. Xu, D. Shen, Y. Kou, and T. Nie, "A synthetic minority oversampling technique based on Gaussian mixture model filtering for imbalanced data classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 3, pp. 3740–3753, 2024. DOI: 10.1109/TNNLS.2022.3197156.
- [160] S. Limanto, J. L. Buliali, and A. Saikhu, "GLoW SMOTE-D: Oversampling technique to improve prediction model performance of students failure in courses," *IEEE Access*, vol. 12, pp. 8889–8901, 2024. DOI: 10.1109/ACCESS.2024.3351569.
- [161] N. M. Nayan, A. Islam, M. U. Islam, E. Ahmed, M. M. Hossain, and M. Z. Alam, "SMOTE oversampling and near miss undersampling based diabetes diagnosis from imbalanced dataset with XAI visualization," in *2023 IEEE Symposium on Computers and Communications (ISCC)*, 2023, pp. 1–6. DOI: 10.1109/ISCC58397.2023.10218281.
- [162] M. Mujahid *et al.*, "Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering," *Journal of Big Data*, vol. 11, no. 1, p. 87, 2024.
- [163] W. Chen, K. Yang, Z. Yu, *et al.*, "A survey on imbalanced learning: Latest research, applications and future directions," *Artificial Intelligence Review*, vol. 57, p. 137, 2024. DOI: 10.1007/s10462-024-10759-6.
- [164] J. Guo, H. Wu, X. Chen, and W. Lin, "Adaptive SV-Borderline SMOTE-SVM algorithm for imbalanced data classification," *Applied Soft Computing*, vol. 150, p. 110986, 2024, ISSN: 1568-4946. DOI: <https://doi.org/10.1016/j.asoc.2023.110986>.
- [165] M. Misdrum, E. Noersasongko, Purwanto, Muljono, and F. Y. Pamuji, "Gaussian based-SMOTE method for handling imbalanced small datasets," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 4, pp. 973–982, 2023, ISSN: 2338-3070. DOI: 10.26555/jiteki.v9i4.26881.
- [166] W. Chen, W. Guo, and W. Mao, "An adaptive over-sampling method for imbalanced data based on simultaneous clustering and filtering noisy," *Applied Intelligence*, vol. 54, no. 22, pp. 11430–11449, Aug. 2024, ISSN: 0924-669X. DOI: 10.1007/s10489-024-05754-x.
- [167] D. Lee and H. Kim, "Adaptive oversampling via density estimation for online imbalanced classification," *Information*, vol. 16, no. 1, 2025, ISSN: 2078-2489. DOI: 10.3390/info16010023.
- [168] Z. Zhang, H.-p. Tian, and J.-s. Jin, "Multiple adaptive over-sampling for imbalanced data evidential classification," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108532, 2024, ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2024.108532>.

- [169] T. Prexawanprasut and T. Banditwattanawong, "Improving minority class recall through a novel cluster-based oversampling technique," *Informatics*, vol. 11, no. 2, 2024, ISSN: 2227-9709. DOI: 10.3390/informatics11020035.
- [170] A. R. Salehi and M. Khedmati, "A cluster-based SMOTE both-sampling (csbboost) ensemble algorithm for classifying imbalanced data," *Scientific Reports*, vol. 14, p. 5152, 2024. DOI: 10.1038/s41598-024-55598-1.
- [171] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *The Stata Journal*, vol. 20, no. 1, pp. 3–29, 2020. DOI: 10.1177/1536867X20909688.
- [172] "Random forest algorithm overview," *Babylonian Journal of Machine Learning*, vol. 2024, pp. 69–79, Jun. 2024. DOI: 10.58496/BJML/2024/007.
- [173] X. Fu, Y. Chen, J. Yan, Y. Chen, and F. Xu, "Bgrf: A broad granular random forest algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 44, no. 5, pp. 8103–8117, 2023. DOI: 10.3233/JIFS-223960.
- [174] Y. Wu, Y. Wang, H. Liu, *et al.*, "Risk assessment of bridge construction investigated using random forest algorithm," *Scientific Reports*, vol. 14, p. 20964, 2024. DOI: 10.1038/s41598-024-72051-5.
- [175] C. N. Noviyanti and A. Alamsyah, "Early detection of diabetes using random forest algorithm," *Journal of Information System Exploration and Research*, vol. 2, no. 1, 2024. DOI: 10.52465/joiser.v2i1.245.
- [176] M. Vadivel, V. B. Marin, S. Balasubramani, S. Hemalatha, S. Murugan, and S. Velmurugan, "Cloud-based passenger experience management in bus fare ticketing systems using random forest algorithm," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2024, pp. 1–6. DOI: 10.1109/ICRITO61523.2024.10522226.
- [177] A. A. Reza and M. S. Rohman, "Prediction stunting analysis using random forest algorithm and random search optimization," *Journal of Information Technology and Education*, vol. 7, no. 2, 2024. DOI: 10.31289/jite.v7i2.10628.
- [178] R. Pavaiyarkarasi, T. Manimegalai, S. Satheeshkumar, K. Dhivya, and G. Ramkumar, "A productive feature selection criterion for bot-iot recognition based on random forest algorithm," in *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, 2022, pp. 539–545. DOI: 10.1109/CSNT54456.2022.9787583.
- [179] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

- [180] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’16, San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794, ISBN: 9781450342322. DOI: 10.1145/2939672.2939785. [Online]. Available: <https://doi.org/10.1145/2939672.2939785>.
- [181] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Frontiers in Neurorobotics*, vol. Volume 7 - 2013, 2013, ISSN: 1662-5218. DOI: 10.3389/fnbot.2013.00021.
- [182] T. Lekan, A. Harry, and H. Rajab, “Comparison of neural networks with traditional machine learning models (e.g., xgboost, random forest),” Nov. 2025.
- [183] P. Zhang, Y. Jia, and Y. Shang, “Research and application of xgboost in imbalanced data,” *International Journal of Distributed Sensor Networks*, vol. 18, no. 6, p. 15501329221106935, 2022. DOI: 10.1177/15501329221106935.
- [184] R. Natras, B. Soja, and M. Schmidt, “Ensemble machine learning of random forest, adaboost and xgboost for vertical total electron content forecasting,” *Remote Sensing*, vol. 14, no. 15, 2022, ISSN: 2072-4292. DOI: 10.3390/rs14153547.
- [185] S. Gündoğdu, “Efficient prediction of early-stage diabetes using xgboost classifier with random forest feature selection technique,” *Multimedia Tools and Applications*, vol. 82, no. 22, pp. 34163–34181, 2023. DOI: 10.1007/s11042-023-15165-8.
- [186] M. Shahhosseini, G. Hu, and H. Pham, “Optimizing ensemble weights and hyperparameters of machine learning models for regression problems,” *Machine Learning with Applications*, vol. 7, p. 100251, 2022, ISSN: 2666-8270. DOI: <https://doi.org/10.1016/j.mlwa.2022.100251>.
- [187] J. M. Ahn, J. Kim, and K. Kim, “Ensemble machine learning of gradient boosting (xgboost, lightgbm, catboost) and attention-based cnn-lstm for harmful algal blooms forecasting,” *Toxins*, vol. 15, no. 10, 2023, ISSN: 2072-6651. DOI: 10.3390/toxins15100608.
- [188] F. Karim, S. Majumdar, H. Darabi, and S. Harford, “Multivariate LSTM-FCNs for time series classification,” *Neural Networks*, vol. 116, pp. 237–245, 2019, ISSN: 0893-6080. DOI: <https://doi.org/10.1016/j.neunet.2019.04.014>.
- [189] F. Karim, S. Majumdar, H. Darabi, and S. Chen, “LSTM fully convolutional networks for time series classification,” *IEEE Access*, vol. 6, pp. 1662–1669, 2018. DOI: 10.1109/ACCESS.2017.2779939.

- [190] H. X. Tan, N. N. Aung, J. Tian, M. C. H. Chua, and Y. O. Yang, "Time series classification using a modified LSTM approach from accelerometer-based data: A comparative study for gait cycle detection," *Gait Posture*, vol. 74, pp. 128–134, 2019, ISSN: 0966-6362. DOI: <https://doi.org/10.1016/j.gaitpost.2019.09.007>.
- [191] R. DiPietro and G. D. Hager, "Chapter 21 - deep learning: RNNs and LSTM," in *Handbook of Medical Image Computing and Computer Assisted Intervention*, ser. The Elsevier and MICCAI Society Book Series, S. K. Zhou, D. Rueckert, and G. Fichtinger, Eds., Academic Press, 2020, pp. 503–519, ISBN: 978-0-12-816176-0. DOI: <https://doi.org/10.1016/B978-0-12-816176-0.00026-0>.
- [192] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017.
- [193] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *CoRR*, vol. abs/1803.01271, 2018. arXiv: 1803.01271.
- [194] Y. Liu, Z. Xu, Y. He, P. Guo, and K. Mu, "Acoustic fault diagnosis method for rotating machinery based on improved spectral subtraction and CNN-TCN model," *Measurement*, vol. 256, p. 118482, 2025, ISSN: 0263-2241. DOI: <https://doi.org/10.1016/j.measurement.2025.118482>.
- [195] P. Tang, P. Du, J. Xia, P. Zhang, and W. Zhang, "Channel attention-based temporal convolutional network for satellite image time series classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022. DOI: 10.1109/LGRS.2021.3095505.
- [196] A. Tarto, *Elcogen and Magnex signed LOI of SOFC commercialization*, en, Apr. 2020. [Online]. Available: <https://elcogen.com/elcogen-and-magnex-signed-loi-of-sofc-commercialization/> (visited on 04/17/2024).
- [197] K. Liu *et al.*, "Transfer learning for battery smarter state estimation and ageing prognostics: Recent progress, challenges, and prospects," *Advances in Applied Energy*, vol. 9, p. 100117, Feb. 2023, ISSN: 2666-7924. DOI: 10.1016/j.adapen.2022.100117.
- [198] E. Mladenova, M. Slavova, B. Abrashev, V. Terziev, B. Burdin, and G. Raikova, "Investigation of Ni- and Co-Based Bifunctional Electrocatalysts for Carbon-Free Air Electrodes Designed for Zinc-Air Batteries," en, *Emerging Science Journal*, vol. 7, no. 3, pp. 991–1003, May 2023, ISSN: 2610-9182. DOI: 10.28991/ESJ-2023-07-03-023. (visited on 10/27/2024).

- [199] I. Oyewole, A. Chehade, and Y. Kim, “A controllable deep transfer learning network with multiple domain adaptation for battery state-of-charge estimation,” *Applied Energy*, vol. 312, p. 118 726, Apr. 2022, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2022.118726. (visited on 10/26/2024).
- [200] J. Zhao and Z. Wang, “Specialized convolutional transformer networks for estimating battery health via transfer learning,” *Energy Storage Materials*, vol. 71, p. 103 668, Aug. 2024, ISSN: 2405-8297. DOI: 10.1016/j.ensm.2024.103668. (visited on 10/26/2024).
- [201] L. Chen *et al.*, “Fault diagnosis for cross-building energy systems based on transfer learning and model interpretation,” *Journal of Building Engineering*, vol. 91, p. 109 424, Aug. 2024, ISSN: 2352-7102. DOI: 10.1016/j.jobe.2024.109424. (visited on 10/26/2024).
- [202] G. Pinto, Z. Wang, A. Roy, T. Hong, and A. Capozzoli, “Transfer learning for smart buildings: A critical review of algorithms, applications, and future perspectives,” *Advances in Applied Energy*, vol. 5, p. 100 084, Feb. 2022, ISSN: 2666-7924. DOI: 10.1016/j.adapen.2022.100084. (visited on 10/26/2024).
- [203] X. Zhan, S. Han, N. Rong, and Y. Cao, “A hybrid transfer learning method for transient stability prediction considering sample imbalance,” *Applied Energy*, vol. 333, p. 120 573, Mar. 2023, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2022.120573. (visited on 10/31/2024).
- [204] Y. Gao, Z. Hu, S. Shi, W.-A. Chen, and M. Liu, “Adversarial discriminative domain adaptation for solar radiation prediction: A cross-regional study for zero-label transfer learning in Japan,” *Applied Energy*, vol. 359, p. 122 685, Apr. 2024, ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2024.122685. (visited on 10/31/2024).
- [205] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010. DOI: 10.1109/TKDE.2009.191.
- [206] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big Data*, vol. 3, no. 1, May 2016. DOI: 10.1186/s40537-016-0043-6.
- [207] F. Zhuang *et al.*, “A comprehensive survey on transfer learning,” *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021. DOI: 10.1109/JPROC.2020.3004555.
- [208] M. Iman, H. R. Arabnia, and K. Rasheed, “A Review of Deep Transfer Learning and Recent Advancements,” in *Technologies*, vol. 11, no. 2, p. 40, Apr. 2023, ISSN: 2227-7080. DOI: 10.3390/technologies11020040.

- [209] S. Song, X. Xiong, X. Wu, and Z. Xue, “Modeling the SOFC by BP neural network algorithm,” *International Journal of Hydrogen Energy*, vol. 46, no. 38, pp. 20 065–20 077, Jun. 2021. DOI: 10.1016/j.ijhydene.2021.03.132.
- [210] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, “Normalization Techniques in Training DNNs: Methodology, Analysis and Application,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 8, pp. 10 173–10 196, Aug. 2023, ISSN: 1939-3539. DOI: 10.1109/TPAMI.2023.3250241. (visited on 10/30/2024).

Appendix A

Transfer Learning for Cross-Domain SOFC Modeling¹

This appendix presents a complementary study on the application of transfer learning (TL) for modeling solid oxide fuel cells (SOFCs) across different cell types and operating domains. While the main chapters of this thesis focus on diagnostics and prognostics using degradation data, this work addresses a related challenge: reducing data and computational requirements when developing predictive models for new SOFC systems.

In particular, the study investigates the transferability of a deep neural network (DNN) model trained on a source tubular SOFC to multiple target domains, including other tubular cells with varying microstructural properties and a planar SOFC with a different physical configuration. The approach leverages fine-tuning techniques to adapt the pretrained model using a limited amount of target-domain data, and systematically evaluates the impact of normalization strategies and data availability on model performance.

The results demonstrate that transfer learning can significantly reduce both training time and data requirements while maintaining high predictive accuracy in current density-voltage (J-V) behavior. This work highlights the potential of TL as an efficient modeling strategy for heterogeneous SOFC systems and provides a foundation for extending data-driven approaches to new cell designs and operating conditions.

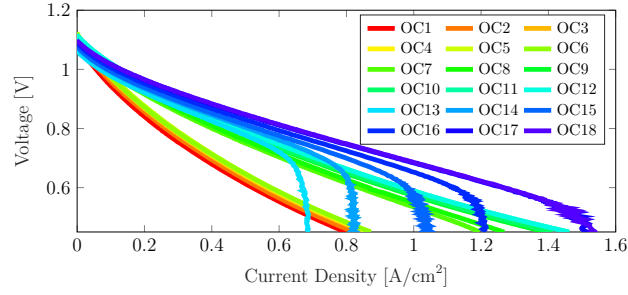
¹This appendix chapter is partially based on the following publication: [6, 7].

A.1 Tubular Cell Healthy Datasets

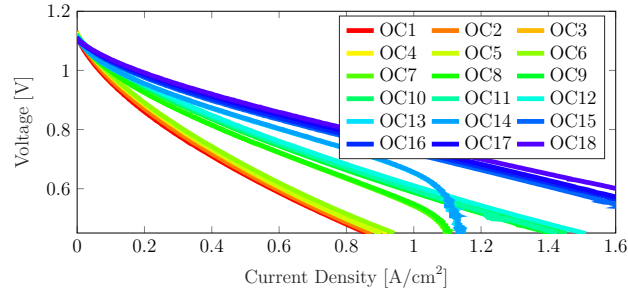
The purpose of the experiment is to characterize the tubular cells so that the resulting dataset can be used to train a neural network model and analyze SOCF performance under different operation conditions. The experiment was conducted by changing the value of the following variables: cell temperature (T_c), hydrogen flow rate ($\dot{V}_{H_2}$), nitrogen flow rate ($\dot{V}_{N_2}$), air flow rate ($\dot{V}_a$), and current density (J). The operating parameters were varied as follows: (1) T_c , ranging from 650 to 750°C in 3 levels, (2) $\dot{V}_{H_2}$, ranging from 20 to 50 SCCM in 3 levels, (3) $\dot{V}_{N_2}$, ranging from 0 to 30 SCCM in 3 levels, (4) $\dot{V}_{Air}$, ranging from 50 to 100 SCCM in 2 levels, and (5) J , ranging from 0 to 1.6 A/cm² with the step of 4.8 mA/cm² resulting in 330 points in each condition. Since the hydrogen outlet gas is inside the water bath, the total pressure of H₂ + N₂ is always close to atmospheric pressure. Here, the partial pressure of the H₂ is a function of its H₂:N₂ ratio. Diluting the fuel mixture reduces H₂ partial pressure. N₂ was added to maintain the total pressure to reduce uncertainty of the experiment. The combined hydrogen and nitrogen flow rates was 50 SCCM, so by setting the hydrogen flow rate to 20, 35, and 50 SCCM, the corresponding nitrogen flow rate is 30, 15, and 0 SCCM. Air flow rate and temperature are changed in 2 and 3 levels, respectively, providing 18 (3×3×2) combinations for data collection. As the current density is changed in 330 levels, this results in 5940 (330×18) data samples for each tubular cell and a combined dataset of 17820 (5940×3) for cells A1, A2, and A3. The measured J-V curves for cells A1, A2, and A3 are shown in Figures A.1a, A.1b, and A.1c respectively and represent a comprehensive database of these fuel cells operation at a steady state.

A.2 Planar Cell Healthy Datasets

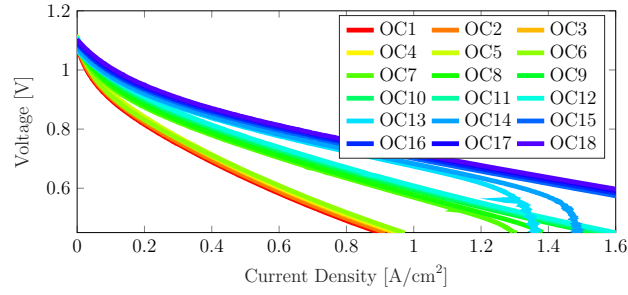
The SOFC stack from Magnex Co. Ltd is combined with anode-supported planar fuel cells fabricated by Elcogen [196]. A single planar cell was assembled using the interconnects, sealing materials, end plates, current collectors, and



(a) Cell A1



(b) Cell A2



(c) Cell A3

Figure A.1: J–V curves of three SOFC cells under 18 operating conditions. Cell temperature (T_c), hydrogen ($\dot{V}_{H_2}$), nitrogen ($\dot{V}_{N_2}$), and air ($\dot{V}_a$) flow rates are varied.

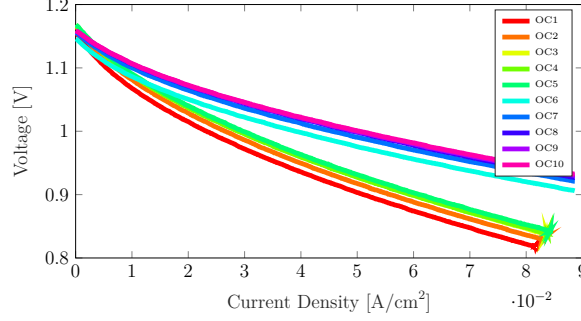


Figure A.2: J-V Curves of the planar cell at 10 Operating Conditions. Cell Temperature (T_c), hydrogen ($\dot{V}_{H_2}$), nitrogen ($\dot{V}_{N_2}$), and Air ($\dot{V}_a$) flow rate conditions are changed.

ceramic springs supplied by Magnex Co. Similar to the tubular cells, steady-state tests were conducted by changing cell temperature (T_c), hydrogen flow rate ($\dot{V}_{H_2}$), nitrogen flow rate ($\dot{V}_{N_2}$), air flow rate ($\dot{V}_a$), and current density (J). J-V curves are recorded for all 10 operating conditions and labeled from OC1 to OC10. The results for the planar cell are shown in Figure A.2. As the current density were changed in 600 levels, so each curve has 600 points, resulting in 6000 (600×10) data samples for the planar cell. The non-linear part of the curve was not measured due to experimental setup limitations.

A.3 DNN Model Development

It was found that a shallow tree-layer network does not capture the voltage drop in J-V curves. So, a DNN model is used to predict the output voltage of Cell A1. The inputs of this model are cell temperature (T_c), air flow rate ($\dot{V}_a$), hydrogen flow rate ($\dot{V}_{H_2}$), and current density (J), and the output of this model is the cell's voltage (V). Nitrogen flow rate depends on the hydrogen flow rate so it is not used. The model consists of 6 hidden layers with 100, 80, 60, 40, 20, and 10 neurons respectively, each utilizing the ReLU activation function. From the 5940 data samples are collected for cell A1 under different operating conditions, 75% of the samples are randomly separated to train the model, and the remaining 25% of samples are used for testing. The numbers of data samples in training and test datasets are 4455 and 1485, respectively.

A validation set from 10% of the training set randomly selected is used to evaluate the generalization capability of the model.

Determining the DNN hyper-parameters is related to the learning process (including optimizer, learning rate, number of epochs, and batch size) and the network architecture (including the number of hidden layers, number of neurons, and activation functions). These parameters are determined manually, and the implementation is done using the **Keras** library in Python.

The network architecture hyper-parameters are: Number of hidden layers: 6; Number of neurons: 4/100/80/60/40/20/10/1; Activation function: ReLU. The training process hyper-parameters are: Optimizer: Adam; Loss function: mean squared error (MSE); Training and test sets: 75/25 % (randomly separated); Validation set: 10% of training set (randomly separated); Learning rate: 0.001 (default of **Keras** for adam optimizer); Number of epochs: 50.

A.4 Transfer Learning

As there are 3 tubular cells with different physical parameters and one planar cell, the effect of design parameters on the cell voltage can be characterized. One approach is to train a DNN for each cell. However, experimental data is often expensive and scarce so Transfer Learning (TL) can be used [197]. TL techniques have the potential to be applied on the other systems beyond SOFCs such as batteries [198–200], smart buildings [201, 202] and others [203, 204] to reduce data requirements and computational effort. TL aims to improve the performance of a target learner exploiting knowledge in related environments. TL is most effective when there are some similarities between the source and target domains, such as shared features, patterns, or tasks [205]. Using TL, the knowledge gained from a source domain, where the data samples characterize the system, can be leveraged to a target domain with limited data points. This can enhance the learning process for the target domain since the model is built using the knowledge that comes from the pre-trained model in the source domain.

In this work, the input features of the source and target domains are the same, and the data distributions in the source and target domains are equal. A simple and widely used Parameter-based Transfer Learning (PBTL) technique is pre-training the network in the source domain and fine-tuning the weights for the target domain [206, 207]. PBTL can be applied to other areas of energy systems research, but in some cases, the features for the source and target domains are different, or the range of target features is a subset of the source features range. In those cases, the other TL strategies such as feature-based and instance-based TL can be applied [205, 208].

The PBTL process is shown schematically in Figure A.3. The DNN is first trained using the data samples from cell A1 and then the trained model is fine-tuned using a limited amount of the target cells (A2, A3, and planar) data. This fine-tuned model is then tested using the remaining portion of the target domain data, considered unknown data and not used for the fine-tuning process. For the normalization of the data before training, these three strate-

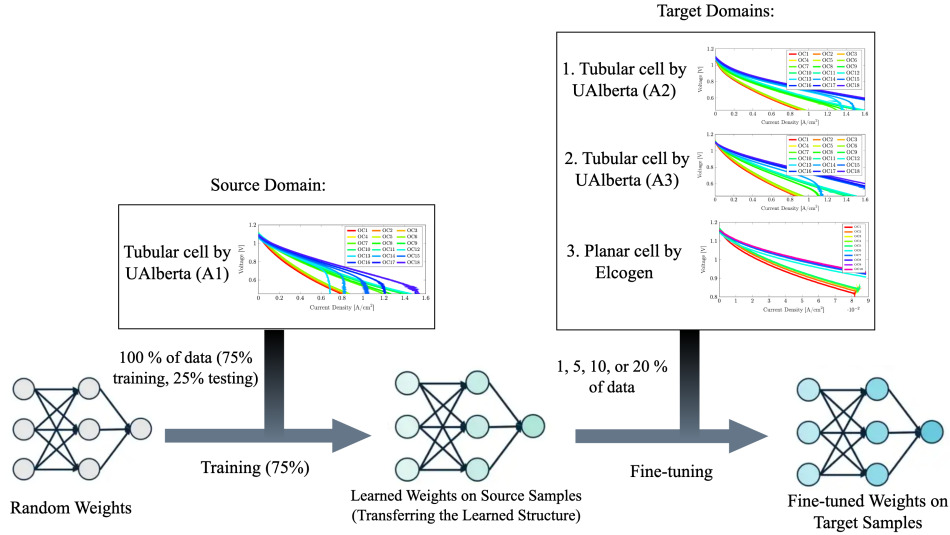


Figure A.3: Transfer Learning process from A1 to A2, A3, and planar cells. After training the DNN using source domain data, the trained network transfers to target domain and its weights are fine-tuned using the 1, 5, 10, and 20% of the target domain data.

gies are tested:

1. Combining source and target data and then normalizing

2. Normalizing the source data, then using that formula for target data
3. Normalizing the source and target data separately

Target data amounts of 1, 5, 10, and 20% for fine-tuning were selected to see how this influenced the TL.

A.5 Results and Discussion

The test results for cell A1 are RMSE=0.0001 V, MAPE= 0.83%, and $R^2=0.997$. The prediction accuracy with RMSE of 0.0001 V is satisfactory for both training and testing datasets, compared to similar predictions [209]. A comparison of the actual and predicted output voltage is shown in Figure A.4 for test dataset. The DNN model's predicted voltage closely aligns with the actual voltage values over the entire measured range. The training and validation losses during the training process are shown in Figure A.5, where the training and validation losses approach zero, indicating no over-fitting in the training process.

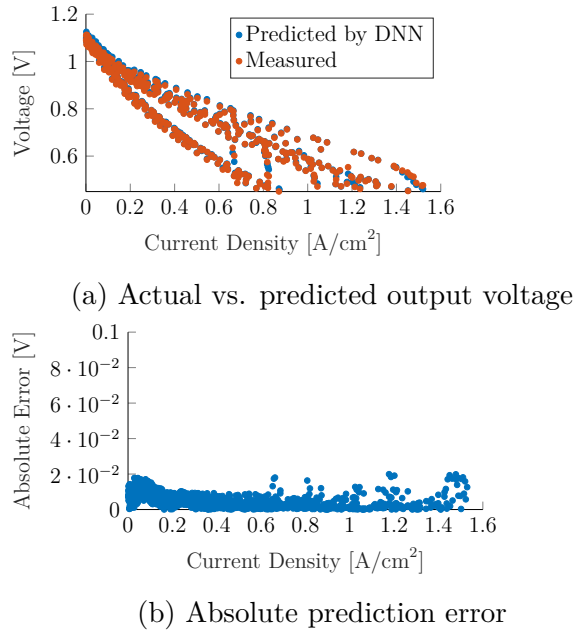


Figure A.4: A1: Actual vs. predicted output voltage and prediction error using DNN for test dataset. DNN model contains one input layer, 6 hidden layers, and one output layer with the number of neurons of 4/100/80/60/40/20/10/1. Experimental data corresponds to Figure A.1a.

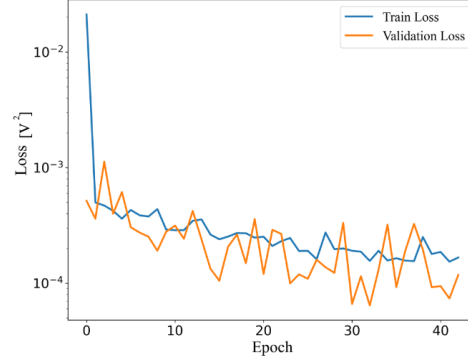


Figure A.5: Cell A1: Training and validation losses during the training process of DNN

The results of the TL process for cells A2 and A3 are listed in Table A.1. As expected, the errors are larger for A3 since cell A3 has a larger physical variation to A1 than A2 to A1. Three different normalization strategies, mentioned in the previous section, are compared in this table. By changing the normalization method, the prediction accuracy does not change substantially. However, strategy 2 achieves better metrics (Equations ?? - ??) than strategies 1 and 3. Transfer learning benefits from the assumption that the target domain shares some characteristics with the source domain. By using the same normalization parameters (mean, standard deviation, etc.) for the target, you're likely maintaining the relationships learned in the source data, which can help the model generalize better to the target domain [210]. Normalization the source data and applying that formula for target data (method 2) is used for the remaining tests.

Next, the effect of the percentage of data for fine-tuning in the target domain is evaluated. Cell A2 and A3 results are reported in Table A.2. The TL results with 20% for A3 are almost the same as those with 10% for A2. This is attributed to the fact that A3 is physically more different than A2 compared to A1. To obtain the MAPE under 1%, and R^2 above 0.999 for both cells, 10% of the data is needed. This results in fine-tuning computation time of 5 seconds. Using the 10% TL network, the prediction results on the other 90% of data for cells A2 and A3 are shown in Figures A.6 and A.7,

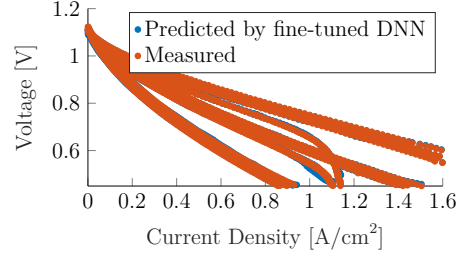
Table A.1: Prediction accuracy of model on future data of target domains (A2 and A3) with transfer learning. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.

Cell	Normal- ization strategy	RMSE (V) $\times 10^3$	MAPE (%)	R^2	Training (A1)/Fine-tuning (A2) Time (50 epochs)
A2	1	0.10	0.593	0.998	40s \ 5s
A2	2	0.00	0.443	0.999	40s \ 5s
A2	3	0.00	0.488	0.999	40s \ 5s
A3	1	0.10	1.008	0.997	40s \ 5s
A3	2	0.10	0.828	0.998	40s \ 5s
A3	3	0.10	0.870	0.997	40s \ 5s

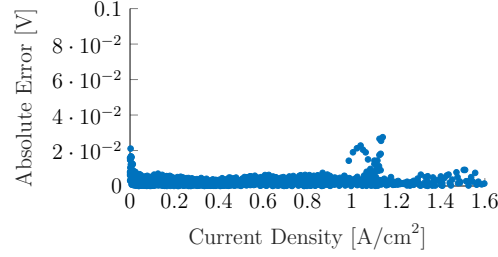
respectively. The decision to use 10% of the target domain data for fine-tuning was based on testing 1%, 5%, 10%, and 20%. It was found 10% provides an optimal balance between accuracy and fine-tuning data requirement. Smaller percentages (e.g., 1%) had reduced accuracy, while larger percentages (e.g., 20%) offered only marginal improvements in performance. This threshold is expected to vary based on the complexity of the target domain. More complex SOFC systems or those with greater differences in design parameters between source and target domains might require a larger portion of the target data to achieve comparable accuracy.

Table A.2: Different percentage of data in target domain for TL fine-tuning for cells A2 and A3, normalization 2. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.

Cell	Fine-tuning data percentage	RMSE (V) $\times 10^3$	MAPE (%)	R^2	Train- ing/Fine- tuning Time (50 epochs)
A2	1%	0.80	2.261	0.978	40s \ 1s
A2	5%	0.10	1.000	0.996	40s \ 3s
A2	10%	0.00	0.443	0.999	40s \ 5s
A2	20%	0.00	0.353	0.999	40s \ 10s
A3	1%	1.00	2.011	0.987	40s \ 1s
A3	5%	0.10	1.091	0.997	40s \ 3s
A3	10%	0.10	0.828	0.997	40s \ 5s
A3	20%	0.00	0.411	0.999	40s \ 10s



(a) Actual vs. predicted output voltage



(b) Absolute prediction error

Figure A.6: Results of transfer learning from cell A1 to cell A2 (90% untrained)

A planar SOFC is the third target domain for which TL is used. This cell is physically quite different from cell A1 and has substantially different operating conditions. The transfer learning process developed above uses 10% of the available data for fine-tuning. The performance of the transferred model on the test dataset in the target domain is listed in Table A.3 and shown in Figure A.8. The transferred model works well in the target domain, predicting the output voltage with a high accuracy of RMSE=0.0001 V. However, the planar cell was only tested in the J-V activation and ohmic loss areas.

Table A.3: Results of transfer learning from tubular cell A1 to planar cell using normalization 2 and 10% of fine-tuning data. Training/fine-tuning time on a MacOS 8-Core CPU processor is reported.

Planar cell	RMSE (V) $\times 10^3$	MAPE (%)	R ²	Training (A1)/Fine-tuning (Planar) Time (50 epochs)
10% fine-tuning data, normalization 2	0.10	0.371	0.996	40s \ 10s

Our findings address the limitations of previous studies discussed in the

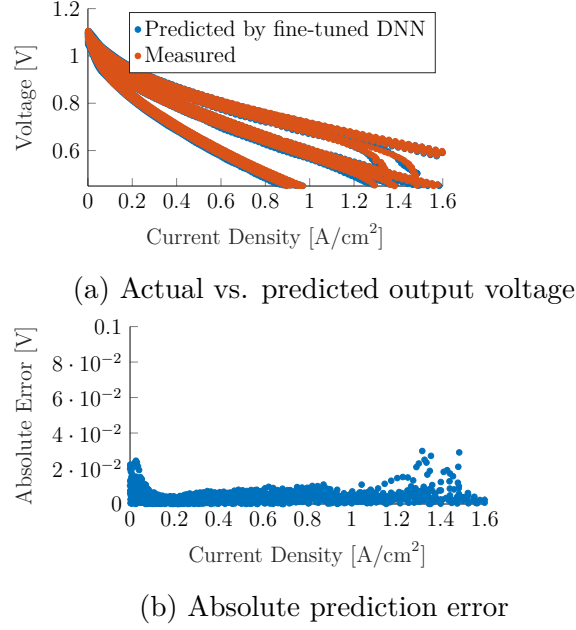
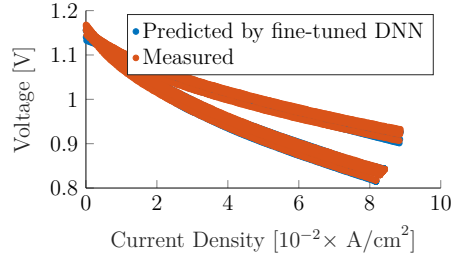


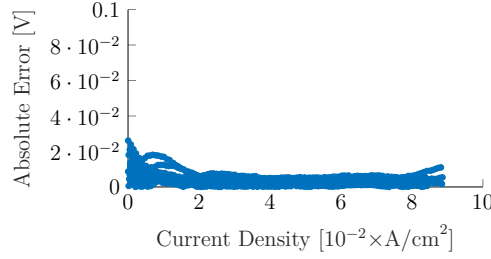
Figure A.7: Results of transfer learning from cell A1 to cell A3 (90% untrained)

literature review. Previous works explored the use of machine learning models in SOFC performance prediction but did not apply TL, leaving a gap in how these models could be adapted for predicting SOFC performance across different cell types and operating conditions with limited data. The results of this study demonstrate the value of TL in addressing this key research gap identified in the introduction. By fine-tuning a DNN model using only 10% of the target domains (cells A2, A3, and planar) data, we achieved predictive accuracy comparable to model trained on the full dataset (cell A1), reducing the training time by 85% and data requirement by 90%. This finding highlights the scalability of the TL approach, as it allows the model to work effectively across different cells or configurations, even when data collection is restricted. This contribution is particularly important given the high costs associated with SOFC testing.

While the transferred DNN model showed high accuracy for the tested cell types, its generalization to other types of SOFCs or cells from different manufacturers could be influenced by several factors. Differences in material and microstructural properties, operational conditions, and manufacturing tech-



(a) Actual vs. predicted output voltage



(b) Absolute prediction error

Figure A.8: Results of transfer learning from tubular cell A1 to planar cell (90% untrained)

niques may introduce additional complexities. Transfer learning could mitigate some of these challenges by fine-tuning the model on smaller datasets from new cells, but significant domain differences may still limit generalization. Future work could explore the extent of this limitation.

A.6 Appendix Summary

A data-driven model for SOFC performance prediction in steady-state operation is developed. The data was collected from 4 different cells, 3 tubular cells tested under 18 operating conditions, and a planar cell tested under 10 operating conditions. The measured experimental data consisting of 23,820 data points is made available as open source for use by other researchers ([link](#)). A DNN is trained using the data collected from one tubular cell and then transferred to the 2 other tubular cells and 1 planar cell. DNN provides higher accuracy than a shallow network due to the nonlinear characteristics of the J-V curves. The DNN is first trained on cell A1, and TL is used for tubular cells A2 and A3, as well as planar cell. The resulting prediction error for all cells is less than 1% with $R^2 > 0.99$. Transfer learning is found to be effective in creating an accurate model when some cells have a limited amount of data. The strategy of normalizing the source data and then using that formula for the target data is found to be most effective in this study. Using 10% of the target domain data was found to be an accepted compromise between accuracy, data availability, and computation time. In summary, by training the network in the source domain (A1) and fine-tuning using the 10% of target data (A2, A3, and planar) in 5 seconds, an 8-fold reduction in computational time, a model is available for prediction in the target domain with a prediction error of less than 1% and R^2 more than 0.99. While this study focuses on laboratory settings, applying the DNN model and transfer learning approach to real-world SOFC systems in the future will test how well it works in the field. Real-world applications introduce additional complexities such as system degradation over time and fluctuating operating conditions, all of which can affect model accuracy.

Appendix B

Benchmark SOH Forecasting Models

All benchmark forecasting models were trained and evaluated using the same DRT-based input features, including the extracted DRT peaks, ohmic resistance, and SOH trajectories, as shown in Figure B.1. The training and testing datasets, look-back window length, forecast horizon, optimizer, and data pre-processing procedures were kept identical across all models to ensure a fair comparison with the proposed GCN-LSTM framework. Specifically, datasets N1-N6 and R1 were used for training, while dataset R2 was reserved for testing. A look-back window of 3 hours and a forecast horizon of 6 hours were applied for all forecasting models. The output focuses on predicting SOH values at specific future time steps. These predictions include points at 10%, 50%, and the full extent of the forecast horizon which is 6 hours.

The hyperparameters of the benchmark models were optimized using a random search strategy within predefined search spaces. Tables B.1-B.3 summarize the hyperparameter search spaces and selected optimal values for each benchmark model. Figure B.2 presents the SOH forecasting results of the benchmark Multi-variate LSTM, GRU, and TCN models on the unseen R2 dataset.

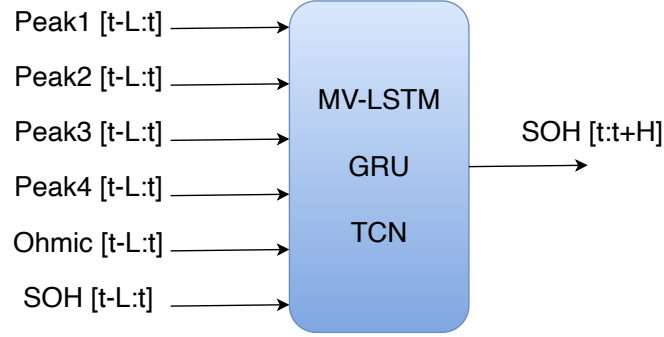


Figure B.1: Schematic representation of the benchmark forecasting models using the same multivariate input features. The inputs include historical DRT peak values, ohmic resistance, and SOH, and the models predict future SOH values over the forecast horizon.

Table B.1: Multivariate LSTM hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of LSTM layers	{1, 2}	2
Number of LSTM units	{32, 64, 128}	64
Learning rate	{1e-2, 1e-3, 1e-4}	1e-4
Batch size	{16, 32, 64}	32

Table B.2: GRU model hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of GRU layers	{1, 2}	2
Number of GRU units	{32, 64, 128}	32
Learning rate	{1e-2, 1e-3, 1e-4}	1e-3
Batch size	{16, 32, 64}	32

Table B.3: TCN model hyperparameter search space and optimal values

Hyperparameter	Search Space	Optimal Value
Number of TCN layers	{2, 3, 4}	4
Number of filters	{32, 64, 128}	64
Kernel size	{2, 3, 5}	3
Dilation rate	{1, 2, 4}	2
Dropout rate	{0, 0.1, 0.2}	0.2
Learning rate	{1e-2, 1e-3, 1e-4}	1e-4
Batch size	{16, 32, 64}	32

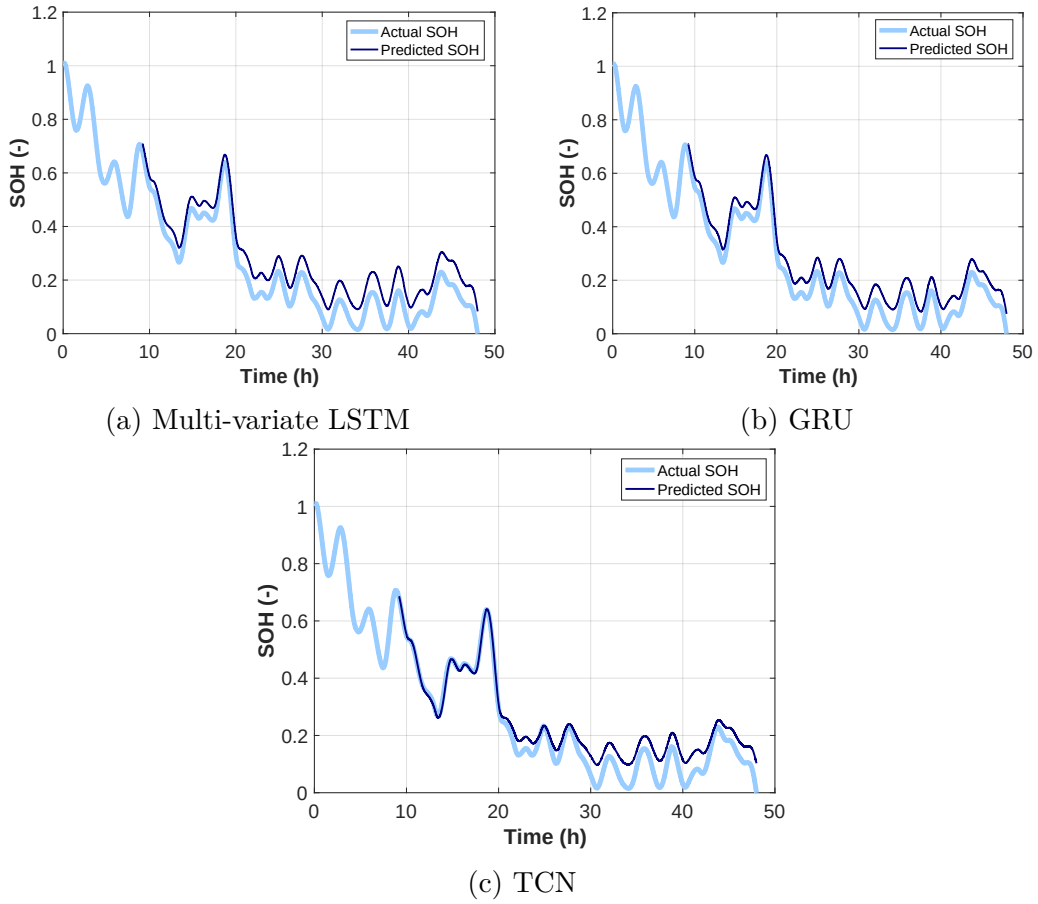


Figure B.2: SOH forecasting results for Multi-variate LSTM (a), GRU (b), and TCN (c) on the R2 dataset.

Appendix C

Ph.D. Publications

Journal Publications from This Thesis

Published:

- **Salehi, Z.**, Fakouri Hasanabadi, M., Smith, D. J., Hanifi, A., Shahbakhti, M., “Bayesian Informer network for probabilistic remaining useful life forecasting of solid oxide fuel cells”, *Control Engineering Practice*, Impact Factor: 4.6, Vol. 173, 11 pages, 2026.
- **Salehi, Z.**, Salahi, A., Fakouri Hasanabadi, M., Hanifi, A., Smith, D. J., Koch, C. R., Shahbakhti, M., “State-of-health forecasting of solid oxide fuel cells using physics-informed temporal graph convolutional network”, *Journal of Power Sources*, Impact Factor: 7.9, Vol. 665, 17 pages, 2026.
- **Salehi, Z.**, Tofigh, M., Kharazmi, A., Smith, D. J., Hanifi, A., Koch, C. R., Shahbakhti, M., “Transfer learning-based deep neural network model for performance prediction of hydrogen-fueled solid oxide fuel cells”, *International Journal of Hydrogen Energy*, Impact Factor: 8.3, Vol. 99, 10 pages, 2025.

Submitted:

- **Salehi, Z.**, Sobhi, S., Hanifi, S., Smith, D. J., Hanifi, A., Shahbakhti, M., “Diagnosis and health monitoring of solid oxide fuel cells using electrochemical impedance spectroscopy under multiple failure modes”, *Journal of Power Sources*, Impact Factor: 7.9, 51 pages in review format, May 6, 2026.
- **Salehi, Z.**, Smith, D. J., Hanifi, A., Shahbakhti, M., “Prognostics of fuel cells based on degradation mechanisms via asymmetric multi-task learning”, *Renewable Energy*, Impact Factor: 9.1, 40 pages in review format, May 17, 2026.

Journal Publications from Other Collaborations

Published:

- Tofigh, M., **Salehi, Z.**, Kharazmi, A., Smith, D. J., Hanifi, A., Koch, C. R., Shahbakhti, M., “Transient modeling of a solid oxide fuel cell using an efficient deep learning HY-CNN-NARX paradigm”, *Journal of Power Sources*, Impact Factor: 7.9, 2024.

Under preparation:

- Vafamand, A., **Salehi, Z.**, Borhan, H., Shahbakhti, M., “Graph Attention Network-Physics-Informed Informer for Joint Battery State of Health Estimation and Degradation Trajectory Prediction”, *Journal of Energy Storage*, Impact Factor: 9.8, 2026.

Peer-Reviewed Conference Papers from This Thesis

Published/Accepted:

- **Salehi, Z.**, Tofigh, M., Vafaeenezhad, S., Kharazmi, A., Smith, D. J., Koch, C. R., Shahbakhti, M., “Performance Prediction of a Range of Diverse Solid Oxide Fuel Cells using Deep Learning and Principal Component Analysis”, *Modeling, Estimation, and Control Conference (MECC), IFAC-Papers-OnLine*, Vol. 58, 6 pages, Chicago, IL, USA, 2024.
- **Salehi, Z.**, Fakouri Hasanabadi, M., Smith, D. J., Hanifi, A., Shahbakhti, M., “Uncertainty-Aware Degradation Trajectory Forecasting for Fuel Cell Prognostics and Health Management”, *The 23rd IFAC World Congress*, 6 pages, Busan, Republic of Korea, 2026.
- **Salehi, Z.**, Sobhi, S., Hanifi, S., Smith, D. J., Hanifi, A. R., Shahbakhti, M., “Attention-Based Failure Mode Classification for Solid Oxide Fuel Cells Toward Predictive Maintenance and Health-Aware Control”, *International Conference on Green Energy and Technology (ICGET)*, Istanbul, Turkey, 2026.

Peer-Reviewed Conference Paper from Other Collaborations

Published:

- Tofigh, M., **Salehi, Z.**, Smith, D. J., Kharazmi, A., Hanifi, A., Koch, C. R., Shahbakhti, M., “Developing a Time-Efficient Model for Solid Oxide Fuel Cells Using Self-Supervised Convolutional Autoencoder and Stateful LSTM Network”, *American Control Conference (ACC)*, Toronto, ON, Canada, 2024.

Appendix D

Thesis Files

The following files were used for this dissertation and are arranged by chapter in the tables below.

D.1 Chapter 1 Source Files

Table D.1: Chapter 1 source files

File Type	File Name	Description
Figure	Figure_1_1.pdf	Figure 1.1
	Figure_1_2.pdf	Figure 1.2
	Figure_1_3.pdf	Figure 1.3
	Figure_1_4.png	Figure 1.4
	Figure_1_5.pdf	Figure 1.5
	Figure_1_6.pdf	Figure 1.6
	Figure_1_7.png	Figure 1.7
	Figure_1_8.png	Figure 1.8
	Figure_1_9.png	Figure 1.9
	Figure_1_10.png	Figure 1.10
	Figure_1_11.pdf	Figure 1.11
	Figure_1_12.pdf	Figure 1.12

D.2 Chapter 2 Source Files

Table D.2: Chapter 2 source files

File Type	File Name	Description
Experimental datasets	redox1.zip	Redox cycling (normal cycle) - N1
	redox2.zip	Redox cycling (normal cycle) - N2
	redox3.zip	Redox cycling (normal cycle) - N3
	redox4.zip	Redox cycling (normal cycle) - N4
	redox5.zip	Redox cycling (normal cycle) - N5
	redox6.zip	Redox cycling (normal cycle) - N6
	redox7.zip	Redox cycling (random cycle) - R1
	redox8.zip	Redox cycling (random cycle) - R2
	starvation1.zip	Fuel Starvation - F1
	starvation2.zip	Fuel Starvation - F2
	starvation3.zip	Fuel Starvation - F3
	shortc1.zip	Short Circuit - S1
	shortc2.zip	Short Circuit - S2
	shortc3.zip	Short Circuit - S3
	thermal1.zip	Thermal Shock - T1
	thermal2.zip	Thermal Shock - T2
	thermal3.zip	Thermal Shock - T3
Code	iv-preparation.py	Script used for processing I-V curves
	eis-preparation.py	Script used for processing EIS curves
Figure	Figure_2_1.png	Figure 2.1
	Figure_2_2a.pptx	Figure 2.2(a)
	Figure_2_2b.pdf	Figure 2.2(b)
	Figure_2_3.pdf	Figure 2.3
	Figure_2_4.pptx	Figure 2.4
	Figure_2_5.pptx	Figure 2.5

D.3 Chapter 3 Source Files

Table D.3: Chapter 3 source files

File Type	File Name	Description
Code	drt-calculation.m	Script used for DRT conversion
	temporal-graph.py	Script used for training the TGCN model
	cnn-lstm.py	Script used for training the CNN-LSTM model
	lstm.ipynb	Script used for training the LSTM model
Figure	Figure_3_1a.drawio	Figure 3.1(a) source
	Figure_3_1a.pdf	Figure 3.1(a)
	Figure_3_1bEIS.png	Figure 3.1(b), EIS
	Figure_3_1bJV.png	Figure 3.1(b), J-V
	Figure_3_2.png	Figure 3.2
	Figure_3_3.drawio	Figure 3.3 source
	Figure_3_3a.pdf	Figure 3.3(a)
	Figure_3_3b.pdf	Figure 3.3(b)
	Figure_3_3c.pdf	Figure 3.3(c)
	Figure_3_4.pptx	Figure 3.4 source
	Figure_3_4a.png	Figure 3.4(a)
	Figure_3_4b.png	Figure 3.4(b)
	Figure_3_4c.png	Figure 3.4(c)
	Figure_3_5a.png	Figure 3.5(a)
	Figure_3_5b.png	Figure 3.5(b)
	Figure_3_5c.png	Figure 3.5(c)
	Figure_3_6a.png	Figure 3.6(a)
	Figure_3_6b.png	Figure 3.6(b)
	Figure_3_7.drawio	Figure 3.7

D.4 Chapter 4 Source Files

Table D.4: Chapter 4 source files

File Type	File Name	Description
Code	bayesian-informer.ipynb	Script used for training the Bayesian Informer model
	bayesian-lstm.ipynb	Script used for training the Bayesian LSTM model
	gru.ipynb	Script used for training the deterministic GRU model
	tcn.ipynb	Script used for training the deterministic TCN model
	informer.ipynb	Script used for training the deterministic Informer model
Figure	Figure_4_1.drawio	Figure 4.1
	Figure_4_2.drawio	Figure 4.2
	Figure_4_3.drawio	Figure 4.3
	Figure_4_4.pdf	Figure 4.4
	Figure_4_5.pdf	Figure 4.5
	Figure_4_6a.pdf	Figure 4.6(a)
	Figure_4_6b.pdf	Figure 4.6(b)
	Figure_4_7a.pdf	Figure 4.7(a)
	Figure_4_7b.pdf	Figure 4.7(b)
	Figure_4_8.drawio	Figure 4.8

D.5 Chapter 5 Source Files

Table D.5: Chapter 5 source files

File Type	File Name	Description
Code	informer-cls.ipynb	Script used for training the Informer classifier
	tcn-cls.ipynb	Script used for training the TCN classifier
	lstm-cls.ipynb	Script used for training the LSTM classifier
	rf-cls.py	Script used for training the RF classifier
	xgb-cls.py	Script used for training the XGBoost classifier
	ensemble-cls.py	Script used for training the RF+XGB ensemble classifier
Figure	Figure_5_1a.pdf	Figure 5.1(a)
	Figure_5_1b.pdf	Figure 5.1(b)
	Figure_5_1c.pdf	Figure 5.1(c)
	Figure_5_2.drawio	Figure 5.2
	Figure_5_2d.pdf	Figure 5.2(d)
	Figure_5_3.drawio	Figure 5.3
	Figure_5_4.drawio	Figure 5.4
	Figure_5_5.drawio	Figure 5.5
	Figure_5_6a.png	Figure 5.6(a)
	Figure_5_6b.png	Figure 5.6(b)
	Figure_5_6c.png	Figure 5.6(c)
	Figure_5_6d.png	Figure 5.6(d)
	Figure_5_6e.png	Figure 5.6(e)
	Figure_5_6f.png	Figure 5.6(f)

D.6 Chapter 6 Source Files

Table D.6: Chapter 6 source files

File Type	File Name	Description
Code	single-rul.ipynb	Script used for training the single RUL estimator
	early-fusion.ipynb	Script used for training the RUL estimator with class-label early fusion
	late-fusion.ipynb	Script used for training the RUL estimator with class-label late fusion
	mtl-informer.ipynb	Script used for training the multi-task learning model
Figure	Figure_6_1a.pdf	Figure 6.1(a)
	Figure_6_1b.pdf	Figure 6.1(b)
	Figure_6_1c.pdf	Figure 6.1(c)
	Figure_6_1d.pdf	Figure 6.1(d)
	Figure_6_2.drawio	Figure 6.2
	Figure_6_3.drawio	Figure 6.3
	Figure_6_4.drawio	Figure 6.4
	Figure_6_5.drawio	Figure 6.5
	Figure_6_6.drawio	Figure 6.6
	Figure_6_7.png	Figure 6.7

D.7 Chapter 7 Source Files

Table D.7: Chapter 7 source files

File Type	File Name	Description
Figure	Figure_7_1.drawio	Figure 7.1