AI Agent for Probabilistic Capacity Degradation Estimation for Varying Lithium-ion Batteries

Sel Ly, Dilip Pandit, Yuliya Preger, Tuan A. Ho, Tu A. Nguyen, Raymond H. Byrne*

***Abstract*—Accurate prediction of lithium-ion battery capacity degradation is central to reliable energy storage system operation and cost-effective maintenance. We propose a hybrid model of quantile Temporal Fusion Transformer (TFT) with Reversible Instance Normalization (RevIN) technique for probabilistic predictions of capacity fades across different cell cathode chemistries, including LFP, NCA, and NMC. In addition, an AI-agent is developed that integrates the hybrid TFT-RevIN model into an autonomous closed-loop AI system enabling continual learning and model adaptation across different datasets. The TFT-RevIN framework demonstrates reliable degradation estimation under substantial distribution and domain shifts, while the integrated AI agent enables adaptive, user-driven model refinement. In particular, the proposed model, trained exclusively on degradation datasets from Sandia National Laboratories (SNL), achieved a mean absolute percentage error (MAPE) of just 0.04% on the test set. Notably, when applied to the Underwriters Laboratories–Purdue University (UL–PUR) dataset, it retained high accuracy with a MAPE of 0.82%. Moreover, fine-tuning the model on a small UL–PUR dataset using the developed AI agent further improved performance, reducing the MAPE to 0.49%.**

***Index Terms*—AI Agent, Battery Capacity Degradation, Distribution Shift, Domain Shift, Reversible Instance Normalization, Temporal Fusion Transformer.**

I. Introduction

Lithium-ion batteries (LIBs) are the dominant energy storage technology across applications, from mobile devices to grid-scale storage. Hence, accurate estimation of the battery degradation process is a necessity for LIBs due to their profound impact on system availability and safety, particularly for establishing effective predictive maintenance strategies. A major challenge in battery prognostics is the variability of degradation across different cell chemistries and operating conditions. In this paper, we focus on estimating capacity degradation across three widely used commercial chemistries: Lithium Iron Phosphate (LFP), Lithium Nickel Cobalt Aluminum Oxide (NCA), and Lithium Nickel Manganese Cobalt Oxide (NMC). Many studies have shown that the cycle life span varies substantially among these chemistries. Crucially, the dependence of capacity fade on variables such as temperature, depth of discharge (DOD), and discharge rate is chemistry-specific and should not be broadly extrapolated across different LIB types [1].

Degradation modeling methodologies generally fall into three categories: physics-driven, data-driven, and hybrid models. Physics-based models rely on fundamental principles and differential equations to predict degradation [2]. While they offer valuable physical insights into mechanisms like Solid–Electrolyte Interphase (SEI) and particle fracture, they are computationally complex, often over-parameterized, and struggle to characterize the complex non-linear degradation mechanisms that dominate chemical wear-out, leading to potential inaccuracies if the underlying physics are incomplete.

On the other hand, data-driven methods, including machine learning (ML) and deep learning (DL) techniques, leverage patterns in historical data to predict system health. References [3]–[6] develop DL degradation models for various cathode chemistries leveraging multivariate regression, neural networks and advanced transformer models. However, a primary disadvantage with the DL models is the necessity for substantial amounts of historical degradation data, which is time-consuming and expensive to acquire due to the extended lifespans of modern batteries. Additionally, DL models face three main challenges in battery applications: (i) **limited interpretability**, as their "black-box" nature hinders understanding and debugging; (ii) **distribution shift**, where non-stationary degradation patterns cause mismatches between training and real-world data; and (iii) **domain shift**, where divergence in battery types and operating conditions degrade model from the original training domain reliability across different systems.

To address existing limitations above, we propose a hybrid modeling framework that integrates the Temporal Fusion Transformer (TFT) for interpretable, high-performance multi-horizon forecasting with Reversible Instance Normalization (RevIN) for robustness under distribution shifts. Here, the TFT architecture [7] fuses the temporal and static features of batteries such as chemistry type, operating conditions, and historical health indicators while capturing the short- and long-term dependencies via attention, and offering interpretability through variable selection and attention weights. In power system applications, the TFT has been employed for short-term load forecasting [8], state-of-charge (SOC) estimation in NaS and Li-ion batteries [9], [10], temperature prediction [11], and capacity estimation [12]. However, these studies overlook the challenge of distribution shift, a common issue arising from the non-stationary nature of battery degradation.

*Corresponding Author. Sel Ly, Dilip Pandit, Yuliya Preger, Tu A. Nguyen, and Raymond H. Byrne are authors with the Department of Energy Storage Technologies and Systems, Sandia National Laboratories, New Mexico 87123, USA, with emails: {*selly, dpandit, ypreger, tunguy, rhbyrne*}*@sandia.gov*. Tuan A. Ho is the author with the Department of Geomechanics and Geochemistry, email: taho@sandia.gov.

This material is based upon work supported by the U.S. Department of Energy, Office of Electricity (OE), Energy Storage Division. This article has been authored by employees of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employees own all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

979-8-3315-8138-1

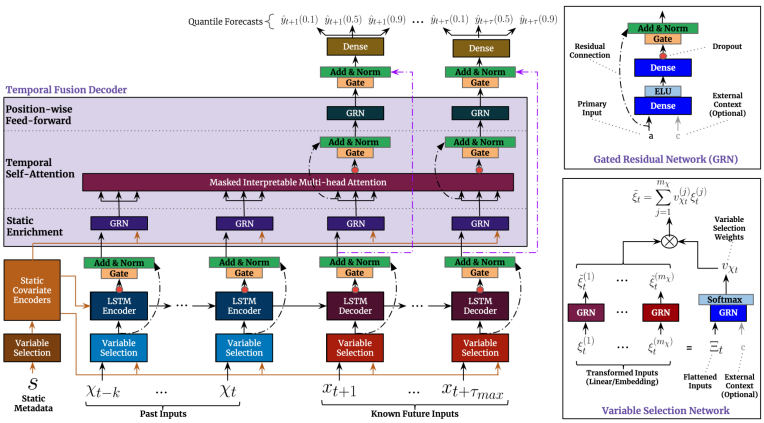


Fig. 1: Temporal Fusion Transformer Model [7].

To address this gap, we integrate RevIN technique [13] into the TFT framework to mitigate the effects of non-stationarity by standardizing and subsequently restoring instance-specific statistical characteristics. This approach effectively reduces discrepancies between training and testing datasets, thereby enhancing model robustness and generalization under varying operating conditions.

Further, we form the proposed hybrid model (TFT + RevIN) as a core module of an AI-agent framework that autonomously accepts user-defined battery scenarios, generates degradation forecasts with uncertainty quantification, and triggers model adaptation or fine-tuning using new dataset when performance degrades, ensuring sustained accuracy and generalization across diverse operational domain shifts. The main contributions of this work are as follows:

- Introduce a hybrid DL model combining TFT and RevIN to accurately estimate capacity degradation across diverse battery chemistries using temporal and static covariates.
- Address non-stationarity and distribution shifts in battery aging via RevIN, improving robustness for long-horizon forecasts.
- Leverage TFT interpretability (attention, variable selection) to identify key degradation drivers and provide actionable insights.
- Develop an AI-agent layer for closed-loop integration, enabling autonomous adaptation, confidence monitoring, and continual learning to enhance reliability and scalability of AI-driven battery health management.

II. Proposed Model: Hybrid TFT with RevIN

We briefly introduce the architecture and mechanisms of the proposed hybrid model (TFT-RevIN) for capacity degradation estimation across **LFP, NCA, and NMC** battery chemistries.

A. Defining Inputs for Degradation Modelling

Accurate prediction of the battery capacity degradation rate requires modelling the nonlinear correlation between capacity and various factors. The TFT architecture, see Fig. 1, is uniquely designed to manage a complex mix of multivariate inputs. In this work, we utilize SNL battery datasets [1] and the inputs for capacity degradation prediction are categorized into three groups. **(i) Static Covariates** $\mathcal{S}_i$ are time-invariant and encode context vectors that condition the temporal dynamics. For each battery i, here we have Battery type (LFP, NCA, or NMC), discharge C-rate, Depth-of-Discharge (DoD), and ambient temperature serving as critical static covariates. **(ii) Observed Past Inputs** ($Y_{i,t}$) are dynamic variables known up to the current time step but unknown for future predictions. These inputs include key health indicators, e.g. historical SOH/Capacity Fade. **(iii) Known Future Inputs** ($\mathcal{X}_{i,t}$) are dynamic but known ahead of time, say, the Time Index or numbers of Equivalent Full Cycles (EFC) ahead that we want to estimate capacity fade rates.

B. TFT Architecture

The core prediction mechanism uses the TFT, an attention-based Deep Neural Network (DNN) favored for high-performance multi-horizon forecasting and interpretable insights into temporal dynamics, see [7] for more details.

1) Gated Mechanisms and Feature Selection: The TFT utilizes two main mechanisms to manage information flow and selection, enhancing efficiency and interpretability. **(i) Gated Residual Network (GRN)** is the foundational building block, providing the flexibility to apply non-linear processing only when necessary. It employs a Gated Linear Unit (GLU) which allows the model to ignore unnecessary input components, thus optimizing information flow. **(ii) Variable Selection Networks (VSN)** identify and select the most salient input features, improving the model's resilience to changing inputs, see [7].

2) Temporal Processing Layers: Standard LSTM sequence-to-sequence layers are used for local processing, while interpretable multi-head attention blocks are used to capture long-term dependencies across the time series. This mechanism provides interpretability by enabling the model to reveal the relative importance of different input features, see [7].

C. RevIN for Non-Stationarity

Battery degradation exhibits **non-stationarity** (distribution shift), where statistical properties of features change over time, leading to discrepancies between training and testing data distributions that degrade model performance. RevIN is a model-agnostic layer implemented to directly mitigate this challenge. Here, RevIN symmetrically applies normalization and denormalization across the temporal dimension of the multivariate time series, see [13] for more details.

1) Normalization Layer: The input sequence $\mathbf{x}^{(i)}$ is normalized by removing instance-specific statistical information (mean $\mu_{\mathbf{x}}$ and standard deviation $\sigma_{\mathbf{x}}$), transforming the data into a mean-centered distribution $\widehat{\mathbf{x}}^{(i)}$. This reduces the distribution discrepancy between different input instances:

$$\widehat{\mathbf{x}}_{k,t}^{(i)} = \gamma_k \cdot \left(\frac{\mathbf{x}_{k,t}^{(i)} - \mu_{\mathbf{x}^{(i)}}}{\sqrt{\sigma_{\mathbf{x}^{(i)}}^2 + \epsilon}} \right) + \beta_k, \tag{1}$$

where γ_k and β_k are learnable parameters, and small $\epsilon > 0$.

2) Denormalization Layer: It is applied to the model's output ($\widetilde{\mathbf{y}}$) to restore the instance-specific statistical information removed during normalization. This symmetrical restoration returns the final prediction ($\widehat{\mathbf{y}}$) to the original distribution, enhancing forecasting stability against non-stationarity

$$\widehat{\mathbf{y}}_{k,t}^{(i)} = \sqrt{\sigma_{\mathbf{x}^{(i)}}^2 + \epsilon} \cdot \left(\frac{\widetilde{\mathbf{y}}_{k,t}^{(i)} - \beta_k}{\gamma_k} \right) + \mu_{\mathbf{x}^{(i)}}. \tag{2}$$

D. Hybrid Modelling Framework

We integrate RevIN layers at the input and output boundaries of the TFT model. RevIN removes per-instance non-stationarity and scale shifts in time-varying covariates, which stabilizes inputs. Then, TFT learns multi-horizon, cross-feature temporal relationships more reliably and yields interpretable attention patterns. This reduces sensitivity to inter-battery distribution shifts and even domain shift, improves calibration and robustness under drifts, and lowers the need for frequent full retraining.

III. AI-Agent Framework

In this premier work, we design an AI-agent, see Fig. 2 that sits atop the pretrained model (TFT-RevIN). Here, our AI-agent app is developed using the LangChain and Streamlit frameworks, with GPT-4o, a Large Language Model (LLM), accessed via the OpenAI API (model id: gpt-4o) and serving as the core engine for understanding user queries and generating contextual responses. The AI-agent presented here is an automation, orchestration and user-interface layer designed to automate dataset preprocessing, model evaluation, and fine-tuning workflows for battery degradation estimation under both distribution and domain shifts. However, it does not itself encode or infer new physical models of battery degradation. In future work, we shall extend the AI-agent by adding an explicit physics-informed decision policy layer that combines domain constraints (safety, thermal, SOC/SOH limits) with pretrained models (SOC/SOH/RUL predictors, battery temperature monitoring, etc.) and reinforcement learning to produce safe, optimal control or scheduling actions (e.g., charge/discharge setpoints, cycling schedules, etc.).

For the current decision logic, our AI-agent continuously monitors model performance and data characteristics and applies a staged retraining policy. Retraining is triggered when a rolling MAPE exceeds a configurable threshold (e.g., 0.5%) and at least a minimum number of new samples are available. Upon trigger, the agent first attempts a lightweight fine-tune using the most recent data window (e.g., the latest 50 cycles) and accepts the update if the validation error falls below the configured threshold. If fine-tuning does not meet this criterion, the agent schedules a full retrain on an expanded dataset. To prevent unbounded retraining, the users can also enforce a maximum number of fine-tune iterations, the retraining process stops when this limit is reached.

Here, the app exposes controls so users can run the pretrained model on a new dataset (e.g., the UL-PUR dataset), inspect reported metrics (MAE, MAPE, RMSE and R^2) and review prediction plots. Then the users can improve the performance by inputing the desired error threshold and maximum fine-tune steps before launching retraining. During fine-tuning the agent reports metrics after each round, so users can monitor convergence (for example, on the UL-PUR dataset the model reached MAPE $< 0.5\%$ after four fine-tune rounds).

Beyond performing the proposed TFT-RevIN model, the agent can also possess comprehensive data analytics capabilities through dynamic Python code generation, execution, and self-debugging. This enables it to perform a wide range of analytical tasks, including computing statistical metrics, generating visualizations, and training ML models autonomously. Users can upload multiple data files in various formats, such as CSV, Excel, and Parquet and specify their tasks. The AI-agent then automatically generates and executes the necessary code based on user's task descriptions, returning the processed results, visual outputs, and insights in real time.

IV. Results and Discussion

A. Data Characteristics and Coverage

The SNL dataset focuses on the performance comparison of three popular commercial battery chemistries, all in the **18650 form factor** [1], including (LFP) A123 Systems (APR18650M1A, 1.1 Ah nominal capacity), and (NCA): Panasonic (NCR18650B, 3.2 Ah nominal capacity), and (NMC) LG Chem (18650HG2, 3 Ah nominal capacity), identified as an Ni-enriched variant of NMC811. Note: Since the cycle data vary in length, we interpolate each sequence to 250 time points using the EFC–Capacity Fade curve as a reference.

The core value of this dataset lies in characterizing battery degradation across varying operational and environmental factors that drive capacity fade, specifically examining the influence of **temperature, Depth of Discharge (DoD), and discharge rate (C-rate)**. These factors are explicitly included as inputs in our model.

B. Model Training Implementation

The proposed TFT-RevIN model is trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and a dropout rate of 0.1 to prevent overfitting. The model configuration includes four attention heads, two LSTM layers, an input sequence length of 50 time steps, an output horizon of 10 steps (prediction ahead), and 11 quantile levels $\{0.05, 0.1, 0.2, \ldots, 0.8, 0.9, 0.95\}$ for probabilistic forecasting. The dataset is divided into 80% (64 batteries) for training and 20% (17 unseen batteries) for testing. To further mitigate overfitting, an early stopping strategy with a patience of five epochs is applied, terminating training when validation performance shows no improvement.

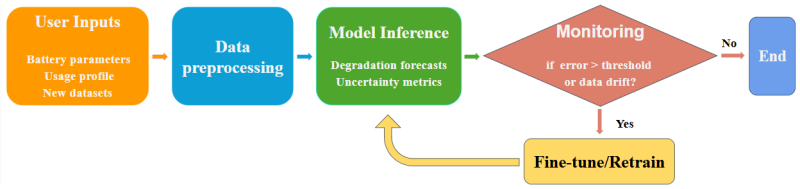


Fig. 2: AI-Agent workflow for degradation prediction and model adaptation.

TABLE I: Performance comparision of the proposed model (TFT-RevIN) vs TFT across different datasets.

Model	Dataset	Retrain	MAE	MAPE	RMSE	R^2
TFT-RevIN	SNL	No	0.04	0.04%	0.06	99.83%
TFT	SNL	No	2.63	3.04%	2.77	-11.63%
TFT-RevIn	UL-PUR	No	0.68	0.82%	1.43	86.92%
		Yes	0.39	0.49%	1.05	90.13%
TFT	UL-PUR	No	5.55	6.28%	5.89	-127.37%
		Yes	2.93	3.40%	3.20	8.24%

C. Model Performance

The proposed TFT-RevIN model achieves excellent performance on the in-domain SNL dataset [1], as shown in Table I, with MAE = 0.0353, MAPE = 0.04%, RMSE = 0.0622, R^2 = 99.83%, reflecting that the model can learn the capacity fade dynamics present in the SNL dataset with very small errors. When the same SNL-trained model is directly applied to the out-of-domain UL-PUR dataset [14], the predictive errors remain relative low, but increased (MAE = 0.6760, MAPE = 0.82%, RMSE = 1.4281, R^2 = 86.92%), demonstrating a clear domain shift between SNL and UL-PUR. Next, by fine-tuning (retraining) the pretrained model on UL-PUR with the help of the built AI-agent using only 5 batteries data, and setting up MAPE of 0.5%, we can reduce the error gap considerably: after retraining the MAE falls to 0.3886 ($\approx 42.5\%$ relative reduction versus the no-retrain UL-PUR run), MAPE drops from 0.82% to 0.49% ($\approx 40.2\%$ relative reduction), RMSE improves by $\approx 26.7\%$, and R^2 increases by 3.21 percentage points (from 86.92% to 90.13%).

Moreover, across both datasets, the proposed TFT-RevIN model substantially outperforms the standard TFT in point-forecast accuracy and explained variance. For instance, on SNL dataset the baseline TFT produces large errors and even negative R^2 (indicative of worse than using mean values for predictions), while TFT-RevIN collapses MAPE to 0.04% and achieve a very high explained variance (99.83%), showing the RevIN technique can handle distribution shift artifacts. Similarly, on UL-PUR dataset, the benefit persists: TFT-RevIN improves the baseline TFT model by lowering MAPE from 6.28% to 0.82% without retraining and to 0.49% with retraining, and increasing R^2 from deeply negative (-127.37%) to 86.92% (no retrain) and 90.13% (retrain). Together these findings support the claim that RevIN substantially improves both average and worst-case forecasting performance under both domain and distribution shifts.

Figure 3 illustrates an example performance that the proposed TFT-RevIN model reproduces the general capacity-fade trend across NCA, NMC and LFP cells, with shaded bands show 95% predictive intervals. We can see that the performance is strongest on NCA, where smooth, slowly-varying degradation is closely tracked. For LFP the model captures long-term slope but shows timing/transition errors around sudden increases in degradation rate. For NMC the model accurately follows the major inflection points with only small local deviations. These patterns indicate the model's strengths on autoregressive, smooth ageing signals and suggest opportunities to improve anticipation of abrupt degradation changes via additional covariates or per-chemistry adaptation.

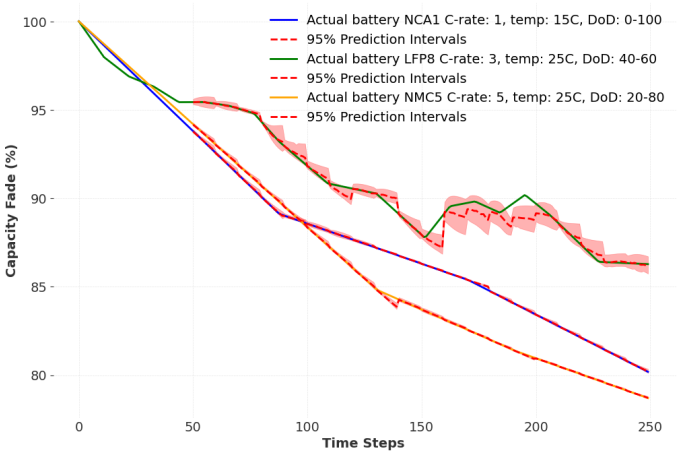
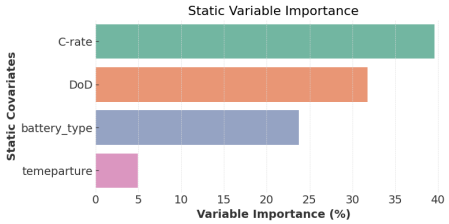
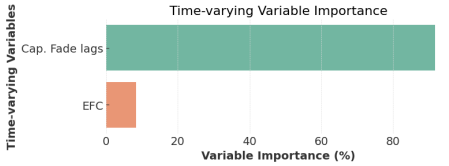


Fig. 3: Prediction performance on different battery types.

Figure 4 shows static and dynamic variable importance results. The model relies overwhelmingly on recent target history (Cap. Fade lags) for forecasting: roughly 91.6% of the dynamic-variable importance is assigned to previous Cap. Fade values. This indicates the predictive signal is dominated by autoregressive behaviour - the recent trajectory of Cap. Fade is the single most informative input for next-step (or horizon) predictions. Of the dynamic covariates, EFC (equivalent full cycles) plays only a small role ($\approx 8.4\%$) relative to Cap. Fade lags. That suggests its information could be already implicitly encoded in Cap. Fade lags (collinearity). For static covariates, the largest single contributor is C-rate ($\approx 39.6\%$), followed by DoD ($\approx 31.8\%$) and battery type ($\approx 23.7\%$). Temperature is assigned very little static importance ($\approx 4.95\%$). High importance of C-rate and DoD matches physical expectations: both directly affect cycling stress and hence capacity fade. Battery type also matters substantially (material/chemistry differences change ageing dynamics). The relatively low importance weight assigned to temperature is noteworthy. This outcome may be attributed to the limited variability of temperature within the dataset, which includes only three distinct values of 15, 25, and 35 °C. Consequently, temperature may contribute minimally to explaining the observed differences in cell degradation behavior.



(a) Static variable Importance.



(b) Temporal variable Importance.

Fig. 4: Importance of static and temporal variables.

TABLE II: Performance comparison of some baseline models.

Model	Dataset	MAE	MAPE	RMSE	R^2
Persistence	SNL	0.70	1.23%	1.07	97.63%
Linear Regression		0.44	0.52%	0.99	86.19%
ARIMA		0.12	0.56%	0.18	92.06%
Persistence	UL-PUR	0.72	0.92%	1.99	87.20%
Linear Regression		0.81	0.96%	1.58	60.21%
ARIMA		0.92	1.21%	1.89	55.94%

D. Comparison with Baseline Models

To quantify the value added by the deep learning predictor, we compare the proposed TFT-RevIN model against simple baselines including persistence ($t + 10 = t$), linear regression model, and ARIMA models selected by AIC. However, we note that the baseline models (except the linear regression model) were trained separately on each battery's historical data and then applied to that battery's remaining records, whereas we only trained a single proposed TFT-RevIN using a pooled data from many batteries and it can be applied to unseen batteries without per-battery retraining.

For a fair comparison, all models are evaluated using the same backtesting procedure and a 10-step forecast horizon. Table II reports averaged MAE, MAPE, RMSE and R^2 for 10-step predictions. We documented that the persistence models yields competitive accuracy compared to linear regression and ARIMA models. However, the proposed TFT-RevIN model still outperforms the persistence model on multi-step forecasts, as shown in Table I.

V. Conclusion

This paper presents a data-driven hybrid model that integrates the strengths of the TFT for interpretable, multi-horizon forecasting of varying complex capacity time-series data with the RevIN technique, which specifically addresses the temporal distribution shift inherent in battery degradation processes across diverse cell cathode chemistries. Evaluations on cell cycling datasets from SNL and UI-PUR demonstrate that the proposed model accurately captures cell degradation under distributional shifts, with MAPE of 0.04% and 0.82%, respectively. In addition, the developed AI-agent effectively automates dataset preprocessing, evaluates model, and retrains the model to address domain shifts, demonstrated by an improved accuracy with MAPE of 0.49% when the TFT-RevIN model trained on SNL data was finetuned on a small UL–PUR dataset of 5 batteries.

Limitation: Our study does not explicitly model dynamic thermal effects. Temperature in the current pipeline is treated as a static/preprocessed covariate, and the training datasets contain only limited ambient temperature variability. As a result, model performance may degrade in outdoor or otherwise thermally variable environments. Future work will address this by (1) augmenting training sets with field data spanning wider temperature ranges, (2) incorporating time-varying temperature covariates and a coupled thermal model (or hybrid physics-informed ML), and (3) applying uncertainty-aware decision policies and field validation to ensure robust performance under thermal stress.

References

[1] Y. Preger, H. M. Barkholtz, A. Fresquez, D. L. Campbell, B. W. Juba, J. Romàn-Kustas, S. R. Ferreira, and B. Chalamala, "Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions," *Journal of The Electrochemical Society*, vol. 167, no. 12, p. 120532, 2020.

[2] S. E. O'Kane, W. Ai, G. Madabattula, D. Alonso-Alvarez, R. Timms, V. Sulzer, J. S. Edge, B. Wu, G. J. Offer, and M. Marinescu, "Lithium-ion battery degradation: how to model it," *Physical Chemistry Chemical Physics*, vol. 24, no. 13, pp. 7909–7922, 2022.

[3] X. Tan, N. Zeng, Y. Deng, S. Sun, B. Rui, and J. Xu, "Early-cycle prediction of battery aging onset across chemistries," *J. Energy Chemistry*, vol. 105, pp. 233–242, 2025.

[4] L. Yang, M. He, Y. Ren, B. Gao, and H. Qi, "Physics-informed neural network for co-estimation of state of health, remaining useful life, and short-term degradation path in lithium-ion batteries," *Applied Energy*, vol. 398, p. 126427, 2025.

[5] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, and D. U. Sauer, "One-shot battery degradation trajectory prediction with deep learning," *Journal of Power Sources*, vol. 506, p. 230024, 2021.

[6] R. Wu, J. Tian, J. Yao, T. Han, and C. Hu, "Confidence-aware quantile transformer for reliable degradation prediction of battery energy storage systems," *Reliability Engineering & System Safety*, vol. 260, p. 111019, 2025.

[7] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International journal of forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.

[8] H. Liao and K. K. Radhakrishnan, "Short-term load forecasting with temporal fusion transformers for power distribution networks," in *2022 IEEE sustainable power and energy conference (iSPEC)*, pp. 1–5, IEEE, 2022.

[9] A. H. Almarzooqi, M. O. Alhusin, I. P. Nikolakakos, A. Husnain, and H. M. Albeshr, "Improved nas battery state of charge estimation by means of temporal fusion transformer," in *2023 IEEE Texas Power and Energy Conference (TPEC)*, pp. 1–6, IEEE, 2023.

[10] A. Pandian, S. Shruthi, V. Dhanush, S. Sundarajan, P. Abishek, S. Subbian, and P. Natarajan, "Temporal fusion transformer-lstm-based improved state of charge estimation of li-ion battery for real drive cycle," in *2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T)*, pp. 592–596, IEEE, 2025.

[11] M. Ma, Q. Xu, Q. Chen, T. Jiang, and H. Wang, "Interpretable and sensorless lithium-ion battery temperature estimation: A temporal fusion transformer approach," *Journal of Energy Storage*, vol. 131, p. 117414, 2025.

[12] W. Gomez, F.-K. Wang, and J.-H. Chou, "Li-ion battery capacity prediction using improved temporal fusion transformer model," *Energy*, vol. 296, p. 131114, 2024.

[13] T. Kim, J. Kim, Y. Tae, C. Park, J.-H. Choi, and J. Choo, "Reversible instance normalization for accurate time-series forecasting against distribution shift," in *International conference on learning representations*, 2021.

[14] D. Juarez-Robles, J. A. Jeevarajan, and P. P. Mukherjee, "Degradation-safety analytics in lithium-ion cells: Part i. aging under charge/discharge cycling," *Journal of The Electrochemical Society*, vol. 167, no. 16, p. 160510, 2020.