Foundation Time-Series Transformers for Lithium-Ion Battery Remaining Useful Life Estimation

Xinyue Su
Stony Brook University
Stony Brook, New York, USA
xinyue.su@stonybrook.edu

Bhavya Jain
North Carolina State University
Raleigh, North Carolina, USA
bjain3@ncsu.edu

Amay Bandodkar
North Carolina State University
Raleigh, North Carolina, USA
ajbandod@ncsu.edu

Tengfei Ma
Stony Brook University
Stony Brook, New York, USA
Tengfei.Ma@stonybrookmedicine.edu

Abstract

Lithium-ion batteries power widely used technologies such as smartphones, laptops, and electric vehicles, making an accurate estimation of battery lifespan critical. Traditional evaluation methods rely on full lifecycle testing, which is expensive and time-consuming. Prior approaches using linear regression for remaining useful life prediction often fail to capture anomalies and nonlinear degradation patterns in battery behavior.

In this paper, we investigate the use of transformer-based foundation models combined with a meta-learning fine-tuning strategy, specifically leveraging First-Order Meta-Learning Algorithms, to improve prediction performance. Our approach enables the model to adapt more effectively to varying battery conditions with limited data.

The models were trained and evaluated on the public "Battery Data Set" from NASA Prognostics Data Repository. Experimental results demonstrate that integrating meta-learning with foundation models enhances adaptability and reduces the RMSE by 98.7% compared to the baseline Ridge regression model. These findings suggest a promising direction for efficient and scalable battery health estimation.

Keywords

foundational model, meta-learning, battery health, transformers

ACM Reference Format:
Xinyue Su, Bhavya Jain, Amay Bandodkar, and Tengfei Ma. 2026. Foundation Time-Series Transformers for Lithium-Ion Battery Remaining Useful Life Estimation. In *Proceedings of 12th Workshop on Mining and Learning from Time Series (MILETS '26).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

MILETS '26, Jeju, South Korea

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Lithium-ion batteries are widely used in major technologies today, including smartphones, laptops, cameras, and electric vehicles. Despite their widespread adoption, battery waste has become a growing environmental concern, making the development of long-lasting and efficient batteries increasingly important.

Traditionally, battery health is evaluated through full lifecycle testing, where a battery is measured from initial use until failure. While accurate, this approach is time-consuming, expensive, and requires specialized equipment such as battery cyclers, environmental chambers, and precision monitoring systems to track performance over thousands of charge–discharge cycles. This makes it impractical for the large-scale evaluation needed for battery engineering. As a result, machine learning methods have been introduced to accelerate battery health assessment by predicting remaining useful life (RUL), enabling accurate lifetime estimation from only using the first few cycle battery charging-discharging behaviors.

Accurate RUL prediction is essential for ensuring the reliability, safety, and longevity of battery systems. However, this task remains challenging due to nonlinear degradation behavior, noisy measurements, and significant variability across batteries caused by differences in usage patterns, manufacturing conditions, and environmental factors. Existing traditional machine learning approaches, including CNN, KNN, random forest, and XGBosst [7], rely on simplified assumptions and struggle to capture complex temporal dynamics. More recent transformer-based models improve sequence modeling capabilities, but they often fail to generalize effectively to unseen batteries, limiting their practical deployment.

Furthermore, battery health prediction suffers from limited data availability due to the high cost of battery testing. One of the widely used NASA Battery Datasets contains only 30 batteries distributed across nine subsets, yielding fewer than four batteries per subset on average. Since each subset was collected under distinct operating conditions, the resulting variations in degradation patterns and charge–discharge behavior pose challenges for generalized machine learning models, which may fail to adequately capture subset-specific characteristics.

To address these challenges, we explore a combination of transformer-based foundation models and meta-learning techniques to improve cross-battery generalization and adaptability under limited data conditions. Public "Battery Data Set" from NASA Prognostics Data

Repository [6], containing batteries with different discharging behaviors, is used to train and evaluate the model. The results show that our meta-learning fine-tuning model significantly improves the battery RUL prediction over traditional OSL Linear Regression models as well as baseline foundational models.

2 Related Work

2.1 Battery RUL Prediction

Remaining Useful Life (RUL) prediction is a fundamental task in battery health management system, aiming to estimate the remaining operational lifetime of a battery before reaching its end-of-life threshold. Traditional approaches relies on collecting statistical feature from voltage, current, capacity, and temperature measurements over time till battery's end-of-life, followed by machine learning models such as linear regression, support vector regression, random forests, and gradient boosting methods. While these methods are computationally efficient and interpretable, they often struggle to capture the nonlinear degradation dynamics and long-term temporal dependencies characteristic of lithium-ion battery aging.

To overcome these limitations, deep learning approaches have become increasingly prevalent for battery prognostics. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) architectures have been employed to automatically learn degradation patterns from battery cycling data. More recently, Transformer-based models [7] [2] and signal detection algorithm [4] have demonstrated strong performance due to their ability to model long-range temporal dependencies through self-attention mechanisms. However, existing approaches are typically trained in a task-specific manner and often require substantial labeled degradation data, limiting their ability to generalize across batteries operating under different conditions.

2.2 Meta-Learning for Time-Series Forecasting

Meta-learning seeks to learn model parameters that can rapidly adapt to new tasks using only a small amount of task-specific data. Among the most widely adopted approaches, Model-Agnostic Meta-Learning (MAML) [3] learns an initialization that can be efficiently adapted through a small number of gradient updates, while First-Order MAML (FOMAML) [5] improves computational efficiency by avoiding second-order gradient computations. These methods have demonstrated success across various few-shot learning problems, including computer vision, natural language processing, and time-series forecasting.

In prognostics and health management applications, meta-learning has been explored as a mechanism for addressing the significant variability observed across machines, sensors, and operating environments [8]. Since individual batteries often exhibit distinct degradation trajectories due to differences in manufacturing processes, usage patterns, and environmental conditions, battery RUL prediction naturally aligns with the meta-learning paradigm, where each battery can be treated as a separate task. However, existing studies primarily focus on task-specific forecasting models and seldom investigate the integration of meta-learning with modern time-series foundation models. In this work, we combine pretrained foundation models with FOMAML-based adaptation to improve battery-specific generalization under limited-data conditions.

3 Method/Approach

3.1 Problem Formulation

Remaining Useful Life (RUL) is a widely used metric in battery health prognostics that estimates the remaining number of operational cycles before a battery reaches its end-of-life threshold. In this work, we define RUL in a normalized form to ensure comparability across batteries with different lifespans. Let $total_c ycle$ denote the total number of charge-discharge cycles a battery can complete before failure, and $current_c ycle$ denote the current observed cycle index. The RUL is defined as: $RUL = \frac{totalCycle - currentCycle}{totalCycle}$ This formulation expresses RUL as a normalized value in $[0, 1]$, where 1 indicates a new battery and 0 indicates end-of-life. This scaling improves model stability and allows direct comparison across different battery trajectories.

3.2 Data Processing

Since each battery is represented as a sequence of cycles, we preprocess the raw signals into cycle-level representations. A full cycle is defined as one charging phase and one discharging phase, with optional impedance measurements. For each cycle, we extract summary statistics from current and voltage signals, including the start, end, and mean values over both charging and discharging phases.

3.2.1 *Feature Construction.*

From raw measurements, we select physically meaningful signals: current, voltage, and discharge capacity.These features are chosen because they directly reflect electrochemical degradation behavior and capacity loss trends.

3.2.2 *Sequence Formation.*

Each battery is represented as: $X = x_1, x_2, x_3, ..., x_n$ where each x_i corresponds to feature values extracted at cycle i.

3.2.3 *Normalization.*

To stabilize training across batteries with different operating ranges, features are normalized at the battery level between $[0, 1]$. This ensures that models learn degradation patterns rather than absolute magnitude differences between batteries, as the number of RUL varies greatly between our batteries.

3.3 Proposed Framework

We propose a framework that integrates pretrained time-series foundation models with FOMAML to enable rapid adaptation to unseen battery degradation trajectories. Each battery is treated as a separate task, and the model is meta-trained to learn an initialization that can be efficiently adapted using only a small number of observed cycles.

3.4 Foundation Model Architectures

3.4.1 *PatchTST-FM-R1.*

We adopt PatchTST-FM R1 [9], a pretrained patch-based time series foundation model, as one of our forecasting backbones. PatchTST-FM R1 divides the input context into non-overlapping patches and processes them through a transformer encoder, capturing both local temporal patterns and long-range dependencies across the degradation trajectory. One thing to note of is the model is pretrained on a sequence of up to 8192 time steps and outputs probabilistic forecasts in the form of quantile predictions across 100 quantile levels. Whereas our pre-cycle RUL series contains only 32-128 cycles per battery. This represents a significant distribution mismatch between pretraining and fine-tuning contexts. We evaluate PatchTST-FM R1 under three training regimes of increasing adaptation intensity: baseline model and meta-learning via FOMAML.

Per-cycle discharge capacity is extracted as the maximum discharge capacity observed during each cycle's discharge phase. Capacity is normalized per battery using min-max scaling:

$$c_i = \frac{c_i - c_{min}}{c_{max} - c_{min} + \epsilon}$$

where $\epsilon = 10^{-8}$ to prevents division by zero. This maps the capacity trajectory to $[0, 1]$, aligning with the model's pretraining distribution and isolating battery-to-battery scale differences.

3.4.2 Chronos-2.

Chronos-2 [1], a pretrained probabilistic multivariate time-series foundation model, is used as one of our primary forecasting backbones. Chronos-2 operates by tokenizing real-valued time series into discrete bins and training a language-model-style transformer to predict future token distributions autoregressively. Unlike task-specific architectures, Chronos-2 is pretrained on a large corpus of heterogeneous time series data, enabling zero-shot generalization to unseen domains without any task-specific architectural modification. For RUL forecasting, we frame the problem as a multivariate time series prediction task, where the forecasting target is the normalized discharge capacity values ($c_i = \frac{c_i - c_{min}}{c_{max} - c_{min} + \epsilon}$ where $\epsilon = 10^{-8}$). To enrich the model's context, we supply past covariates derived from each cycle's electrochemical measurements — including discharge voltage statistics (mean, start, end, drop), discharge current statistics (mean, max), and charge voltage and current features. These covariates are observed over the same historical window as the target and provide the model with degradation signals beyond what the discharge capacity series alone encodes. Linear probing is applied to map the predicted discharge capacity to estimated RUL ($\frac{currentRUL}{totalRUL}$) in the range $[0, 1]$, which is then denormalized by multiplying back by total RUL to recover realistic cycle estimates. We evaluate Chronos-2 under three training regimes of increasing adaptation intensity: zero-shot inference, standard fine-tuning, and meta-learning via FOMAML.

3.4.3 MAML (FOMAML) Fine Tuning.

A central challenge in battery RUL forecasting is that degradation trajectories vary significantly across batteries due to differences in manufacturing, usage conditions, and ambient temperature. A model trained on a population of batteries may fail to capture the idiosyncratic degradation curve of a specific test battery. To address this, we apply First-Order Model-Agnostic Meta-Learning (FOMAML) [5] on top of both Chronos-2 and PatchTST-FM, framing per-battery adaptation as a few-shot learning problem.

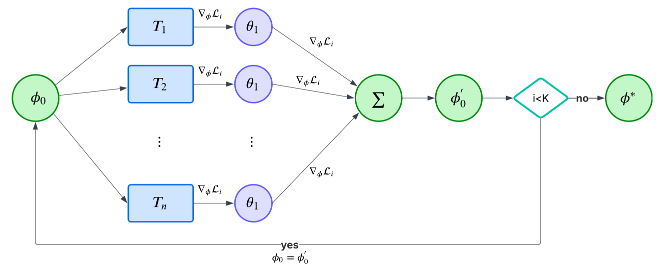

Figure 1: FOMAML Architecture

Task Construction. We define each battery as a separate meta-learning task τ. Given a set of training batteries, the goal of FOMAML is to learn a model initialization θ^* such that a small number of gradient steps on the support set of a new, unseen battery yields strong predictive performance on that battery's query set. For each meta-training episode, a random contiguous window of $K = 24$ cycles is sampled from a training battery as the support set, followed by the next $H = 8$ cycles as the query set. The support set is further partitioned into an inner context of $k - shot = 16$ cycles and inner labels of H cycles, which form the supervision signal for the inner-loop loss. This window-based sampling provides multiple task instantiations per battery per epoch, effectively augmenting the number of available meta-training tasks.

Meta Training. **Inner loop**. For each sampled task, a copy of the current meta-model θ is adapted via $N_{inner} = 10$ steps of SGD with inner learning rate $=10^{-5}$. The inner loss is computed using the backbone's native training objective over the support context and labels. **Outer loop**. The query loss is computed on the adapted parameters θ_i' and its gradient is accumulated onto the meta-model following the FOMAML approximation. A meta-batch of $H = 8$ tasks is sampled per outer update, with gradients averaged across the batch. The meta-model is then updated with Adam at outer learning rate $=10^{-8}$, with gradient norm clipping at 0.01 to prevent instability. Training proceeds for 300 meta-epochs, sampling 5 random windows per battery per epoch. To preserve pretrained representations and limit overfitting on the small battery dataset, only the final two encoder blocks, the final layer normalization, and the output projection are unfrozen during meta-training — approximately 19% of total model parameters.

Initialization Variant. FOMAML is directly applied to our foundational baseline models - PatchTST FM R1, Zero-shot Chronos-2, and Fine-tuned Chronos-2. Zero-shot with MAML tests whether meta-learning alone can bridge the domain gap. PatchTST FM R1 and Fine-tuned Chronos-2 tests whether FOMAML further shapes it for rapid per-battery adaptation.

Evaluate FOMAML. At test time, the first $K = 24$ cycles of each test battery serve as the support set. The meta-model is adapted using $N_{inner} = 10$ SGD steps at $\alpha = 10^{-5}$, producing a battery-specific parameter set θ'. The adapted model then forecasts the subsequent $H = 8$ cycles. No query labels from the test battery are used at any point during adaptation.

4 Experiments

4.1 Dataset and Experimental Settings

We evaluate our models on the publicly available NASA Prognostics Data Repository battery dataset, which is widely used in battery degradation research. The NASA dataset consists of 30 batteries grouped into 9 sets, each representing different operational and cycling conditions. It provides more complex time-series data, including ambient and battery temperature measurements. The datasets also include optional impedance cycles, during which no charging or discharging occurs. Each battery is represented as a sequence of cycles, where each cycle contains sampled sensor readings. The goal is to model degradation over time using these sequential observations.

Following prior work, batteries are split at the battery level into training, validation, and testing sets using a 70/15/15 ratio to prevent information leakage across batteries.

For all experiments, the context length is set to 32 cycles and the prediction horizon is set to 8 cycles. Performance is evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), reported in cycle units after denormalization.

4.2 Baselines

4.2.1 *Ridge Regression.*

We applied Ridge regression to predict RUL across an 8-cycle horizon. At each cycle, a fixed set of per-cycle summary statistics is extracted from the raw time-series measurements. Discharge features include the starting voltage, ending voltage, mean voltage, and voltage drop, as well as mean and peak absolute current. Charge features include the terminal charge voltage, mean charge voltage, and mean charge current. This yields a 9-dimensional feature vector per cycle. The target RUL is normalized by total battery lifetime: $\frac{currentRUL}{totalRUL}$, mapping targets to $[0, 1]$. Rather than predicting a single future RUL value, we formulate the regression as a direct multi-output problem. For each cycle i, the model predicts the normalized RUL at the next 8 cycles: $[i + 1, i + 2, i + 3, ...i + 8]$. Samples where a full 8-cycle lookahead is unavailable are dropped. This avoids the error accumulation typical of recursive forecasting strategies. Features are standardized using z-score normalization (zero mean, unit variance) fit on the training set and applied to validation and test. A MultiOutputRegressor wrapping Ridge($\alpha = 1.0$) is trained on the multi-output targets. The Ridge penalty provides regularization against multicollinearity among the voltage and current features, which tend to be correlated across charge and discharge phases. Train/validation/test splits are performed at the battery level (70/15/15), ensuring no data leakage across batteries.

4.2.2 *PatchTST-FM R1.*

PatchTST-FM R1 is used as a capacity forecaster. The backbone takes a context window of normalized discharge capacity as input and produces a forecast of future capacity, which is used to map the capacity forecast to the RUL prediction target. We append a trainable RUL head that maps the backbone's full quantile forecast distribution — all 99 quantiles across the prediction horizon, flattened into a 792-dimensional vector — to a normalized RUL prediction. Using the full quantile distribution rather than the median alone provides the head with richer information, including the model's uncertainty estimates across future capacity values. The RUL head is a three-layer MLP:

$$Linear(792{\to}256){\to}\text{ReLU}{\to}\text{Dropout}(0.2){\to}\text{Linear}(256{\to}64)$$
$${\to}\text{GeLU}{\to}\text{Dropout}(0.1){\to}\text{Linear}(64{,}8){\to}\text{Sigmoid}$$

The Sigmoid output bounds predictions to $[0, 1]$, consistent with the normalized RUL target space. During linear probing, the backbone is almost entirely frozen. Only the final four transformer blocks $blocks.16$, $blocks.17$, $blocks.18$ and $blocks.19$ and the full RUL head is trainable, representing approximately 19.61% of total parameters. The model is trained with Adam at a learning rate 10^{-4}, MSE loss, for up to 50 epochs with early stopping at patience 7, using a batch size of 16.

4.2.3 *Chronos-2 Zero-Shot.*

In the zero-shot setting, the pretrained Chronos-2 checkpoint is used directly without any parameter updates. The normalized discharge capacity history of each battery serves as the context, and the model predicts the next 8 cycles purely from its pretraining. This establishes an upper bound on what the model can achieve without observing any battery degradation data, and serves as a reference for measuring the benefit of adaptation.

4.2.4 *Chronos-2 Fine-Tuning.*

In the fine-tuning setting, the pretrained Chronos-2 model is further trained on the battery training split using the native $pipeline.fit()$ interface. The model is trained for $2,000$ steps with a learning rate of $5 * 10^{-6}$ and a batch size of 32, with validation loss monitored on the held-out validation batteries. Fine-tuning allows the model to shift its internal representations toward the electrochemical degradation patterns specific to lithium-ion batteries, at the cost of requiring labeled training data.

4.3 Variants of Proposed methods

In addition to standard baselines, we evaluate several variants of our proposed framework that combine pretrained time-series foundation models with meta-learning-based adaptation for battery RUL prediction. Specifically, we investigate three model families under increasing levels of adaptation:

4.3.1 *PatchTST-FM R1 + FOMAML.*

We apply First-Order Model-Agnostic Meta-Learning (FOMAML) on top of the pretrained PatchTST-FM R1 backbone. The model is first pretrained on large-scale time-series data and then meta-trained across battery-level tasks. Each battery is treated as an independent task, enabling the model to learn a shared initialization that can be rapidly adapted to unseen battery degradation trajectories using a small number of observed cycles.

4.3.2 *Chronos-2 + FOMAML (Zero-Shot Initialization).*

We further evaluate FOMAML applied to Chronos-2 in a zero-shot initialization setting, where the pretrained model is directly used without task-specific fine-tuning. Meta-learning is used to adapt the frozen pretrained representations to the battery domain, enabling rapid personalization at test time with limited support data.

4.3.3 *Chronos-2 Fine-Tuned + FOMAML.*

In this variant, Chronos-2 is first fine-tuned on the training battery set and then further adapted using FOMAML. This setting evaluates the complementarity between supervised domain adaptation

Table 1: Model MAE and RMSE Table

Name	MAE	RMSE
Ridge Regression	10.3971	12.2505
Chornos-2 Zero-shot	1.1073	1.7653
Chornos-2 Zero-shot with MAML	0.4368	0.4847
Chornos-2 Fine-tuned	0.8246	1.5434
Chornos-2 Fine-tuned with MAML	0.1330	0.1601
PatchTST FM R1	11.2423	14.0791
PatchTST FM R1 with MAML	1.6844	1.6844

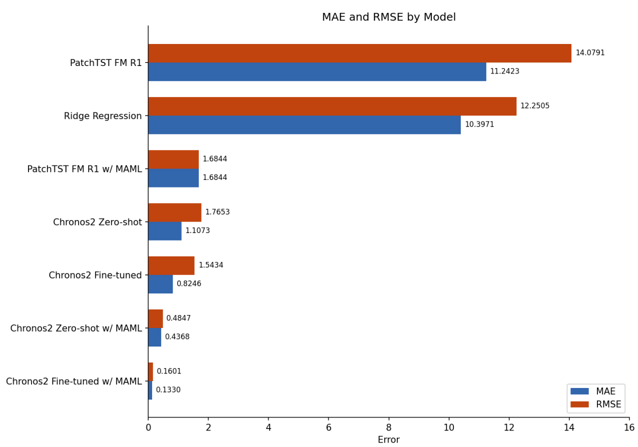

Figure 2: Bar Graph Model Performance Comparison

and meta-learning, where fine-tuning aligns the model to battery-specific distributions while meta-learning improves fast adaptation to unseen batteries.

5 Results

5.1 Overall Performance

We evaluate all models using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The evaluation is done by splitting the battery dataset into training, validation, and testing sets. Each model is asked to predict the next 8 remaining lives. Table 1 summarizes the predictive performance across ridge regression, Chornos-2, and PatchTST FM R1 baselines, and meta-learning fine-tuning approaches.

5.2 Observation

The traditional Ridge Regression model exhibits a high average prediction error of 10.3971 cycles, indicating a limited ability to capture complex battery degradation patterns. Although PatchTST FM R1 was pre-trained on a diverse dataset, a baseline PatchTST model underperformed Ridge Regression with an average of 11.2423 cycles – failing to recognize the non-linear battery degradation patterns. In contrast to PatchTST-FM R1, the zero-shot Chronos2 model already achieves a dramatic reduction in error, having an average error of 1.1073 cycles, suggesting that pretrained multivariate time-series foundation models encode strong general temporal patterns. Application of FOMAML improves all the baseline foundational models (e.g., PatchTST-FM R1 MAE drops from 11.2423 to 1.6844, an 85% improvement), showing that meta-learning helps adapt quickly to new battery trajectories. Out of all the baseline models, fine-tuned Chronos2 yields the best performance improvement, reducing MAE to 0.8246 cycle. This indicates that task-specific adaption is critical for high-precision RUL prediction. The best performance is achieved by combining fine-tuning with MAML ($MAE = 0.1330$), demonstrating that: meta-learning and fine-tuning are complementary rather than redundant. The large performance gap between traditional models and Chronos2-based methods may be attributed to an architectural mismatch between PatchTST-FM's pretraining distribution (long sequences) and the short degradation trajectories characteristic of battery RUL datasets. Chronos-2, by contrast, demonstrates robustness to short context lengths, making it better suited for battery aging applications where cycle counts are inherently limited. We ensured fair comparison by using the same train/test splits, evaluating all models on denormalized cycle units, and reporting both MAE and RMSE.

6 Conclusion / Discussion

6.1 Key Findings

Our experiments demonstrate that transformer-based foundation models significantly outperform traditional machine learning approaches for battery Remaining Useful Life (RUL) prediction which aligns with our hypothesis. In particular, methods based on Chronos2 achieve substantially lower error compared to Ridge Regression and PatchTST-FM R1, highlighting the effectiveness of large-scale pretrained multivariate time-series models for degradation forecasting. We also observe that the zero-shot performance of Chronos2, while competitive, is limited in capturing fine-grained battery-specific degradation patterns. This suggests that although foundation models encode general temporal structures, they lack domain-specific adaptation when applied directly to battery data. Importantly, incorporating meta-learning and fine-tuning leads to substantial performance improvements. This enables rapid adaptation to different batteries and varying cycle dynamics, demonstrating the importance of task-specific calibration even when using powerful pretrained models.

6.2 Insights

Our results provide several key insights. First, transformer-based architectures are particularly well-suited for modeling complex, nonlinear temporal degradation processes such as battery aging. Second, the combination of transfer learning and meta-learning is highly effective in low-data regimes, where traditional supervised learning methods struggle due to limited training samples. Finally, our findings support the idea that general-purpose time-series foundation models can serve as a strong starting point for battery analytics, but require adaptation mechanisms to fully realize their potential in domain-specific applications.

6.3 Challenges

Despite strong empirical performance, several challenges remain. The limited size of available battery datasets constrains both training stability and generalization ability. In addition, significant variability across individual batteries introduces difficulty in learning consistent degradation patterns, particularly for models trained without adaptation mechanisms.

6.4 Future Work

Future work will focus on addressing several limitations. First, we plan to evaluate larger and more diverse datasets to improve robustness and generalization. Second, we aim to incorporate additional physical and operational features, such as temperature and charging conditions, to better model real-world battery behavior. Third, we are also interested in reducing the required historical context length, enabling accurate predictions with fewer observed cycles. Finally, improving generalization to entirely unseen battery chemistries and datasets remains an important direction for real-world deployment.

References

[1] Jaris Küken Andreas Auer Boran Han Pedro Mercado Syama Sundar Rangapuram Huibin Shen Lorenzo Stella Xiyuan Zhang Mononito Goswami Shubham Kapoor Danielle C. Maddix Pablo Guerron Tony Hu Junming Yin Nick Erickson Prateek Mutalik Desai Hao Wang Huzefa Rangwala George Karypis Yuyang Wang Michael Bohlke-Schneider Abdul Fatir Ansari, Oleksandr Shchur. 2025. Chronos-2: From Univariate to Universal Forecasting. arXiv. https://arxiv.org/abs/2510.15821

[2] Durga Prasad Mainali Bishwash Paneru Krishna Bikram Shah Biplov Paneru, Bipul Thapa. 2024. Remaining Useful Life Prediction for Batteries Utilizing an Explainable AI Approach with a Predictive Application for Decision-Making. arXiv. https://arxiv.org/abs/2409.17931

[3] Sergey Levine Chelsea Finn, Pieter Abbeel. 2017. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. arXiv. https://arxiv.org/abs/1703.03400

[4] Lin Z. Wang F. et al Fan, Y. 2025. A hybrid approach for lithium-ion battery remaining useful life prediction using signal decomposition and machine learning. Sci Rep. https://www.nature.com/articles/s41598-025-92262-8#citeas

[5] Alex Nichol, Joshua Achiam, and John Schulman. 2017. On First-Order Meta-Learning Algorithms. arXiv. https://arxiv.org/abs/1703.03400

[6] B. Saha and K. Goebel. 2007. Battery Data Set. NASA Ames Prognostics Data Repository. http://ti.arc.nasa.gov/project/prognostic-data-repository

[7] Erdinc O. Yenioglu M, Aycicek E. 2026. A Data-Driven Framework for Lithium-Ion Battery Remaining Useful Life Prediction Using CNN and Machine Learning Models. MDPI. https://doi.org/10.3390/batteries12040135

[8] Jingsong Xie b Tianci Zhang a Haixin Lv a Tongyang Pan a Yong Feng a, Jinglong Chen a. 2024. Meta-learning as a promising approach for few-shot cross-domain fault diagnosis: Algorithms, applications, and prospects. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0950705121009084

[9] Phanwadee Sinthong Jayant Kalagnanam Yuqi Nie, Nam H. Nguyen. 2023. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. arXiv. https://arxiv.org/abs/2211.14730