Learning from Charging Time Series: Transferable EV Battery Capacity Estimation

Andrei Zakharov[1, 2], **Ilya Makarov**[1, 2, 3]

[1]AI Research Institute, Moscow, Russia [2]Research Center of the Artificial Intelligence Institute, Innopolis University, Innopolis, Russia [3]ISP RAS, Moscow, Russia

Abstract

Accurate estimation of electric vehicle (EV) battery capacity from operational data is critical for range prediction, maintenance planning, and fleet management. Existing battery management systems provide proprietary, vehicle-specific estimates that are often inaccessible to external stakeholders. We propose a transferable, time-series modeling framework that leverages per-cycle charging traces to predict battery capacity across vehicles and manufacturers. Our approach uses Fourier Neural Operators to map multivariate charging time-series into accurate capacity estimates, while incorporating sequence-level constraints to ensure smooth and temporally consistent degradation curves. Evaluated on a real-world dataset of 83 EVs and over 15,800 charging cycles, the model achieves a relative RMSE of 1.6%, outperforming standard baselines. Furthermore, we demonstrate that the learned representations transfer effectively across manufacturers via light fine-tuning on independent subsets. These results highlight the potential of spectral time-series learning for fleet-scale, interpretable, and generalizable battery health estimation, enabling practical applications beyond proprietary systems.

1 Introduction

Monitoring and predicting the remaining capacity of electric vehicle (EV) batteries is essential for ensuring safe operation, maximizing driving range, and optimizing charging strategies. As EV adoption grows, data-driven prediction of capacity fade from operational time-series can improve fleet management, enable preventive maintenance, and reduce safety risks (Noura, Boulon, and Jemeï 2020; Sulzer et al. 2021; Reza et al. 2024). In practice, however, battery management systems (BMS) provide proprietary, vehicle-specific SOH estimates, while widely available telemetry—such as per-cycle charging traces—is noisy, irregular, and highly heterogeneous across vehicles and usage patterns. Environmental factors, including temperature fluctuations and driving behavior, further complicate reliable capacity estimation (Shao et al. 2024). These challenges motivate the development of transparent, transferable time-series models that can operate robustly on real-world, cross-manufacturer data.

Real-world EV charging records exhibit multiple challenges for time-series modeling, such as variable sampling rates, partial or interrupted cycles, measurement noise, and latent environmental influences. Subtle degradation trends are often masked by this variability, requiring models that can capture long-range dependencies and smooth temporal patterns in irregular sequences. Traditional approaches include physics-based models, such as equivalent-circuit or coulomb-counting methods, which depend on precise parameters and perform best under controlled laboratory conditions. Data-driven models, including tree ensembles, shallow neural networks, and recurrent architectures like LSTMs or graph-based CNNs, can capture temporal dynamics but often fail to generalize to unseen data for vehicles from a different manufacturer or usage conditions and may produce non-smooth, temporally incoherent predictions (Ren and Du 2023; Zhang et al. 2023; Khaleghi et al. 2024; Al-Selwi et al. 2024).

To address these limitations, we propose a Fourier Neural Operator (FNO)–based framework (Li et al. 2020) for EV battery capacity estimation. By transforming charging time-series into the frequency domain, the model learns global spectral operators that capture long-range correlations and recurring patterns in the data. Each FNO layer applies a Fourier transform, multiplies by learned spectral weights, and then performs an inverse transform, allowing every output timestep to incorporate information from the entire input sequence. To enforce physically plausible capacity trends, we include a smoothness-aware loss that penalizes abrupt fluctuations, complementing the inherent spectral bias of FNOs and improving robustness in noisy or small-data regimes.

We evaluate our approach on the EVBattery benchmark (He et al. 2022), using a subset of about 80 vehicles and 16K charging cycles. Our FNO-based model achieves superior accuracy compared to the strong baselines, provides interpretable frequency-domain insights, and operates directly on universally available charging traces, without requiring proprietary BMS signals. Moreover, the learned representations transfer across manufacturers via light fine-tuning, demonstrating sample-efficient adaptation to heterogeneous fleets.

The main contributions of this work are: (i) A spectral operator-learning architecture for time-series battery capac-

ity estimation, capturing global temporal structure and improving efficiency on noisy operational data. (ii) A dual-head, smoothness-aware training design that delivers per-timestep and aggregated predictions, yielding physically consistent and interpretable degradation curves. (iii) An empirical demonstration of cross-manufacturer transferability, showing that light fine-tuning recovers SOTA accuracy while accelerating convergence.

Together, these contributions provide a time-series–focused AI methodology for accurate, transferable, and interpretable EV battery health estimation, addressing practical challenges in real-world fleet-scale applications.

2 Related Work

Time-series modeling for battery health estimation has attracted significant attention due to the sequential and noisy nature of operational data. Classical approaches relied on handcrafted features extracted from charge/discharge cycles, combined with linear models (Severson et al. 2019), support vector machines (Nuhic et al. 2013), Gaussian processes (Richardson, Osborne, and Howey 2017; Zhang et al. 2020), or ensemble trees (Zhu et al. 2022). While these models are interpretable and robust on small datasets, they often fail to capture long-range temporal dependencies inherent in battery degradation.

Deep learning methods, particularly recurrent architectures such as LSTMs and GRUs, were introduced to explicitly model the temporal structure of battery time-series (Jiao, Wang, and Qiu 2020; Li et al. 2021). These models can capture sequential dependencies, but they often focus on local patterns, and their predictions may accumulate error over long horizons, especially on irregular, noisy real-world charging traces. CNNs and graph neural networks have been explored to capture local correlations and inter-cell relationships (Wang et al. 2023; Zhou, Qin, and Yuen 2024), while Transformer-based models show promise for modeling long-term dependencies (Zhao et al. 2025). However, most prior work has been limited to laboratory or simulated datasets, and achieving robustness on heterogeneous field data remains challenging.

Physics-informed and hybrid approaches incorporate domain knowledge to improve consistency. Physics-informed neural networks embed electrochemical equations into learning (Wang et al. 2024), but they require accurate parameters and often struggle with large-scale, noisy operational datasets.

Operator learning offers an alternative framework for time-series modeling by capturing global temporal structure. DeepONet (Lu et al. 2021) and Fourier Neural Operators (FNOs) (Li et al. 2020) learn mappings between function spaces, enabling spectral learning and long-range correlations in sequential data. FNOs, in particular, perform global convolutions in the frequency domain, naturally modeling smooth, long-term patterns that are difficult for local or sliding-window methods.

Recent work has begun to leverage operator learning for battery modeling. Panahi et al. (Panahi et al. 2025) used FNOs for lithium-ion dynamics, and Kwak et al. (Kwak et al. 2023) applied an FNO-based approach to EV state-of-charge estimation. These studies demonstrate the potential of spectral operator learning for time-series problems but have focused on lab-scale or simulated data. In contrast, our work applies an FNO1D architecture to real-world, heterogeneous EV charging time-series, introducing smoothness-aware training to enforce physically plausible degradation trends. This combination allows the model to capture long-range correlations and transfer across vehicles and manufacturers, addressing key challenges in practical time-series forecasting for EV battery health.

3 Methodology

3.1 Formal description

Let the dataset consist of cycles collected from N batteries. For battery $i \in \{1, \dots, N\}$ we observe T_i cycle segments. Each cycle (segment) j for battery i is represented by a multivariate time-series

$$\mathbf{X}^{(i,j)} = \left[x_{t,c}^{(i,j)}\right]_{t=1,\dots,L^{(i,j)};\ c=1,\dots,C} \in R^{L^{(i,j)} \times C},$$

where $L^{(i,j)}$ is the number of timesteps in the cycle, C is the number of input channels. Each cycle carries a scalar capacity label $y^{(i,j)} \in R$. Each input cycle $\mathbf{X}^{(i,j)}$ is normalized channel-wise using dataset means, μ, and standard deviations, σ, as $\tilde{\mathbf{X}}^{(i,j)} = \frac{\mathbf{X}^{(i,j)} - \mu}{\sigma + \varepsilon}$, where $\varepsilon > 0$ prevents division by zero. Cycles have varying length $L^{(i,j)}$ and within a batch are padded to the maximum length $L_{\max}$ and a binary mask $m^{(i,j)} \in \{0,1\}^{L_{\max}}$ indicates valid timesteps. For the cycle-level regularization, cycles from the same battery are grouped into small windows via the batching routine, so that a training batch contains multiple cycles of the same battery.

The learning model is an operator f_θ parameterized by θ, implemented as a one-dimensional Fourier Neural Operator (FNO1d). For a padded and normalized cycle input $\tilde{\mathbf{X}} \in R^{L \times C}$, the model produces $\hat{\mathbf{y}}_{\text{seq}} \in R^L$ and $\hat{y} \in R$, where $\hat{\mathbf{y}}_{\text{seq}}$ is the per-timestep sequence output produced by the sequence head, and $\hat{y}$ is the per-cycle scalar capacity estimate produced by the scalar head after pooling. Formally, $\hat{\mathbf{y}}_{\text{seq}},\ \hat{y} = f_\theta\big(\tilde{\mathbf{X}},\ m,\ \mathbf{c}\big)$, where m is the valid-timestep mask and $\mathbf{c} \in R^{d_c}$ denotes optional cycle-level features (e.g. duration, max/min temperature) concatenated to the pooled latent representation before the scalar head.

Each FNO layer projects the per-timestep features into a latent width W, applies a spectral convolution in Fourier space multiplying low-frequency Fourier coefficients by learned complex weights, combines the spectral output with a local convolution (residual connection), and applies normalization and nonlinearity. The final latent is mapped to the two heads. One is a convolutional sequence head that returns $\hat{\mathbf{y}}_{\text{seq}}$ and the second is a linear scalar head that returns $\hat{y}$.

The full loss minimized during training is a weighted sum

$$\mathcal{L}(\theta) = \alpha \mathcal{L}_{\text{scalar}} + \beta \mathcal{L}_{\text{seq}} + \gamma_{\text{t}} \mathcal{L}_{\text{time}} + \gamma_{\text{c}} \mathcal{L}_{\text{cycle}},$$

with hyperparameters $\alpha, \beta, \gamma_{\text{t}}, \gamma_{\text{c}}$. The per-cycle scalar loss is a Smooth-ℓ_1 (Huber-like) loss applied to normalized

scalar predictions

$$\mathcal{L}_{\text{scalar}} = \frac{1}{B}\sum_{b=1}^{B} SmoothL1\Big(\frac{\hat{y}^{(b)}-\mu_y}{\sigma_y+\varepsilon}, \frac{y^{(b)}-\mu_y}{\sigma_y+\varepsilon}\Big),$$

where μ_y, σ_y are the training-set mean and standard deviation of the capacity labels. The within-cycle supervised loss is a mask-aware mean squared error:

$$\mathcal{L}_{\text{seq}} = \frac{\sum_{b=1}^{B}\sum_{t=1}^{L_{\max}} m_t^{(b)}\big(\hat{y}_{\text{seq},t}^{(b)} - y_{\text{seq},t}^{(b)}\big)^2}{\sum_{b=1}^{B}\sum_{t=1}^{L_{\max}} m_t^{(b)} + \delta},$$

with a small $\delta > 0$ to avoid division by zero. The sequence head is regularized to discourage large adjacent differences:

$$\mathcal{L}_{\text{time}} = \frac{1}{B}\sum_{b=1}^{B} \frac{\sum_{t=1}^{L_{\max}-1} m_{t+1}^{(b)} m_t^{(b)} \big(\hat{y}_{\text{seq},t+1}^{(b)} - \hat{y}_{\text{seq},t}^{(b)}\big)^2}{\sum_{t=1}^{L_{\max}-1} m_{t+1}^{(b)} m_t^{(b)} + \delta}.$$

And finally, to encourage temporal coherence across cycles of the same battery, for each battery i let c_i index the cycles of battery that appear in the batch, ordered by cycle number n. If the ordered scalar predictions are $\{\hat{y}_{n_k}^{(i)}\}_{k=1}^{c_i}$, we define

$$\mathcal{L}_{\text{cycle}} = \frac{1}{N_{\text{batt}}}\sum_{i}\frac{1}{c_i-1}\sum_{k=1}^{c_i-1}\big(\hat{y}_{n_{k+1}}^{(i)} - \hat{y}_{n_k}^{(i)}\big)^2,$$

where the outer sum is taken only over batteries with $c_i \geq 2$. This term is implemented by grouping predictions by car (battery) number and ordering by the cycle number.

3.2 Overall framework

The proposed framework (Fig.1) end-to-end maps raw per-cycle charging time-series into per-cycle capacity estimates and produces temporally coherent degradation trajectories for each vehicle. It is organized as a modular pipeline with four components: (i) data input and preprocessing, (ii) grouped batching and normalization, (iii) the FNO1d model with dual heads, and (iv) loss-driven training with per-battery aggregation.

3.3 Training setup

The model was trained end-to-end using mini-batch gradient descent with the AdamW optimizer (learning rate $1e-4$, weight decay $1e-6$) and a ReduceLROnPlateau scheduler (factor 0.5, patience 5) for a total 50 epochs. We used dataset-level train/validation five-fold split by car numbers shuffled with random seed. Each cycle segment was represented as a variable-length multivariate time-series and carries a scalar label for capacity. Cycles were padded to the maximum length within each batch and a binary mask marked valid timesteps. The effective training batch size was set to 32. The FNO model used in experiments was instantiated with width 128, modes 32, depth 4, and dual heads producing a per-timestep sequence and a pooled per-cycle scalar. In all runs we set random seed 42. The loss weights were chosen as $\alpha, \beta = 1$, $\gamma_{\text{t}} = 1e-2$,$\gamma_{\text{c}} = 1e-3$. The best model checkpoint with the lowest validation RMSE was used for inference.

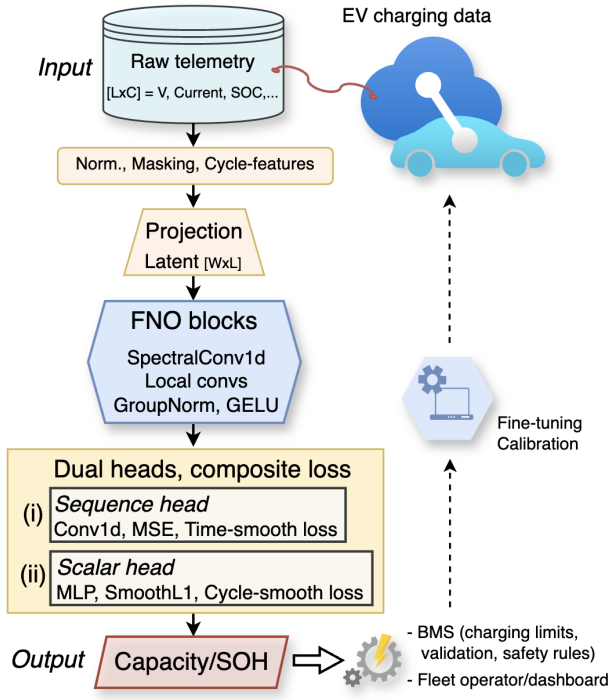

Figure 1: Schematic of the proposed Fourier Neural Operator (FNO)–based framework for electric vehicle battery capacity prediction. Spectral convolution layers map input time-series to capacity forecasts, while a composite loss ensures both accuracy and smooth, physically consistent degradation trajectories.

3.4 Electric Vehicle Dataset

We used Dataset 2 from He et al. (He et al. 2022) as the empirical basis for our experiments. After preprocessing and filtering, the subset employed in this study comprises total of 83 vehicles and 15816 labeled charging cycles. Each cycle is represented as a variable-length multivariate time series (voltage, current, SOC, max/min voltages, max/min temperatures, elapsed time) together with an engineer-provided scalar capacity label and per-cycle handcrafted features (duration, mean SOC, peak current, voltage std) that are also computed for the model. The data are real-world and therefore noisy and heterogeneous with cycle records being available only for a subset of charging events. Cycles are also unevenly distributed across vehicles, and charging durations vary substantially. Figure 2 illustrates the distribution of capacity values across the dataset (top left), the capacity versus cycle number for three representative vehicles (bottom left), and within-cycle dynamics of voltage, current and capacity (right), motivating the use of spectral modeling and smoothness regularization.

4 Results

4.1 Prediction accuracy

In Table 1, we report the aggregate test performance of our FNO1d model compared with standard baselines. While en-

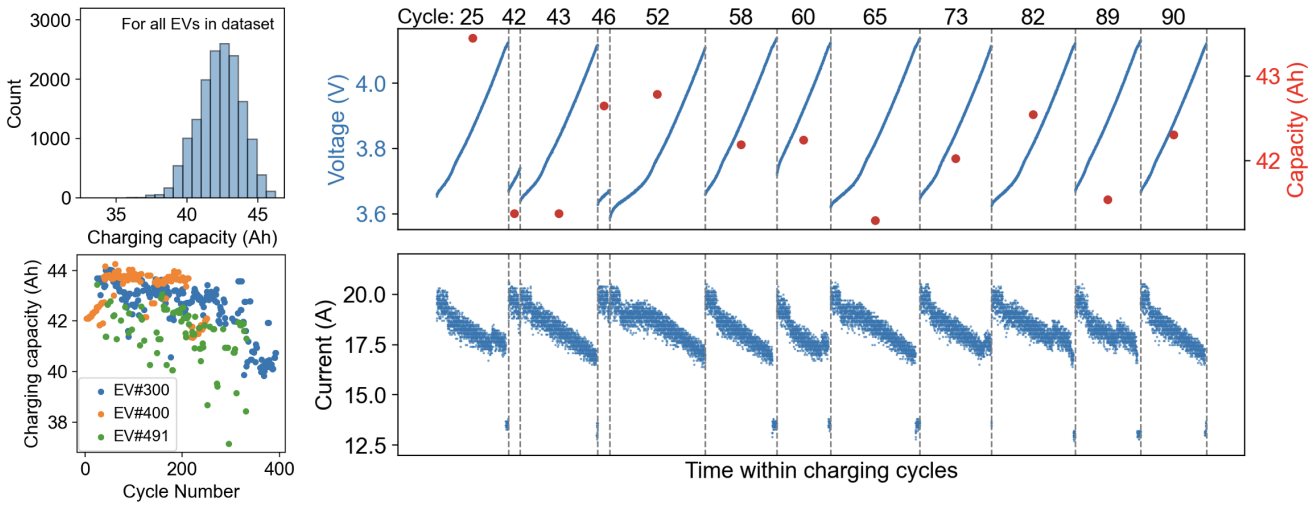

Figure 2: (Top left) The distribution of charging capacity values within EVs dataset. (Bottom left) Dependence of the charging capacity on the number of charging cycles for three representative EVs. (Right) Dynamics of voltage, current and charging capacity during charging cycles. Note that data are available only for some of the charging cycles that are unevenly distributed and their charging time significantly varies.

semble and conventional sequence models produce similar errors (RMSE≈1.39-1.43), our FNO1d with dual heads and a weighted composite loss achieves a substantially lower RMSE of 0.68±0.18, roughly a two-fold reduction versus the best baseline. This corresponds to an average relative RMSE of about 1.6% with respect to the mean capacity. We attribute this improvement to several complementary design choices. First, the FFT-based spectral convolutions successfully capture global, low-frequency patterns and long-range dependencies in noisy cycle traces that local time-domain operators struggle to learn. Second, the dual-head architecture supplies dense per-timestep supervision (sequence head) while the pooled latent plus handcrafted cycle features drive a robust scalar estimate (scalar head). These two smoothness regularizers (within-cycle and cycle-to-cycle) suppress spurious high-frequency fluctuations and remove implausible jumps across consecutive cycles. Also, grouped batching by battery enables the cycle regularizer to operate effectively, further improving temporal coherence of the estimated degradation curves.

We next examined model performance in more detail (Fig. 3). The parity (predicted vs true) plot shows a strong overall agreement with predictions clustered tightly around the identity line, indicating low systematic bias and good per-sample fidelity. To explore temporal behavior we plot prediction error versus cycle number, and observe that errors remain small and stable for the first roughly 400 cycles, after which the model develops a positive bias and begins to slightly overestimate capacity, with mean error rising to about 1.0 Ah and the variance increasing. The per-car RMSE cumulative distribution also shows that most EVs achieve low RMSE, while a tail of cars, typically those with larger observed cycle numbers or irregular usage, account for the larger errors. Finally, three representative vehicle trajectories (for EVs with numbers 306, 334, 464) plotting predicted and true capacity versus cycle number show that despite noisy, irregular measurements and strongly non-linear degradation patterns the model closely tracks the ground truth and recovers the main decline and salient features in each case.

Table 1: **Comparison of model accuracy in EV capacity estimation against benchmark baselines.**

Model	RMSE (Ah)	Model configuration
Rand.Forest	1.42 ± 0.11	Trees/depth=4/-/ RMSE loss
XGBoost	1.39 ± 0.13	Trees/depth=4/-/RMSE loss
MLP	1.43 ± 0.12	MLP/hid=32/Adam/MSE
GCNN	1.42 ± 0.11	4 gate+FC+resid/Adam/MSE
LSTM	1.41 ± 0.11	LSTM+2FCL/hid=32/Adam/MSE
Our model	$\mathbf{0.68 \pm 0.18}$	FNO1d + DualHeads

4.2 Ablation study

We evaluated three modifications to our reported FNO1d model summarized the results in Table.2. On the full dataset consisting 83 EVs with 17 of them in validation set, the FNO baseline, the FNO variant trained without time smoothing ($\gamma_t = 0$) and the FNO variant with the spectral operator replaced by a local convolution produce very similar patterns. All three variants show a roughly 35% increase in mean rolling standard deviation, a moderate reduction in counted extreme short term spikes and an increase in high frequency energy fraction, which together indicate slightly larger short term variability and fewer oscillatory flips compared to the true capacity curves. The variant without time smoothing has essentially the same runtime and behavior as

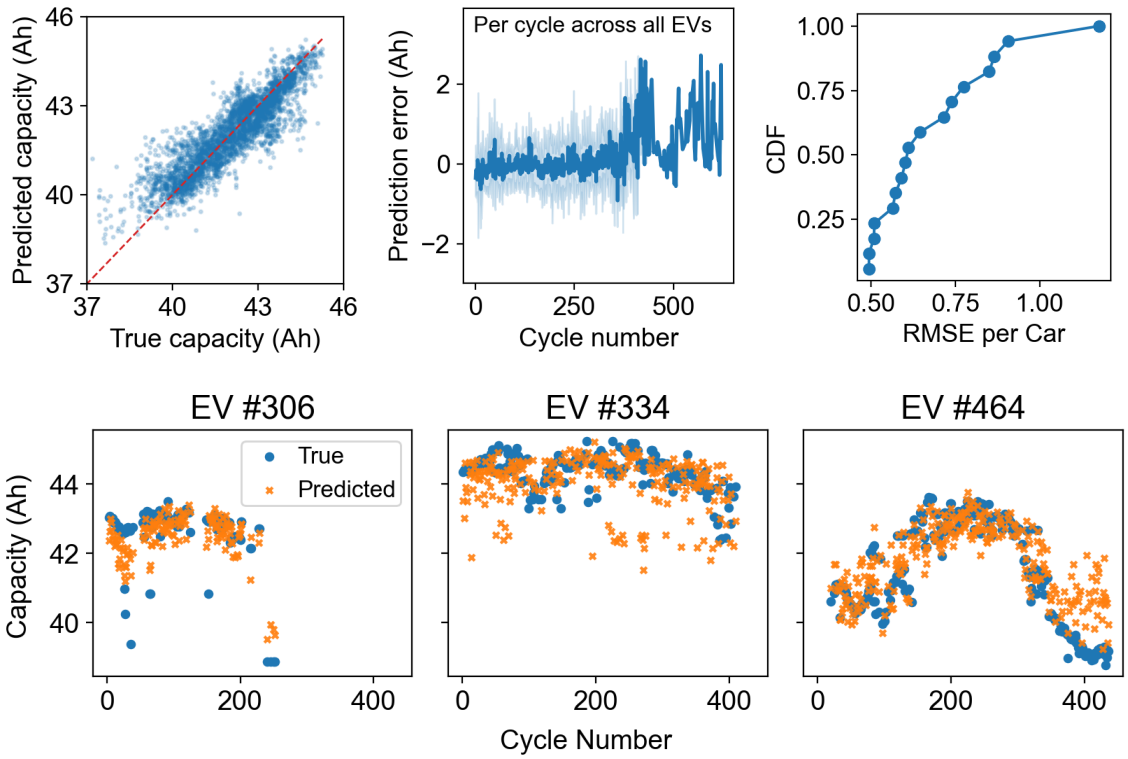

Figure 3: Predicted versus true capacity and diagnostic plots summarizing model performance. (Top left) A parity plot with predictions clustered near the identity line indicating low bias. (Top middle) Prediction error versus cycle number, with an increasing positive bias only in EVs with larger number of cycles. (Top right) The cumulative distribution of per-car RMSE, indicating most vehicles achieve low error. (Bottom) Three representative vehicle trajectories, with predicted and true capacity versus cycle number, illustrating that the model tracks non-linear degradation trends and noisy measurements for typical vehicles.

the baseline, which indicates that the smoothing term carried little weight in the learned solution. However, the local unshared convolution run is much slower, with training time per epoch about three times larger than the baseline, because filters require distinct weights at each position and therefore many more parameters and compute, but these heavier computations do not translate to improved fidelity. The smaller dataset trained on 20 of 83 EVs completes far faster per epoch, and it produces closer agreement on some aggregate measures but a dramatic increase in high frequency energy and larger extreme spikes, a pattern consistent with overfitting on the limited subset.

In Fig.4a we also show how the validation RMSE curves steadily improve with epochs for all variants and reach stable convergence, but the model trained on the much smaller subset starts with a much higher error and remains substantially worse, demonstrating that reducing the training set harms performance. The full data runs reach very similar final RMSE, suggesting the light smoothing term has little effect in a model that already benefits from the spectral operator's implicit low frequency inductive bias, while replacing the spectral operator with a local convolution greatly increases computational cost without improving accuracy.

Table 2: **Relative changes in temporal and high-frequency metrics of predicted compared to true capacity values across different model variants and dataset sizes.**

Metric change relative to true curve	FNO	FNO, no time smoothing	No spectral	Smaller data set (25% EVs)
Mean rolling STD	35.7%	35.0%	35.5%	5.0%
Max rolling STD	-13.7%	-12.6%	-7.6%	33.2%
N sign changes	-8.1%	-10.7%	-8.7%	-12.5%
High freq. energy	30.9%	26.9%	19.7%	171.6%
Epoch train time	695s	697s	1931s	182s

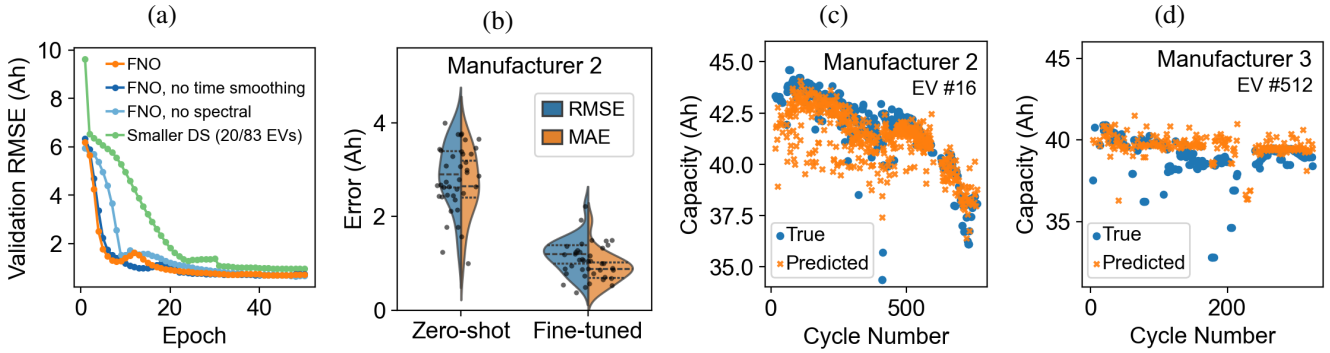

Figure 4: (a) Validation RMSE curves showing similar convergence for full-data runs of different variants of the model and degraded accuracy with reduced data. (b) Transfer learning results on new Manufacturer 2 dataset, showing substantially reduced metrics after light fine-tuning. (c) Example EV capacity trajectory from Manufacturer 2 dataset demonstrating close agreement between predicted and true capacities for over 750 cycles. (d) Prediction results for example EV capacity trajectory from Manufacturer 3 dataset after a light fine-tuning.

4.3 Transfer learning across manufacturers

Because BMS algorithms are proprietary and not generally transferable across EV/ battery manufacturers and charging operators, we investigate whether our pretrained FNO-1D can generalize across manufacturers by testing it on Datasets 1 and 3 from He et al. We start with Datasets 1 comprising 100 vehicles and about 27,000 charging cycles, from which we randomly held out 20 vehicles as a validation set and performed zero-shot inference by applying the model pretrained on Dataset 2 directly to these held-out vehicles without any further adaptation. To measure adaptation gains, we then fine-tuned the same pretrained model for additional 20 epochs using training data drawn from Dataset 1. Concretely, we selected another disjoint set of 20 vehicles from Dataset 1 for fine-tuning and used a five-fold split within that set to form fine-tune training/validation folds; the fine-tuned model was evaluated on the original 20 held-out validation vehicles. Then, we performed the same test on Datasets 3 from a different manufacturer.

Results are summarized in Table 3 and depicted in Figure 4 (b-c panels). Even in the zero-shot setting, the model achieves a mean RMSE of 2.915 and mean MAE of 2.684, which already compares favorably with baseline models (the best prior result was LSTM with mean RMSE 1.69 on Dataset 1). Fine-tuning reduces errors substantially to mean RMSE 1.186 and mean MAE 0.903, well below baseline performance and similar to the gains observed on Dataset 2. Figure 4c shows a representative validation vehicle with over 750 cycles of history, where the fine-tuned model closely tracks the true capacity trajectory, demonstrating improved accuracy and temporal coherence. Similar high prdicton accuracy is achieved on trajectories from Manufacturer 3, for which we show results for a representative EV in Figure 4d. These findings indicate that the spectral priors and latent representations learned on one EV Dataset transfer effectively to a different manufacturer's data and that a small amount of in-domain fine-tuning is sufficient to recover high accuracy.

Table 3: Transfer learning results on new Manufacturer 2 dataset. Performance is reported as mean RMSE and MAE over held-out validation EVs.

Dataset, Model	EVs	Cycles	RMSE ↓	MAE ↓
EV manufacturer 1				
Pretrain, FNO	83	∼15.8k	**0.68**	**0.516**
EV manufacturer 2				
LSTM (He et al. 2022)	100	∼27k	1.690	–
Zero-shot, FNO	20	∼5.4k	2.915	2.684
Finetuned 20 ep, FNO	20	∼5.4k	**1.186**	**0.903**
EV manufacturer 3				
Zero-shot, FNO	8	∼1.2k	3.275	3.004
Finetuned 5 ep, FNO	8	∼1.2k	**1.760**	**1.092**

5 Conclusion

This work advances practical battery prognostics by showing that spectral operator learning can make EV battery capacity estimation more accurate and more robust to noisy, irregular field data. Beyond the numerical gains showed in the experiments, the main value of the approach lies in the combination of three properties. First, the spectral model provides a natural inductive bias for slowly varying degradation signals, which leads to stable predictions. Second, the dual output design supports dense supervision and improves representation learning for per cycle and per timestep tasks. Third, the training penalties and grouped batching by vehicle encourage temporal coherence that aligns model outputs with physical processes of gradual capacity loss.

These properties open several concrete application opportunities. The model can improve range estimation and driver facing range confidence, which in turn can raise user trust and reduce unexpected range shortfalls. It can support predictive maintenance workflows by flagging vehicles with atypical degradation trajectories and prioritizing inspections or controlled cell testing. The method can also inform warranty and fleet management decisions by providing more reliable lifetime forecasts when aggregated over many vehicles. Finally, the spectral diagnostics that arise from the learned frequency content can help engineers identify shifts in usage patterns and potential failure modes in a more interpretable way.

There are practical considerations to address before deployment. The current work treats each vehicle independently and does not yet exploit fleet wide covariates that may capture shared operating regimes. The model does not provide calibrated uncertainty estimates, which are important for risk sensitive decisions such as charging policy or maintenance scheduling. Although the dataset used in this study has uneven coverage and missing readings, production systems will need robust schemes for irregular sampling and missingness that go beyond simple padding and masking.

Several research directions appear especially promising. For example, incorporating lightweight physics informed constraints could combine domain knowledge with the spectral inductive bias and improve interpretability. Adding probabilistic output layers or ensembling could yield calibrated uncertainty that is useful for decision making. Online adaptation and continual learning could allow the model to update as new EV data arrive, reducing the need for frequent full retraining. Federated or privacy preserving training strategies could enable model sharing across manufacturers while keeping raw data private. Finally, expanding evaluation to larger and more diverse EV fleets and to multimodal data that includes driving behavior or ambient conditions will help clarify the generality and limits of the approach.

Overall, spectral operator learning offers a practical and interpretable path toward more reliable battery capacity estimation at large scale. The results reported in this work suggest that combining spectral priors with modest regularization and vehicle aware batching can deliver useful improvements for real world EV management, and they point to a number of extensions that can bridge the gap between research prototypes and production battery management systems.

To support reproducibility and community adoption, we will release the processed datasets and pre-trained model checkpoints in a public repository upon publication.

Acknowledgements

The work was supported by the Ministry of Economic Development of the RF (agreement No. 139-10-2025-034 dd. 19.06.2025, IGK 000000C313925P4D0002).

References

Al-Selwi, S. M.; Hassan, M. F.; Abdulkadir, S. J.; Muneer, A.; Sumiea, E. H.; Alqushaibi, A.; and Ragab, M. G. 2024. RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. *Journal of King Saud University-Computer and Information Sciences*, 36(5): 102068.

He, H.; Zhang, J.; Wang, Y.; Jiang, B.; Huang, S.; Wang, C.; Zhang, Y.; Xiong, G.; Han, X.; Guo, D.; et al. 2022. EVBattery: A large-scale electric vehicle dataset for battery health and capacity estimation. *arXiv preprint arXiv:2201.12358*.

Jiao, M.; Wang, D.; and Qiu, J. 2020. A GRU-RNN based momentum optimized algorithm for SOC estimation. *Journal of Power Sources*, 459: 228051.

Khaleghi, S.; Hosen, M. S.; Van Mierlo, J.; and Berecibar, M. 2024. Towards machine-learning driven prognostics and health management of Li-ion batteries. A comprehensive review. *Renewable and Sustainable Energy Reviews*, 192: 114224.

Kwak, M.; Jin, H. S.; Lkhagvasuren, B.; and Oyunmunkh, D. 2023. A Robust State of Charge Estimator Based on the Fourier Neural Operator for xEV Batteries. *Journal of The Electrochemical Society*, 170(10): 100504.

Li, W.; Sengupta, N.; Dechent, P.; Howey, D.; Annaswamy, A.; and Sauer, D. U. 2021. Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *Journal of power sources*, 482: 228863.

Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2020. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*.

Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; and Karniadakis, G. E. 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3): 218–229.

Noura, N.; Boulon, L.; and Jemeï, S. 2020. A review of battery state of health estimation methods: Hybrid electric vehicle challenges. *World Electric Vehicle Journal*, 11(4): 66.

Nuhic, A.; Terzimehic, T.; Soczka-Guth, T.; Buchholz, M.; and Dietmayer, K. 2013. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *Journal of power sources*, 239: 680–688.

Panahi, A. A.; Luder, D.; Wu, B.; Offer, G.; Sauer, D. U.; and Li, W. 2025. Fast and Generalizable parameter-embedded

Neural Operators for Lithium-Ion Battery Simulation. *arXiv preprint arXiv:2508.08087*.

Ren, Z.; and Du, C. 2023. A review of machine learning state-of-charge and state-of-health estimation algorithms for lithium-ion batteries. *Energy Reports*, 9: 2993–3021.

Reza, M.; Mannan, M.; Mansor, M.; Ker, P. J.; Mahlia, T. I.; and Hannan, M. 2024. Recent advancement of remaining useful life prediction of lithium-ion battery in electric vehicle applications: A review of modelling mechanisms, network configurations, factors, and outstanding issues. *Energy Reports*, 11: 4824–4848.

Richardson, R. R.; Osborne, M. A.; and Howey, D. A. 2017. Gaussian process regression for forecasting battery state of health. *Journal of Power Sources*, 357: 209–219.

Severson, K. A.; Attia, P. M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M. H.; Aykol, M.; Herring, P. K.; Fraggedakis, D.; et al. 2019. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5): 383–391.

Shao, Y.; Zheng, Y.; Zhang, J.; Han, X.; Jin, B.; and Sun, Y. 2024. A cloud capacity estimation method for electric vehicle lithium-ion battery independent of cloud SOC. *Journal of Energy Storage*, 85: 110998.

Sulzer, V.; Mohtat, P.; Aitio, A.; Lee, S.; Yeh, Y. T.; Steinbacher, F.; Khan, M. U.; Lee, J. W.; Siegel, J. B.; Stefanopoulou, A. G.; et al. 2021. The challenge and opportunity of battery lifetime prediction from field data. *Joule*, 5(8): 1934–1955.

Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; and Chen, X. 2024. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nature Communications*, 15(1): 4332.

Wang, Z.; Yang, F.; Xu, Q.; Wang, Y.; Yan, H.; and Xie, M. 2023. Capacity estimation of lithium-ion batteries based on data aggregation and feature fusion via graph neural network. *Applied Energy*, 336: 120808.

Zhang, M.; Yang, D.; Du, J.; Sun, H.; Li, L.; Wang, L.; and Wang, K. 2023. A review of SOH prediction of Li-ion batteries based on data-driven algorithms. *Energies*, 16(7): 3167.

Zhang, Y.; Tang, Q.; Zhang, Y.; Wang, J.; Stimming, U.; and Lee, A. A. 2020. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. *Nature communications*, 11(1): 1706.

Zhao, J.; Wang, Z.; Wu, Y.; and Burke, A. F. 2025. Predictive pretrained transformer (PPT) for real-time battery health diagnostics. *Applied Energy*, 377: 124746.

Zhou, K. Q.; Qin, Y.; and Yuen, C. 2024. Graph neural network-based lithium-ion battery state of health estimation using partial discharging curve. *Journal of Energy Storage*, 100: 113502.

Zhu, J.; Wang, Y.; Huang, Y.; Bhushan Gopaluni, R.; Cao, Y.; Heere, M.; Mühlbauer, M. J.; Mereacre, L.; Dai, H.; Liu, X.; et al. 2022. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nature communications*, 13(1): 2261.