Machine Learning with Applications 22 (2025) 100729

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

Machine-learning based Li-Ion Cell state prediction using Impedance spectroscopy

Carl Philipp Klemm [a,b,*], Till Frömling [c]

[a] rhd instruments GmbH, Otto-Hesse-Straße 19, 64293 Darmstadt, Germany
[b] Technische Universität Darmstadt, Alarich-Weiss-Straße 2, 64287 Darmstadt, Germany
[c] Forschungszentrum Jülich GmbH, Institute of Energy Technologies Fundamental Electrochemistry, Wilhem-Johnen-Straße 802, 52428 Jülich, Germany

ARTICLE INFO

Dataset link: https://uvos.xyz/git/uvos/CoincellHellExpiramentData, https://github.com/orgs/EIS-KISS/repositories

Keywords:
Machine learning
Lithium ion battery
Electrochemical impedance spectroscopy

ABSTRACT

Accurate and reliable monitoring of battery state parameters is crucial for ensuring optimal battery performance, safety, and lifetime. Existing methods have limitations, such as requiring modeling of each degradation mechanism involved or relying on direct measurement techniques that impose restrictions on field studies or end-user use. In this paper, we propose a machine learning-based approach that combines the strengths of electrochemical impedance spectroscopy (EIS) and machine learning algorithms to predict battery state parameters. We have developed an efficient prediction system that can learn from EIS data and accurately predict battery state parameters. Our approach is trained on an open dataset comprising of over 30,000 spectra, generated using an automated measurement technique that outperforms current machine learning-based models, particularly in terms of generalization across different cells and measurement setups.

1. Introduction

The increasing use of battery-powered devices and the need to electrify vehicles has made it urgent to monitor battery state parameters accurately and reliably. Real-time tracking of state of health (SOH), state of charge (SOC), and temperature are critical for ensuring optimal performance, safety, and lifetime of batteries. As the demand for high-performance and sustainable energy storage solutions continues to grow, the ability to accurately predict battery state parameters has become a bottleneck in developing advanced battery management systems. Existing methods rely on modeling the underlying physics of degradation, which requires modeling each degradation mechanism involved (Reniers et al., 2019), or on direct measurement techniques, which are more restrictive in terms of field studies or end-user application. As an alternative, data-driven systems have been proposed. However, their performance in generalizing to situations significantly different from those in the training set has not been evaluated (Babaeiyazdi et al., 2023; Buchicchio et al., 2023; Liu et al., 2024; Wu et al., 2020; Zhang et al., 2020). The simple and reliable acquisition of battery state parameters could be used to improve manufacturing protocols for quality assurance, accelerate battery research, provide notice for timely maintenance of end-user devices This could also improve the economics of identifying which cells that may have a useful second life after decommissioning (Montes et al., 2022). Electrochemical impedance spectroscopy (EIS) has emerged as a powerful tool for characterizing battery behavior, offering insights into the underlying electrochemical processes that govern battery performance and degradation. However, analysis of data generated by EIS measurements poses significant challenges, as most require manual analysis for interpretation (Hu et al., 2023; Liu et al., 2024; Patrizi et al., 2024; Zhu et al., 2024). Machine learning algorithms have demonstrated remarkable capabilities in pattern recognition and prediction. By combining the strengths of machine learning and EIS, we have developed an efficient prediction system that can learn from EIS data and accurately predict battery state parameters. We have compared it to two baselines: the method proposed by Zhang et al. (2020), which we consider to be an excellent representation of the current state of modeling in this field; and a method trained with XGBoost (Chen & Guestrin, 2016). Zhang et al. used a GPR model which achieved excellent predictive power. The method and is validated against a variety of cell temperatures. This method has the advantages of the increased degree of explainability of GPR models as well as the inherent ability to estimate prediction tolerance (Rasmussen & Williams, 2005). Although better prediction metrics have previously been achieved on the dataset used Zhang et al. have been previously achieved using CNNs (Babaeiyazdi et al., 2023), a better combination of metrics, explainability, and verified robustness against changes in battery state had not previously been achieved. In pursuit of this research and to enable further development in this area,

* Corresponding author at: rhd instruments GmbH, Otto-Hesse-Straße 19, 64293 Darmstadt, Germany.
E-mail addresses: klemm@rhd-instruments.de (C.P. Klemm), t.froemling@fz-juelich.de (T. Frömling).

https://doi.org/10.1016/j.mlwa.2025.100729
Received 18 July 2025; Received in revised form 21 August 2025; Accepted 30 August 2025
Available online 8 September 2025

Acronyms

R^2	Coefficient of determination
CNN	Convolutional neural network
EIS	Electrochemical impedance spectroscopy
GPR	Gaussian process regression
KISS	KI-basiertes Sortieren von Spektren
KissA	Dataset comprised of 90% of the measurements taken from cells 1–3, 6 and 8–15 of the KISS dataset
KissB	Dataset comprised of cells 0, 4, 7 and 16 of the KISS dataset
KissVal	Dataset comprised of 10% of the measurements taken from cells 1–3, 6 and 8–15 of the KISS dataset that are not in KissA
PFI	Permutation feature importance
RMSE	Root mean square error
SHAP	Shapley additive explanations
SOC	State of charge
SOH	State of health
ZangA	Dataset comprised of cells cycled at room temperature and 45 °C by Zhang et al.
ZangB	Dataset comprised of cells cycled at 35 °C by Zhang et al.

we have created an open dataset of over 30,000 spectra, generated using automated measurement setup for battery cell condition variation. This dataset provides a valuable resource for the research community, particularly for validating advanced battery state prediction systems and accelerating the development of energy storage solutions.

2. Datasets

Five datasets were used to train, validate and test the methods considered in this paper. First, the dataset as measured by Zhang et al. (2020) was used. This dataset was split into two parts in the same way as Zhang et al. did when training and validating their GPR model for multi-temperature prediction. These datasets were named ZangA and ZangB. Secondly, a dataset was generated by cycling commercial LR2032 cells under varying conditions, see Section 5.1. This dataset was also split into two by separating it into cells 1–3, 5, 6, 8–15 and 0, 4, 7, 16, see Table A.3. The first set of cells was further subdivided by random sampling into two datasets with a 90/10 percent split. These three datasets were labeled KissA, KissVal and KissB. The KissVal dataset was only used during network architecture selection for the convolutional neural network CNN model, while ZangA, ZangB, KissA and KissB were used to train the methods described below.

3. Results

We consider the task of capacity estimation. In this case, the user needs to know the capacity of the cells without knowing the SOC, SOH or temperature of his cell. Four models were trained to perform this task:

- A reproduction of the GPR model by Zhang et al. trained on the ZangA dataset.
- The GPR model trained on the KissA dataset.
- Our CNN model trained on the ZangA dataset.
- Our CNN model trained on the KissA dataset.
- A XGBoost model trained on the ZangA dataset.
- A XGBoost model trained on the KissA dataset.

The models were validated by testing them on the corresponding test dataset. Finally, models were examined for generalization by testing them against the test portion of the opposite dataset. This allowed us to check for the presence of generalization across different cells and measurement setups.

3.1. In-distribution performance

As can be seen in Figs. 1, and 2, when evaluated on the same distribution as the training data — in this case, ZangA for training and ZangB for testing — all methods achieve respectable and broadly similar results. The coefficient of determinations (R^2s) is 0.959, 0.859 and 0.794 and the root mean square error (RMSE) of 0.00078, 0.00106 and 0.002 for the CNN, XGBoost and GPR models, respectively. These results are comparable to the $R^2 = 0.953$ reported by Babaeiyazdi et al. using a deep neural network on the same dataset (Babaeiyazdi et al., 2023).

With a larger number of cells, a greater variation in SOC and temperature, as well as more noise, the KissA and KissB datasets present a significantly greater challenge to the proposed regression methods. This greater challenge results in greater separation between the results, as can be seen in Fig. 3. The R^2 and RMSE also demonstrate this with the values of 0.8071, 0.68, 0.5930 and 0.00452, 0.00134, 0.00688 respectively.

3.2. Cross-dataset generalization

So far, all methods show acceptable performance in both tasks. However, when it comes to actual application it is of course of utmost importance that the models generalize to new cells and measurement setups. While the holdout method used to generate the KissB dataset has been shown to be effective in determining when overfitting has occurred (Roelofs et al., 2019; Steyerberg et al., 2003), external validation of models using independently acquired datasets remains the gold standard for predicting generalization in the final application environment (Goodfellow et al., 2016). To this end we externally validated all four models using opposite datasets.

As can be seen in Fig. 4 none of the models retained any predictive power when trained on the ZangA dataset and tested against the KissB dataset with the R^2 of all combinations being below zero.

When trained on the KissA dataset and tested against the ZangB dataset the GPR and XGBoost models again perform poorly, see Fig. 5, with a R^2 below zero. However, the CNN model retained predictive power with an R^2 of 0.212 and an RMSE of 0.0325.

As is intuitive and suggested by the data scaling law (Hestness et al., 2017), when trained simultaneously on the KissA and ZangB datasets the CNN model gains overall better performance on the combined testing dataset with an R^2 of 0.8413 compared to 0.6505, this can also be seen in Fig. 6. This suggests that the well-known performance scaling rule with increasing dataset size and diversity remains effective (Rajput et al., 2023; Zhang & Ling, 2018). However, this single sample is insufficient to prove this in the general case.

4. Discussion

The failure of all models to generalize when trained on the ZangA dataset shows, that while the holdout method is effective at predicting generalization across examples from a similar population, generalization across different cells and measurement setups requires more rigorous validation. When trained on a larger, more varied dataset, the deep CNN begins to generalize while the performance of the GPR and XGBoost models remains poor. Since the KissA dataset has a significantly higher noise floor, the difference could potentially be due to the well-known regularization effect of noise when applied to the input features (Noh et al., 2017), or to the more varied distribution of examples in the larger dataset. To further understand these differences

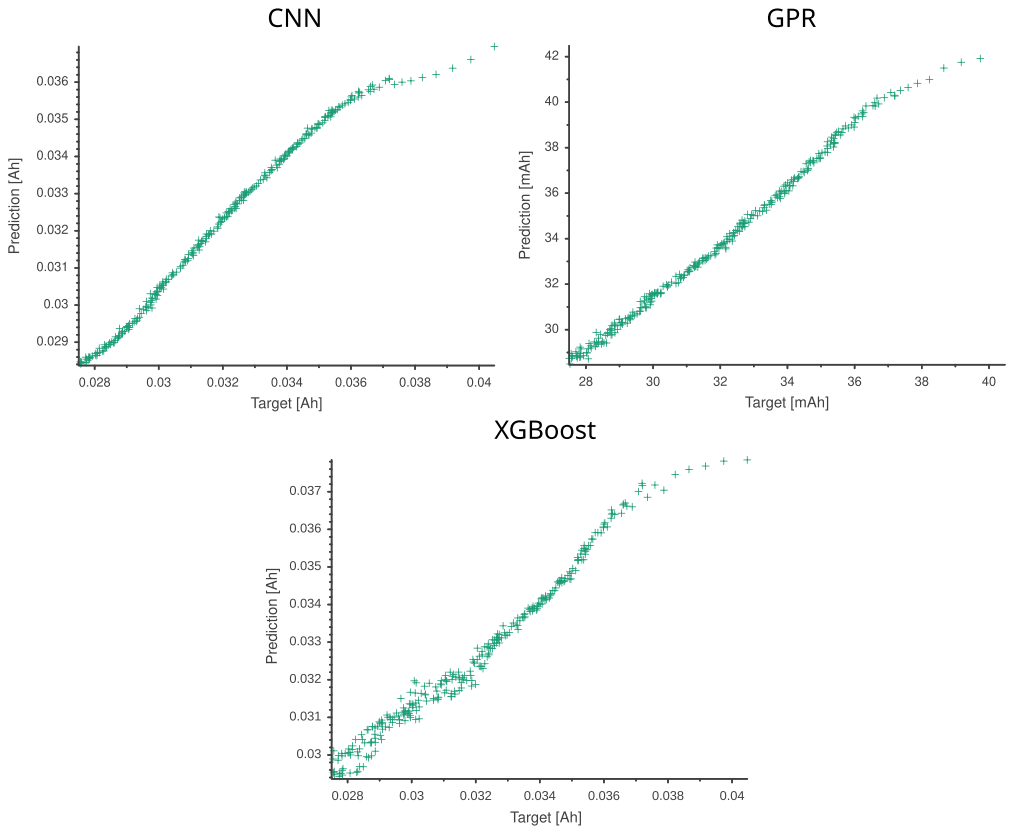

Fig. 1. ZangB results from the CNN, GPR and XGBoost models trained on the ZangA dataset.

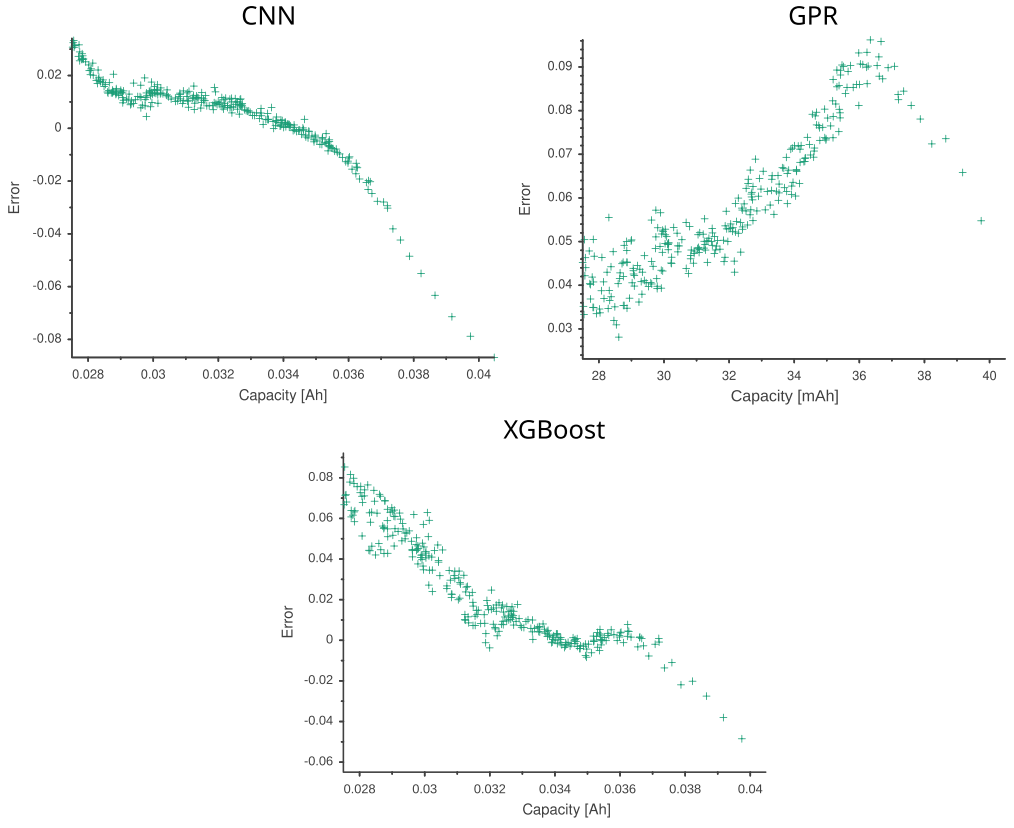

Fig. 2. Relative error of the CNN, GPR and XGBoost models trained on the ZangA dataset.

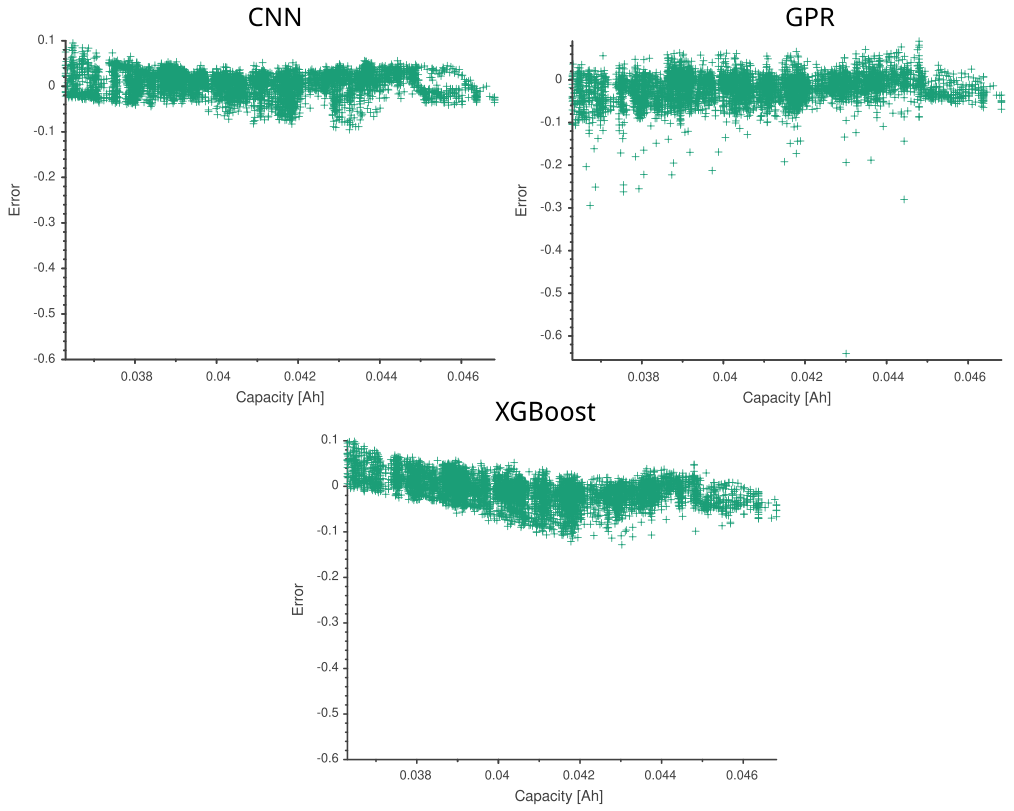

Fig. 3. Relative error of the CNN, GPR and XGBoost models on the KissB dataset when trained on the KissA dataset.

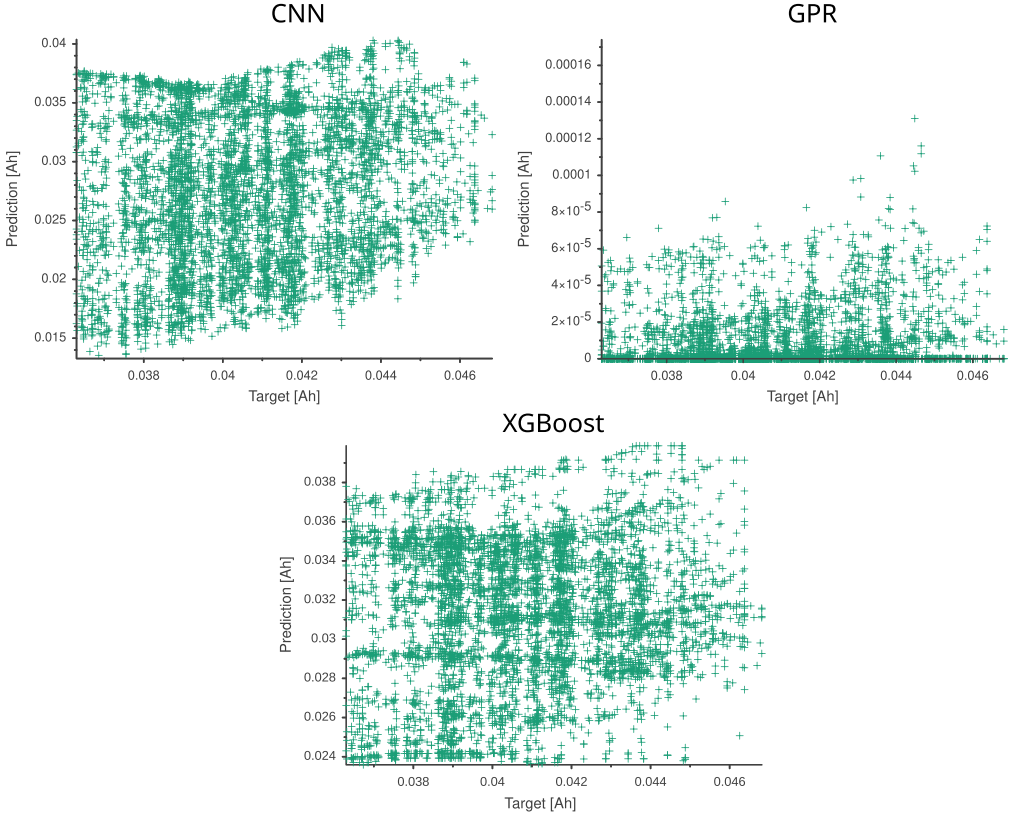

Fig. 4. KissB results from the CNN, GPR and XGBoost models trained on the KissA dataset.

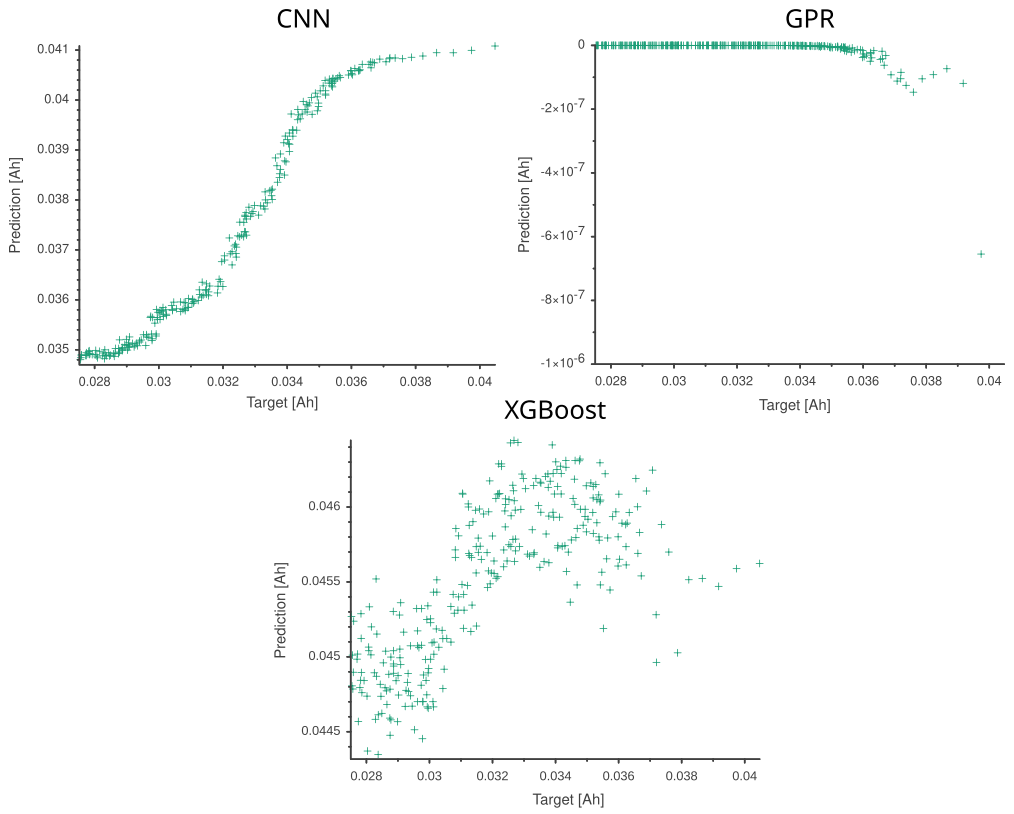

Fig. 5. KissB results from the CNN, GPR XGBoost models trained on the KissA dataset.

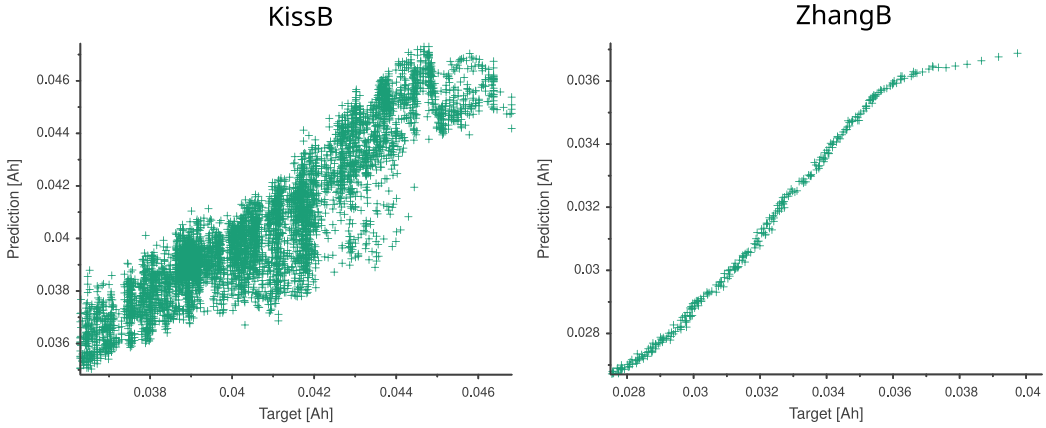

Fig. 6. Resulting predictions by the CNN model trained on the combined KissA + ZangA dataset tested on KissB and ZangB.

in generalization we analyzed the feature importance of the CNN models using a variation of permutation feature importance (PFI), as well as SHAP (Lundberg & Lee, 2017), see Section 5.3.

As first discovered by Zhang et al. for the ZangA dataset, it is sufficient to observe only the low frequency real part of the impedance spectra, see Fig. 7.

Targeted drop-column importance was used to determine whether the real part of the 1 Hz feature alone could predict the capacity of a cell in the ZangA dataset. Indeed, a R^2 of 0.731 was obtained by training the CNN on input vectors zero padded to size 10 using only this feature. Since the 1 Hz real part of the resistance is strongly correlated with the DC internal resistance of the cell, without electrochemical changes due to long-range ion diffusion, it is reasonable to assume that the predictive power of the methods trained on the ZangA dataset is largely based on the observation that the DC cell internal resistance is lower at higher capacity (Morse & Sargent, 1911). Therefore, it is not entirely surprising that when these methods are then applied to a different cell with a different DC internal resistance, that the models fail to correctly predict the capacities of these new cells.

Different behavior is observed in the CNN that was trained on the KissA dataset, which also exhibits a degree of generalization over the ZangB dataset. As can be seen in Figs. 8 and 9, a wide variety of features is used to make the prediction in this case with particular emphasis on the medium-high frequency imaginary part of the spectra. This suggests that a more complex relationship exists between the features and the output parameters compared to previous models. At these high frequencies, processes such as charge transfer between the electrode and the collectors, the ionic conductivity of the electrolyte, and the

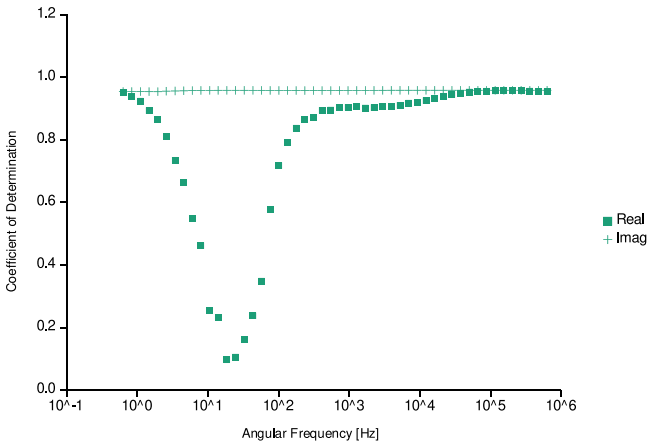

Fig. 7. Feature importance of the CNN network trained on the ZangA dataset.

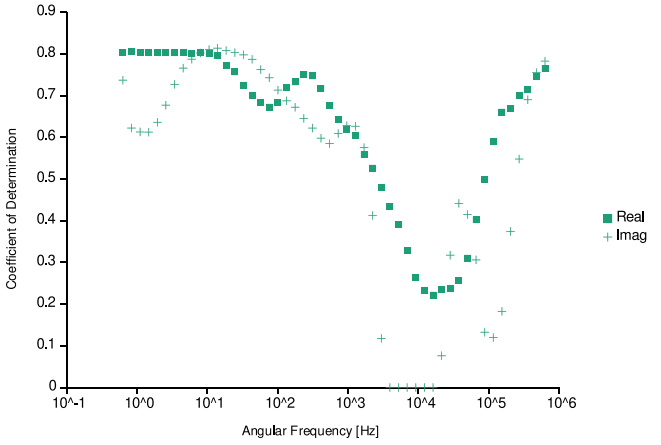

Fig. 8. Permutation feature importance of the CNN network trained on the KissA dataset.

electronic conductivity in the carbon black contribute to the electrical response. These are thus the candidate processes that have an impact on the models final prediction.

4.1. Limitations

Although our model shows promising performance, some key limitations remain. While our results demonstrate the potential for generalization across certain parameters, further study is required to validate generalization across parameters not covered by the available datasets, such as cells with larger capacities or different chemistries. Although we have ragained interpetability by applying PFI and SHAP, the GPR model as proposed by Zhang et al. has the advantage of being simpler to interpret and providing intrinsic uncertainty estimates. Future study on applying interpritable layers to CNN models used for this task may be warrented.

4.2. Applications

The presented methodology enables rapid, non-invasive capacity estimation from a single EIS sweep. The approach is suitable for on-board BMS integration, high-throughput laboratory screening or quality control. We hope the open dataset and accompanying code will provide a benchmark for future research on data-driven battery diagnostics.

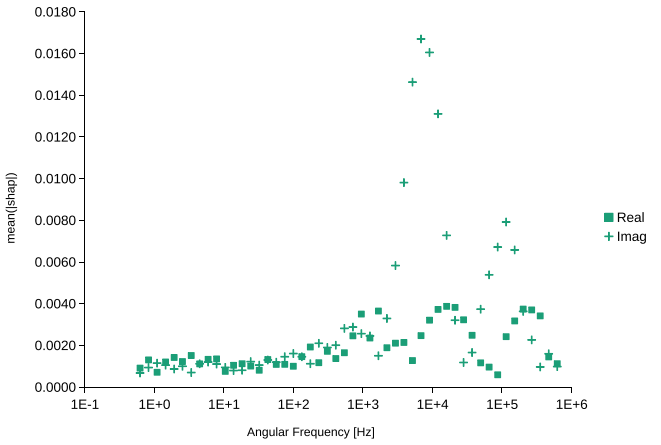

Fig. 9. Mean absolute SHAP values of the CNN network trained on the KissA dataset.

Table 1
Networks considered.

Network	Varied hyperparameters	Range
Fully connected encoder network	Encoder layers	[5, 25]
	Downsample gap	[0, 10]
	Extra layers at end	[0, 5]
	Batch normalization (Ioffe & Szegedy, 2015)	True/False
	Dropout (Srivastava et al., 2014)	{0, 0.1, 0.2}
Partially convolutional encoder–decoder network.	Decoder, Encoder layers	[5, 25]
	Downsample gap	[0, 10]
	Extra layers at end	[0, 5]
	Batch normalization	True/False
	Dropout	{0, 0.1, 0.2}
Partially convolutional network, with skip connections as proposed by He et al. (2015)	Layer count	[10, 100]
	Downsample gap	[1, 5]
	Batch normalization	True/False
	Dropout	{0, 0.1, 0.2}

Table 2
XGBoost hyperparameters, best combination of parameters highlighted in bold.

Parameter	Range
Maximum depth	{2, 4, **6**, 8, 10}
Minimum child weight	{1, **3**, 6}
Observation subsampling	{0.2, 0.6, 0.7, 0.8, **0.9**, 1.0}
Feature subsampling	{0.01, 0.05, 0.1, **0.2**, 0.3, 0.5, 0.6, 0.7, 1}
L1 regularization	{**0**, 0.1}
L2 regularization	{0.3, **0.5**, 0.8, 1.0}

5. Methods

5.1. Dataset generation

The measurement of 17 commercial Voltcraft lr2032[1] cells was conducted using the Biologic SP-240 with a single channel and 3 A power amplifier. The single channel was multiplexed through an array of relays using six tiles of the setup shown in Fig. B.10. 50 logarithmically spaced samples were taken ranging from 100 mHz to 100 kHz. The SOC, SOH and temperature parameters were varied according to Table A.3. The measurements were randomly distributed across cycles, so that all varied parameters were sampled in a uniformly at random. For cells that were thermally cycled and SOC cycled, ten thermal cycles

[1] UPC-A: 4064161214337.

Table A.3
Cells in the KISS dataset.

KISS cell	Thermal cycles	Temperature range [°C]	SOC cycled	Min SOC	Max SOC	Charge rate
0	Y	35–55	Y	0	1	0.85 C–1 C
1	Y	35–55	Y	0	1	0.85 C–1 C
2	Y	35–55	Y	0	1	0.85 C–1 C
3	Y	35–55	Y	0	1	0.85 C–1 C
4	Y	35–55	N	0.33	0.33	0.85 C–1 C
5	Y	35–55	N	0.33	0.33	0.85 C–1 C
6	Y	35–55	N	0.66	0.66	0.85 C–1 C
7	Y	35–45	N	0.66	0.66	0.85 C–1 C
8	Y	35–45	Y	0	1	0.85 C–1 C
9	Y	35–45	Y	0	1	0.85 C–1 C
10	Y	35–45	Y	0	1	0.85 C–1 C
11	N	35–35	Y	0	1	0.85 C–1 C
12	N	35–35	Y	0	1	0.85 C–1 C
13	N	35–35	Y	0	1	0.85 C–1 C
14	N	45–45	Y	0	1	0.85 C–1 C
15	N	45–45	Y	0	1	0.85 C–1 C
16	N	45–45	Y	0	1	0.85 C–1 C

were performed for each SOC cycle. The charge rate was 0.85 C for all samples, except for every 10th cycle in which 1 C was used. Charge rates are relative to the average capacity of the cells at the beginning of the test.

5.2. Model architecture selection

For our CNN model, the architecture was selected using a grid search along a set of hyperparameters. These candidate models were divided into three overarching candidate architectures and varied parameters (see Table 1):

The networks were trained on the KissA dataset and selected based on the R^2 score obtained on the KissVal dataset. The configuration with 50 layers, a downsample gap of 2, batch normalization set to true, and a dropout rate of 0.1 performed best against the KissVal dataset and was thus selected. The parameters were optimized using the popular AdamW algorithm (Loshchilov & Hutter, 2019) with a learning rate of $5e{-}3$, a weight decay of 0.02, $\beta_1 : 0.90$ and $\beta_2 : 0.99$.

For our XGBoost model, we also performed a grid search to select the best hyperparameters (see Table 2):

5.3. Permutation feature importance

In the general case, PFI is determined by an arbitrary function $f : D \to S$, where $D_{i,\{0\ldots n\}}$ are n uncorrelated features of an example D_i in a dataset D which contains K examples. The sensitivity of the function f to $D_{\{1,\ldots,K\},k}$ is estimated by observing the change in $\frac{1}{K}\sum_{i=0}^{K} f(D_i)$ when features $D_{\{1,\ldots,K\},k}$ are shuffled along i (Altmann et al., 2010; Breiman, 2001).

In the present case, $f(D) = s(g(D))$, where $g(D)$ regression model learnt by the CNN, s is the scoring function R^2 and D is our dataset of K impedance spectra. However, this general approach is poorly applicable, because in an impedance spectrum, neighboring features are strongly correlated. Additionally, for a small dataset D, the current values for each feature $D_{\{1,\ldots,K\},k}$ may not be representative of the possible values for the given feature encountered in the broader case.

For these reasons we made two modifications to the established method of PFI: First, instead of shuffling the features k along the dataset examples i we replace $D_{k,i}$ with a random variable $\{Z \in \mathbb{R} \mid \min(D_{\{1,\ldots,K\},k}) \leq Z \leq \max(D_{\{1,\ldots,K\},k})\}$. Secondly, instead of replacing a single feature at a time we use a sliding window of width $w \in \{1, 3, 5, 7, 9, \ldots\}$ and stride 1, where all features within the window are replaced by random variables. We call the examples of this thus modified dataset $Dp_{w,i}$, see Fig. C.11 for examples. The width of the window w is then optimized such that $\left|\frac{\delta}{\delta w}\frac{1}{K}\sum_{i=0}^{K}\max\{f(Dp_{w,i}), 0\}\right|$ begins to decrease with increasing window width w. For our model and dataset this occurred at $w = 5$. Next, $s(g)$ is computed with the window centered around each feature i. The result is plotted in Figs. 8 and 7.

6. Resource availability

All code related to this article is open source under GPLv3, LGPLv3 or MIT license, depending on the subproject.

- The code implementing the networks, training and validation can be found at https://github.com/EIS-KISS/TorchKissAnnPostdict.
- The code implementing the XGBoost model can be found at https://github.com/EIS-KISS/XGBoost.
- The code applying the shap explainer can be found at https://github.com/EIS-KISS/sharp_analysis.
- The code used to translate the datasets can be found at https://github.com/EIS-KISS/dataformaters
- The code used to handle, filter and combine the datasets can be found at https://github.com/EIS-KISS/KissDatasetGenerator
- The source data for the KissA and KissB datasets is available at https://uvos.xyz/git/uvos/CoincellHellExpiramentData.

CRediT authorship contribution statement

Carl Philipp Klemm: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Visualization. **Till Frömling:** Supervision, Project administration, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carl Klemm reports financial support was provided by rhd-instruments. Carl Klemm reports financial support was provided by Hessisches Ministerium für Digitalisierung und Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the support of this work in form of the grant 21_0005_2A from the German State of Hesse as part of the Distr@l program. This work was supported by rhd instruments GmbH & Co. KG,[2] as part of the project KISS.

[2] https://www.rhd-instruments.com.

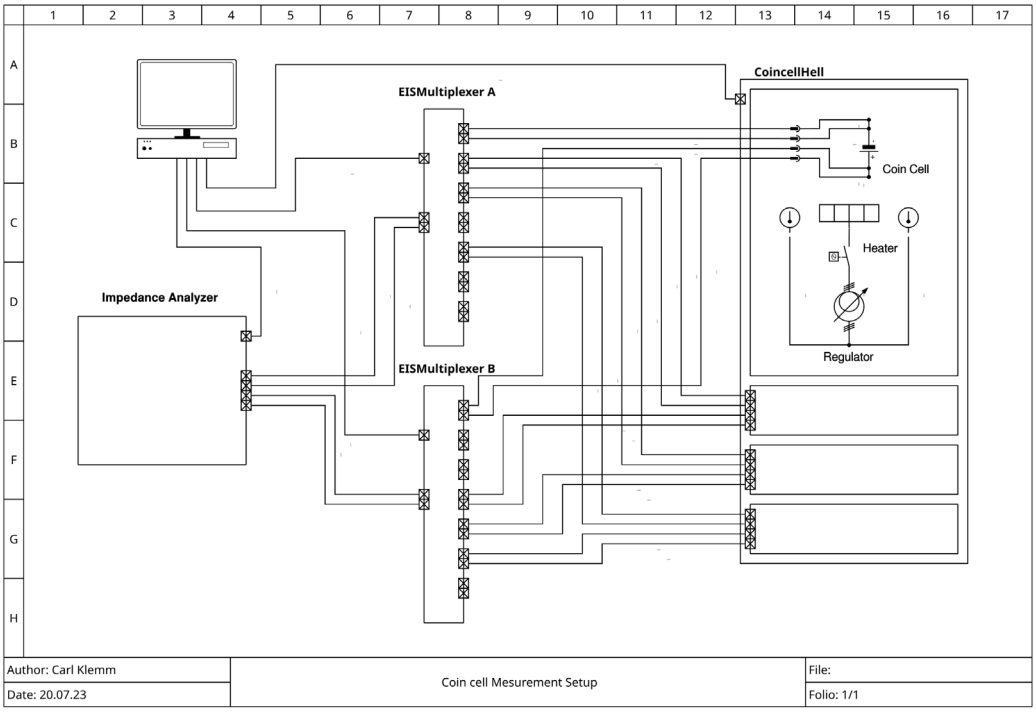

Fig. B.10. Measurement setup for the KISS dataset (repeated 6 times).

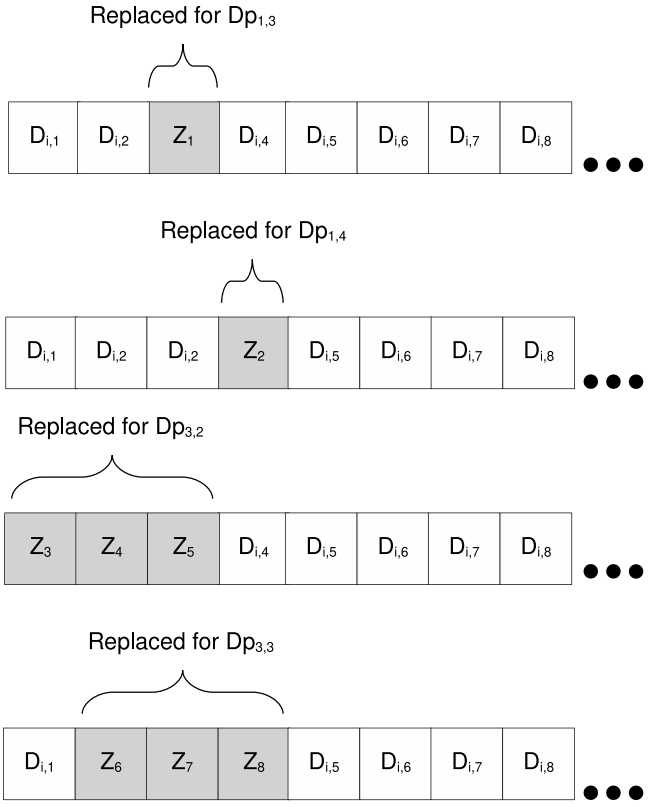

Fig. C.11. Examples of dataset permutation.

Appendix A. Kiss dataset structure

See Table A.3.

Appendix B. Measurement setup

See Fig. B.10.

Appendix C. PFI

See Fig. C.11.

Data availability

Data is available at https://uvos.xyz/git/uvos/CoincellHellExpiramentData, code to the varoius utilities used is available at https://github.com/orgs/EIS-KISS/repositories.

References

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, *26*, http://dx.doi.org/10.1093/bioinformatics/btq134.

Babaeiyazdi, I., Rezaei-Zare, A., & Shokrzadeh, S. (2023). Transfer learning with deep neural network for capacity prediction of Li-ion batteries using eis measurement. *IEEE Transactions on Transportation Electrification*, *9*, http://dx.doi.org/10.1109/tte.2022.3170230.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, http://dx.doi.org/10.1023/A:1010933404324.

Buchicchio, E., Angelis, A. D., Santoni, F., Carbone, P., Bianconi, F., & Smeraldi, F. (2023). Battery soc estimation from eis data based on machine learning and equivalent circuit model. *Energy*, *283*, http://dx.doi.org/10.1016/j.energy.2023.128461.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). New York, NY, USA: Association for Computing Machinery, ISBN: 9781450342322, http://dx.doi.org/10.1145/2939672.2939785.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press, http://www.deeplearningbook.org.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. CoRR, abs/1512.03385. URL http://arxiv.org/abs/1512.03385.

Hestness, J., Narang, S., Ardalani, N., Diamos, G. F., Jun, H., Kianinejad, H., Patwary, M. M. A., Yang, Y., & Zhou, Y. (2017). Deep learning scaling is predictable, empirically. http://dx.doi.org/10.48550/arXiv.1712.00409, ArXiv.

Hu, W., Peng, Y., Wei, Y., & Yang, Y. (2023). Application of electrochemical impedance spectroscopy to degradation and aging research of lithium-ion batteries. *The Journal of Physical Chemistry C*, *127*, http://dx.doi.org/10.1021/acs.jpcc.3c00033.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. URL https://arxiv.org/abs/1502.03167.

Liu, Y., Li, Q., & Wang, K. (2024). Revealing the degradation patterns of lithium-ion batteries from impedance spectroscopy using variational auto-encoders. *Energy Storage Materials*, *69*, http://dx.doi.org/10.1016/j.ensm.2024.103394.

Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. URL https://arxiv.org/abs/1711.05101.

Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. Curran Associates, Inc., URL http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.

Montes, T., Etxandi-Santolaya, M., Eichman, J., Ferreira, V. J., Trilla, L., & Corchero, C. (2022). Procedure for assessing the suitability of battery second life applications after ev first life. *Batteries*, *8*, http://dx.doi.org/10.3390/batteries8090122.

Morse, H. W., & Sargent, L. W. (1911). The internal resistance of the lead accumulator. *Proceedings of the American Academy of Arts and Sciences*, *46*(21), 589–612, URL http://www.jstor.org/stable/20022681.

Noh, H., You, T., Mun, J., & Han, B. (2017). Regularizing deep neural networks by noise: Its interpretation and optimization. CoRR, abs/1710.05179. URL http://arxiv.org/abs/1710.05179.

Patrizi, G., Canzanella, F., Ciani, L., & Catelani, M. (2024). Life cycle analysis of lithium batteries for smart grids using electrochemical impedance spectroscopy data. In *2024 IEEE international workshop on metrology for Industry 4.0 and IoT* (pp. 298–303). http://dx.doi.org/10.1109/MetroInd4.0IoT61288.2024.10584119.

Rajput, D., Wang, W.-J., & Chen, C.-C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, *24*, http://dx.doi.org/10.1186/s12859-023-05156-9.

Rasmussen, C. E., & Williams, C. K. I. (2005). *Gaussian processes for machine learning*. The MIT Press, ISBN: 9780262256834, http://dx.doi.org/10.7551/mitpress/3206.001.0001.

Reniers, J. M., Mulder, G., & Howey, D. A. (2019). Review and performance comparison of mechanical–chemical degradation models for lithium-ion batteries. *Journal of the Electrochemical Society*, *166*, http://dx.doi.org/10.1149/2.0281914jes.

Roelofs, R., Shankar, V., Recht, B., Fridovich-Keil, S., Hardt, M., Miller, J., & Schmidt, L. (2019). A meta-analysis of overfitting in machine learning. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, & R. Garnett (Eds.), *vol. 32*, *Advances in neural information processing systems*. Curran Associates, Inc., URL https://proceedings.neurips.cc/paper_files/paper/2019/file/ee39e503b6bedf0c98c388b7e8589aca-Paper.pdf.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958.

Steyerberg, E. W., Bleeker, S. E., Moll, H. A., Grobbee, D. E., & Moons, K. G. (2003). Internal and external validation of predictive models: A simulation study of bias and precision in small samples. *Journal of Clinical Epidemiology*, *56*(5), 441–447. http://dx.doi.org/10.1016/S0895-4356(03)00047-7, URL https://www.sciencedirect.com/science/article/pii/S0895435603000477.

Wu, B., Widanage, W. D., Yang, S., & Liu, X. (2020). Battery digital twins: Perspectives on the fusion of models data and artificial intelligence for smart battery management systems. *Energy and AI*, *1*, http://dx.doi.org/10.1016/j.egyai.2020.100016.

Zhang, Y., & Ling, C. (2018). A strategy to apply machine learning to small datasets in materials science. *Npj Computational Materials*, *4*, http://dx.doi.org/10.1038/s41524-018-0081-z.

Zhang, Y., Tang, Q., Zhang, Y., Wang, J., Stimming, U., & Lee, A. A. (2020). Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. *Nature Communications*, *11*, http://dx.doi.org/10.1038/s41467-020-15235-7.

Zhu, H., Evans, T. A. P., Weddle, P. J., Colclasure, A. M., Chen, B.-R., Tanim, T. R., Vincent, T. L., & Kee, R. J. (2024). Extracting and interpreting electrochemical impedance spectra (eis) from physics-based models of lithium-ion batteries. *Journal of the Electrochemical Society*, *171*, http://dx.doi.org/10.1149/1945-7111/ad4399.