2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Scenario-Aware Machine Learning Pipeline for Battery Lifetime Prediction*

Huang Zhang[1,2], Faisal Altaf [1] and Torsten Wik [2]

***Abstract*— Advanced machine learning (ML) models have been developed for battery lifetime prediction in different use cases at all stages of a battery's life. As the first step to enable the transferability of ML models for battery lifetime prediction across multiple use cases, a scenario-aware machine learning pipeline is proposed, in which two feature engineering methods that have been able to generate input features with outstanding predictive power are used to learn the best ML model for battery lifetime prediction in a chosen usage scenario. The experimental results show that the histogram-based feature engineering method is able to generate input features with predictive power generalized across two usage scenarios (i.e., identical cycling and protocol cycling). Thus, to enable transferability of ML models for battery lifetime prediction across different scenarios, and even battery chemistries, this histogram-based feature engineering method will be further investigated together with online fine-tuning strategies.**

I. INTRODUCTION

In order to decarbonize transportation systems in Europe, EU-wide CO_2 emission standards for light-duty vehicles (Regulation 2019/631) and heavy-duty vehicles (Regulation 2019/1242) are set in 2019, in which emission reduction goals require technologies to enable significant transitions from internal combustion engine vehicles to zero-emission vehicles [1]. As one of the key enabling technologies, lithium-ion batteries are widely adopted by major automotive companies due to their decreasing costs, high energy densities, and long lifetimes [2]. To further improve lithium-ion battery technology, battery lifetime prediction has gained considerable academic attention and research interest in recent years. As a result, a large number of battery lifetime prediction methods have been proposed for various use cases at all stages of a battery's life [3].

More recently, spurred by the ever-increasing availability of battery data and the success of machine learning (ML), substantial efforts have been made in data-driven methods for battery lifetime prediction. As a matter of fact, there are many different use cases for these battery lifetime prediction ML models spanning battery design, manufacturing, usage, and repurposing stages, and the feature engineering method is generally specific to each use case. To exemplify, 7 use cases for battery lifetime prediction are listed here. Firstly, the elemental compositions of the liquid electrolytes were used as features to develop ML models that assist the design of high-performance electrolytes [4]. Secondly, experimental variables were used as features to develop ML models for the design of electrode manufacturing processes [5]. Thirdly, features extracted from discharge data were used to develop ML models for the evaluation of intrinsic cell-to-cell variations in quality control [6] [7]; Fourthly, features extracted from half discharge data were used to develop ML models for battery grading before its usage [8]; Fifthly, features extracted from early degradation data were used to develop ML models for protocol optimization (e.g., formation protocol optimization [9], fast-charging protocol optimization [10] [11]); Sixthly, the time spent in specific voltage, current and temperature ranges were used to develop ML models for battery health estimation and prediction in the battery management system (BMS) [12] [13] [14]; Lastly, features extracted from incremental capacity curves were used to develop ML models for second-life battery health estimation and prediction at the repurposing stage [15].

Different ML models have proven to give highly accurate lifetime prediction in the aforementioned use cases with possibly quantified uncertainty. However, the input features generated in the feature engineering process in one use case can vary from another. Therefore, the goal of this work is to find feature engineering methods that are capable of generating input features with generalized predictive power across different use cases. As a result, the transferability of ML models for battery lifetime prediction can be enabled across multiple use cases at different stages of a battery's life.

Specifically, our **key results and contributions** are summarized as follows:

- We first classify typical use cases for battery lifetime prediction into three scenarios based on battery usage profiles, i.e., identical cycling, protocol cycling, and dynamic cycling (see Table I), then a scenario-aware machine learning pipeline (see Fig. 1) is proposed to automate the process of selecting the best feature engineering method for developing the best ML model in a chosen battery usage scenario.
- Gaussian process regression using input features generated by the histogram-based feature engineering method is found to be the best choice in two of the usage scenarios (i.e., identical cycling and protocol cycling). Thus, to enable the transferability of ML models for battery lifetime prediction across multiple use cases, this histogram-based feature engineering method will be further investigated.

*This work was supported by Volvo Group

[1] Department of Electromobility, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden huang.zhang@volvo.com, faisal.altaf@volvo.com

[2] Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden huangz@chalmers.se, torsten.wik@chalmers.se

The remaining of the paper is organized as follows: Section II classifies typical use cases for battery lifetime prediction into three scenarios and then introduces the scenario-aware pipeline; Section III validates the effectiveness of our proposed scenario-aware pipeline for battery lifetime prediction in two scenarios; Section IV draws conclusions and recommends future work.

II. Scenario-Aware Machine Learning Pipeline

In the literature, many machine learning (ML) models have been developed to provide accurate and reliable battery lifetime prediction in various use cases. In Table I, we classify typical use cases in the literature into three scenarios based on their battery usage profiles, i.e., identical cycling, protocol cycling, and dynamic cycling.

- Identical cycling: Batteries are repeatedly cycled under identical or nearly identical conditions to quantify their intrinsic cell-to-cell variations, such as battery lifetime spread, capacity spread, etc.
- Protocol cycling: Batteries are repeatedly cycled under various cycling protocols to characterize the impact of different protocols (e.g., formation protocols or fast-charging protocols) on battery health degradation and lifetime.
- Dynamic cycling: Batteries are cycled under varying conditions to characterize battery aging in real-world applications.

Based on use case classification results for battery lifetime prediction in Table I, a scenario-aware machine learning pipeline that automates the process of producing the best ML model for each scenario is proposed in Fig. 1.

A. Feature engineering

The main objective of feature engineering is to reduce data dimensionality, improve model performance, and enhance model interpretability [13]. To develop ML models for battery lifetime prediction in various use cases (see Table I), different feature engineering methods have been proposed in the literature. Among them, two feature engineering methods have been able to generate input features with excellent predictive power while retaining a good level of interpretability. As listed in Table II, the first feature engineering method that extracts features from discharge data of the first 100 cycles was initially proposed by Severson et al. [16], and this MIT 6-feature set is selected as the first feature set in this work. In contrast, the second feature engineering method that extracts features from full cycling data was initially proposed by Greenbank et al. [13], and a histogram-based 3-feature set was selected with an optimal tradeoff between model accuracy and complexity. This Oxford 3-feature set is selected as the second feature set in this work.

B. Train-test split

In our prior experience, the train-test split has been observed to have a significant impact on ML model performance. Therefore, in order to improve model generalization performance and ensure reliable model evaluation on the test set, the stratified random sampling method [17] is used to split the dataset, with 80% in a training set, and 20% in a test set. Specifically, retrieved battery cells are first classified into different groups based on a desired criterion for each scenario. Then, equal ratios of cells that belong to different groups are maintained in the training and test set at each split. Note that the specific criterion for classifying cells into different groups may differ from one use case to another (see Section III). Lastly, to reduce the randomness of experiments, the stratified random sampling is repeated 5 times, and the results of 5 train-test splits are averaged.

C. Model selection

1) Elastic net: The elastic net is a regularized method used in the fitting of linear or logistic regression models [16]. The elastic net linearly combines the L1 and L2 penalties of the lasso and the ridge methods. The combination will result in a sparse model where only a few coefficients are non-zero. The elastic net has been shown to perform well when there are high correlations between the features, as is often the case for capacity predictions.

2) Gaussian process regression: The Gaussian process regression (GPR) has been widely employed to address battery health prognostic problems due to its advantages of being nonparametric, probabilistic, and customizable [12] [18] [19]. For the input $\boldsymbol{x}$, a Gaussian process (GP) defines a probability distribution of a function $f(\boldsymbol{x})$. Then the property of $f(\boldsymbol{x})$ is specified by its mean function $m(\boldsymbol{x})$ and covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$. For a finite number of input samples from the training set, $\{x_i\}_{i=1}^{N_D}$, the GPR calculates the joint Gaussian probability distribution $p(f(\boldsymbol{x}_1), ..., f(\boldsymbol{x}_{N_D}))$, with mean function $\boldsymbol{m}$ and covariance function $\boldsymbol{K}$. For simplicity, the mean function is usually assumed to be zero [20]. The Matérn 5/2 covariance function is used as it has previously shown good performance in battery lifetime prediction [12].

3) Quantile regression forest: The quantile regression forest (QRF) provides conditional quantiles of an output variable given input features [21]. To construct the conditional distribution of the output variable, all observations of the output variable in every leaf of each tree are stored in the QRF [22]. The 95% prediction intervals are constructed from the predictive quantiles of the output variable, i.e., $\hat{I}(\boldsymbol{x}) = [\hat{Q}_{.025}(\boldsymbol{x}), \hat{Q}_{.975}(\boldsymbol{x})]$.

D. Model performance evaluation

1) Point prediction evaluation: To evaluate the quality of battery lifetime point predictions, three metrics are used in this work, namely, root-mean-square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). They are defined as

$$\mathrm{RMSE}(y_j, \hat{y}_j) = \sqrt{\frac{1}{N_T}\sum_{j=1}^{N_T}(y_j - \hat{y}_j)^2} \qquad (1)$$

$$\mathrm{MAPE}(y_j, \hat{y}_j) = \frac{1}{N_T}\sum_{j=1}^{N_T}\left|\frac{y_j - \hat{y}_j}{y_j}\right| \qquad (2)$$

TABLE I

USE CASES CLASSIFICATION FOR BATTERY LIFETIME PREDICTION

Scenario	Battery usage profile	Use case	Battery life stage
1	Identical cycling	High-performance electrode design [4]	Design
		Graphite electrode manufacturing process design [5]	Manufacturing
		Product quality control [6] [7]	Manufacturing
		Battery grading [8]	Manufacturing
2	Protocol cycling	Formation protocol optimization [9]	Manufacturing
		Fast-charging protocol optimization [10]	Usage
3	Dynamic cycling	Onboard battery health estimation and prediction [14]	Usage
		Second-life battery health estimation and prediction [15]	Repurposing

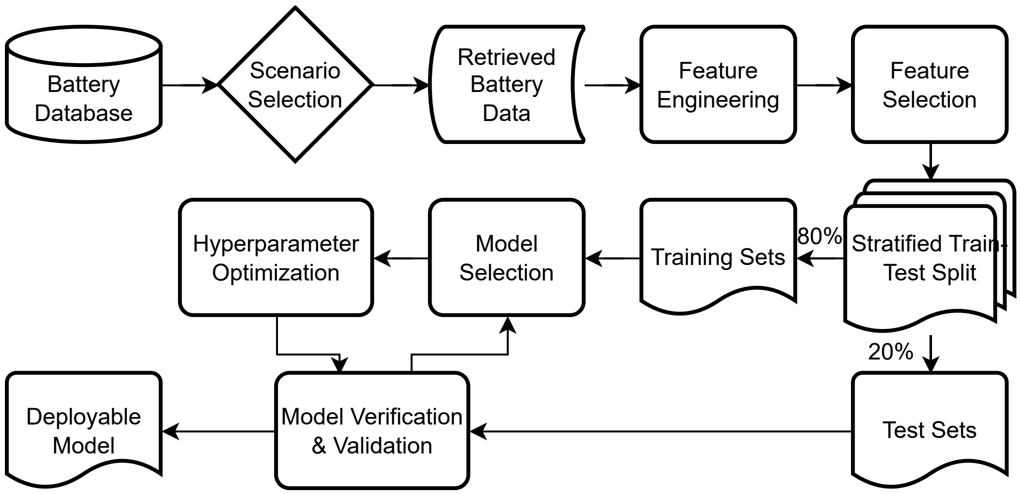

Fig. 1. The scenario-aware machine learning pipeline for battery lifetime prediction.

TABLE II

TWO FEATURE SETS BY TWO FEATURE ENGINEERING METHODS

Feature engineering method	Data type	Feature	Output variable
MIT 6-feature set [16]	Discharge data	Minimum, variance, skewness, and kurtosis of difference of the discharge voltage curve between cycle 100 and cycle 10 (i.e., $\Delta Q_{100-10}(V)$)	Cycle life
		Discharge capacity at cycle 2 (i.e., Q_2)	
		Difference between maximum discharge capacity within the first 100 cycles and discharge capacity at cycle 2 (i.e., $\Delta Q_{\mathrm{Max}-2}$)	
Oxford 3-feature set [13]	Full cycling data	Time spent between voltages corresponding to 1st and 33rd percentiles over every 12 hours (i.e., V_{12})	Capacity change over every 12 hours (i.e., ΔQ)
		Time spent between voltages corresponding to 33rd and 67th percentiles over every 12 hours (i.e., V_{23})	
		The calendar time (i.e., t)	

$$R^2(y_j, \hat{y}_j) = 1 - \frac{\sum_{j=1}^{N_T}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{N_T}(y_j - \bar{y})^2} \tag{3}$$

where N_T denotes the number of samples to be evaluated in the test set, and y_j and $\hat{y}_j$ denote the observed cycle life and the predicted cycle life of cell j, respectively. The average cycle life for in total N_T samples in the test set is calculated as $\bar{y} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j$.

2) Range prediction evaluation: To evaluate the quality of battery lifetime range prediction, three metrics are used in this work, namely, prediction interval coverage probability (PICP), mean prediction interval width (MPIW), and averaged interval score (AIS). They are expressed by

$$\mathrm{PICP} = \frac{1}{N_T}\sum_{j=1}^{N_T} c_j \tag{4}$$

$$\mathrm{MPIW} = \frac{1}{N_T}\sum_{j=1}^{N_T} |u_j - l_j| \tag{5}$$

$$\begin{aligned}\mathrm{AIS} =& \frac{1}{N_T}\sum_{j=1}^{N_T}((u_j - l_j) + \frac{2}{0.05}(l_j - y_j)\mathbb{1}_{\{y_j < l_j\}} \\ &+ \frac{2}{0.05}(y_j - u_j)\mathbb{1}_{\{y_j > u_j\}})\end{aligned} \tag{6}$$

where c_j is a binary variable. Specifically, if the observed battery lifetime of cell j in the test set is within the range constructed by the lower bound l_j and the upper bound u_j, then $c_j = 1$; otherwise, $c_j = 0$. $\mathbb{1}_{\{\cdot\}}$ is an indicator function and equal to 1 if the specified condition is satisfied; otherwise equal to 0.

III. Experiments and Results

A. Battery dataset

To demonstrate the effectiveness of our proposed scenario-aware machine learning pipeline for battery lifetime prediction, we apply it to a battery dataset generated by Toyota Research Institute together with Stanford University and Massachusetts Institute of Technology [16]. In total, this dataset consists of 124 lithium iron ferrous phosphate (LFP)/graphite cylindrical cells manufactured by A124 Systems (model APR18650M1A, 1.1 Ah nominal capacity). The test purpose is to characterize the impact of different fast-charging protocols on battery health degradation and then identify the best high-cycle-life fast-charging protocol. All the cells were charged with either a one-step or a two-step charging protocols from 0% to 80% state-of-charge (SoC), and then charged with a 1C constant current–constant voltage (CC-CV) charging step from 80% to 100% SoC. Thereafter, all the cells were identically discharged with a 4C CC-CV discharging step to 0% SoC. All the tests were conducted in an environmental chamber at a constant temperature of 30°C. The cells were cycled until they reached the end of life (EoL) threshold (80% of initial nominal capacity in this dataset).

In terms of data availability, time-series cell voltage, current, and (surface) temperature in each cycle were continuously measured, while two battery health metrics, i.e., rated capacity (4C discharge, 30 °C) and internal resistance (±3.6 C pulse current, 30 or 33 ms pulse width, 80% SoC) were measured per cycle. To evaluate the effectiveness of two feature engineering methods (see Table II) on this dataset, features versus output variables at one stratified train-test split are illustrated in Fig. 2 and Fig. 3, respectively. It can be observed from both figures that there is a strong overlap between training and test data.

B. Battery lifetime prediction in scenario 1

In scenario 1, batteries are repeatedly cycled under identical or nearly identical conditions with the objective of quantifying their intrinsic cell-to-cell variation. In order to capture the cell-to-cell variations of a larger population, a sample size of 8-10 cells is recommended in Ref. [23]. In the chosen dataset, one charging protocol, i.e., ”5.3C(54%)-4C” has been repeated on 8 cells. Therefore, for this charging protocol, the battery lifetime is used as the criterion to first classify these 8 cells into long-lived (i.e., greater than or equal to observed median lifetime) cells and short-lived (i.e., less than observed median lifetime) cells. Then the battery data is split into a training set (6 cells) and a test set (2 cells). Moreover, equal ratios of long-lived cells and short-lived cells are kept in the training and test set at each split.

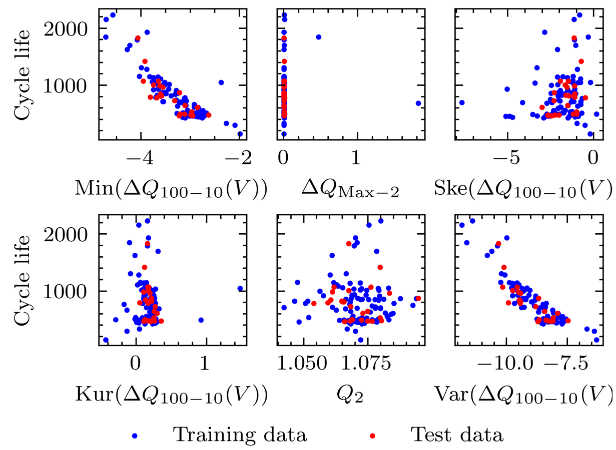

Fig. 2. Cycle life versus MIT 6 features at one stratified train-test split.

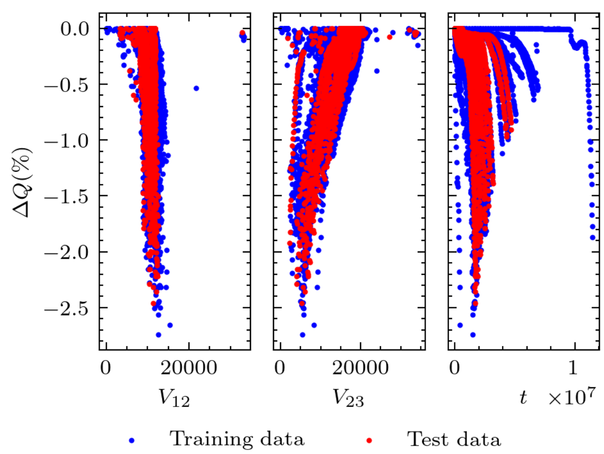

Fig. 3. Capacity change versus Oxford 3 features at one stratified train-test split.

The battery lifetime point and range prediction results for three ML models using two different feature sets in scenario 1 are summarized in Table III. In terms of battery lifetime point prediction quality, as measured by RMSE, MAPE, and R^2, it can be seen that the elastic net using MIT 6-feature set performs the best, while the GPR using Oxford 3-feature set performs the second best among all ML models using two feature sets. In terms of battery lifetime range prediction quality, as measured by PICP, MPIW, and AIS, it can be seen that the GPR performs better than QRF using the same MIT 6-feature set. Moreover, the GPR using Oxford 3-feature set is also capable of predicting battery capacity fade with uncertainty quantified by confidence intervals (see Fig. 4).

In scenario 1, the objective is to quantify intrinsic cell-to-cell variation, and then these information is used for refining battery design and manufacturing process, grading batteries before their usage, etc. In line of this, the experimental results in Table III suggest that the elastic net using MIT 6-feature would be the best choice if only battery lifetime spread is to be quantified, while the GPR using Oxford 3-feature set would be preferred if both battery lifetime spread and capacity spread are to be quantified.

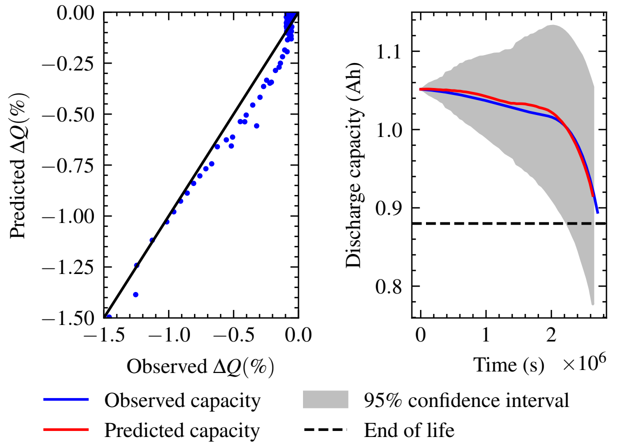

Fig. 4. Predicted versus observed ΔQ (left) and predicted capacity versus time (right) of a sample cell [b3c39] in the test set. Note that the GPR with Matérn 5/2 covariance function is selected for battery lifetime prediction here.

C. Battery lifetime prediction in scenario 2

In scenario 2, batteries are repeatedly cycled under intentionally varied conditions with the objective of characterizing the impact of different protocols (e.g., formation protocols or fast-charging protocols) on battery health degradation and lifetime. In the chosen dataset, there are 72 different charging protocols whose nominal charging time from 0% to 80% SoC ranges from 9 to 13.3 min. Therefore, the nominal charging time from 0% to 80% SoC is used as the criterion to first classify cells into fast-charged (less than 10.5 min) cells, medium-charged (between 10.5 and 11.7 min) cells, and slow-charged (greater than 11.7 min) cells. Then the battery data is split with 80% in a training set (99 battery cells) and 20% in a test set (25 battery cells). Moreover, equal ratios of fast-charged cells, medium-charged cells, and slow-charged cells are kept in the training and test set at each split.

The battery lifetime point and range prediction results for three ML models using two different feature sets in scenario 2 are summarized in Table IV. In terms of battery lifetime point prediction quality, as measured by RMSE, MAPE, and R^2, it can be seen that the GPR using Oxford 3-feature set performs the best, while the elastic net using Oxford 3-feature set performs the second best among all ML models using two feature sets. In terms of battery lifetime range prediction quality, as measured by PICP, MPIW, and AIS, it can be seen that the QRF performs better than the GPR using the same MIT 6-feature set. Instead of proving battery lifetime range prediction, the GPR using Oxford 3-feature set predicts battery capacity fade trajectories with uncertainty quantified as confidence intervals (see Fig. 5).

In scenario 2, the objective is to characterize the impact of different protocols (e.g., formation protocols or fast-charging protocols) on battery health degradation and lifetime, and then this information is used for protocol optimization. Therefore, the experimental results in Table IV suggest that the QRF model using MIT 6-feature set, or the GPR using Oxford 3-feature set, would be preferred in scenario 2 as they both provide more information for decision-making under uncertainty via prediction intervals or confidence intervals than point prediction alone. For an example of how uncertainty information facilitates decision-making that reduces the occurrence of unacceptably short cycle life, we refer to Ref [24].

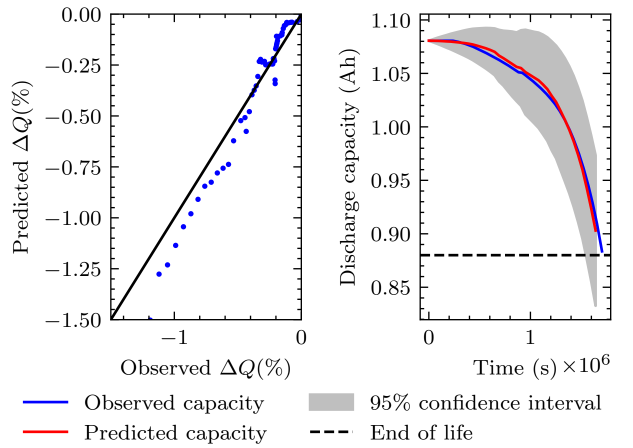

Fig. 5. Predicted versus observed ΔQ (left) and predicted capacity versus time (right) of a sample cell [b2c26] in the test set. Note that the GPR with Matérn 5/2 covariance function is selected for battery lifetime prediction here.

IV. Conclusion

Various machine learning (ML) models have been developed for battery lifetime prediction in different use cases at all stages of a battery's life. To enable the transferability of ML models across multiple use cases, a pipeline-based approach was proposed to automatically select the best feature engineering method for developing the best ML models in a chosen usage scenario. With the aid of our proposed pipeline, it has been demonstrated that the Gaussian process regression (GPR) using Oxford 3-feature set would be the overall best choice as it provided uncertainty information via confidence intervals in addition to predicted capacity fade trajectories with high accuracy, which was desired in both usage scenarios (i.e., identical cycling and protocol cycling). Moreover, this histogram-based feature engineering method was found to be able to generate input features (e.g., Oxford 3-feature set in this work) with predictive power generalized across two usage scenarios (i.e., identical cycling and protocol cycling).

These findings lead to our future research, i.e., to further validate the generalized predictive power of input features generated by this histogram-based feature engineering method in usage scenario 3 (i.e., dynamic cycling), a battery dataset with cells aged under real-driving scenarios is needed. To enable transferability of ML models for battery lifetime prediction across different scenarios, online fine-tuning strategies also need to be investigated.

TABLE III

BATTERY LIFETIME PREDICTION PERFORMANCE IN SCENARIO 1

Feature set +	**Point prediction evaluation**			**Range prediction evaluation**		
ML models	RMSE (cycles)	MAPE (%)	R^2	PICP (%)	MPIW (cycles)	AIS (cycles)
MIT 6-feature set + Elastic net	37	3.2	0.79	-	-	-
MIT 6-feature set + GPR	82	7.0	0.28	0.7	260	906
MIT 6-feature set + QRF	79	7.2	0.07	0.7	315	963
Oxford 3-feature set + Elastic net	333	21.5	0.01	-	-	-
Oxford 3-feature set + GPR	54	4.7	0.62	-	-	-

The charging protocol "5.3C(54%)-4C" has been repeated with 8 cells in the dataset. Hence, these 8 cells are retrieved from the battery database in this scenario.

TABLE IV

BATTERY LIFETIME PREDICTION PERFORMANCE IN SCENARIO 2

Feature set +	**Point prediction evaluation**			**Range prediction evaluation**		
ML models	RMSE (cycles)	MAPE (%)	R^2	PICP (%)	MPIW (cycles)	AIS (cycles)
MIT 6-feature set + Elastic net	190	20.7	0.72	-	-	-
MIT 6-feature set + GPR	165	17.2	0.78	95.2	607	865
MIT 6-feature set + QRF	151	12.2	0.83	93.6	451	635
Oxford 3-feature set + Elastic net	73	7.3	0.94	-	-	-
Oxford 3-feature set + GPR	41	4.9	0.99	-	-	-

In total, 124 cells are retrieved from the battery database in this scenario.

ACKNOWLEDGMENT

The authors would like to thank Volvo Group for funding this work.

REFERENCES

[1] L. Albertsen, J. L. Richter, P. Peck, C. Dalhammar, and A. Plepys, "Circular business models for electric vehicle lithium-ion batteries: An analysis of current practices of vehicle manufacturers and policies in the eu," *Resources, conservation and recycling*, vol. 172, p. 105658, 2021.

[2] R. Schmuch, R. Wagner, G. Hörpel, T. Placke, and M. Winter, "Performance and cost of materials for lithium-based rechargeable automotive batteries," *Nature energy*, vol. 3, no. 4, pp. 267–278, 2018.

[3] V. Sulzer, P. Mohtat, A. Aitio, S. Lee, Y. T. Yeh, F. Steinbacher, M. U. Khan, J. W. Lee, J. B. Siegel, A. G. Stefanopoulou *et al.*, "The challenge and opportunity of battery lifetime prediction from field data," *Joule*, vol. 5, no. 8, pp. 1934–1955, 2021.

[4] S. C. Kim, S. T. Oyakhire, C. Athanitis, J. Wang, Z. Zhang, W. Zhang, D. T. Boyle, M. S. Kim, Z. Yu, X. Gao *et al.*, "Data-driven electrolyte design for lithium metal anodes," *Proceedings of the National Academy of Sciences*, vol. 120, no. 10, p. e2214357120, 2023.

[5] S. X. Drakopoulos, A. Gholamipour-Shirazi, P. MacDonald, R. C. Parini, C. D. Reynolds, D. L. Burnett, B. Pye, K. B. O'Regan, G. Wang, T. M. Whitehead *et al.*, "Formulation and manufacturing optimization of lithium-ion graphite-based electrodes via machine learning," *Cell Reports Physical Science*, vol. 2, no. 12, 2021.

[6] T. Baumhöfer, M. Brühl, S. Rothgang, and D. U. Sauer, "Production caused variation in capacity aging trend and correlation to initial cell performance," *Journal of Power Sources*, vol. 247, pp. 332–338, 2014.

[7] A. Geslin, B. van Vlijmen, X. Cui, A. Bhargava, P. A. Asinger, R. D. Braatz, and W. C. Chueh, "Selecting the appropriate features in battery lifetime predictions," *Joule*, 2023.

[8] Y. Yuan, X. Kong, J. Hua, Y. Pan, Y. Sun, X. Han, H. Yang, Y. Li, X. Liu, X. Zhou *et al.*, "Fast grading method based on data driven capacity prediction for high-efficient lithium-ion battery manufacturing," *Journal of Energy Storage*, vol. 73, p. 109143, 2023.

[9] A. Weng, P. Mohtat, P. M. Attia, V. Sulzer, S. Lee, G. Less, and A. Stefanopoulou, "Predicting the impact of formation protocols on battery lifetime immediately after manufacturing," *Joule*, vol. 5, no. 11, pp. 2971–2992, 2021.

[10] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang *et al.*, "Closed-loop optimization of fast-charging protocols for batteries with machine learning," *Nature*, vol. 578, no. 7795, pp. 397–402, 2020.

[11] B. Jiang, W. E. Gent, F. Mohr, S. Das, M. D. Berliner, M. Forsuelo, H. Zhao, P. M. Attia, A. Grover, P. K. Herring *et al.*, "Bayesian learning for rapid prediction of lithium-ion battery-cycling protocols," *Joule*, vol. 5, no. 12, pp. 3187–3203, 2021.

[12] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Battery health prediction under generalized conditions using a gaussian process transition model," *Journal of Energy Storage*, vol. 23, pp. 320–328, 2019.

[13] S. Greenbank and D. Howey, "Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 2965–2973, 2021.

[14] Y. Zhang, T. Wik, J. Bergström, M. Pecht, and C. Zou, "A machine learning-based framework for online prediction of battery ageing trajectory and lifetime using histogram data," *Journal of Power Sources*, vol. 526, p. 231110, 2022.

[15] E. Braco, I. San Martin, P. Sanchis, A. Ursúa, and D.-I. Stroe, "Health indicator selection for state of health estimation of second-life lithium-ion batteries under extended ageing," *Journal of Energy Storage*, vol. 55, p. 105366, 2022.

[16] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.

[17] Z. Reitermanova *et al.*, "Data splitting," in *WDS*, vol. 10, 2010, pp. 31–36.

[18] K. Liu, X. Hu, Z. Wei, Y. Li, and Y. Jiang, "Modified gaussian process regression models for cyclic capacity prediction of lithium-ion batteries," *IEEE Transactions on Transportation Electrification*, vol. 5, no. 4, pp. 1225–1236, 2019.

[19] A. A. Chehade and A. A. Hussein, "A collaborative gaussian process regression model for transfer learning of capacity trends between li-ion battery cells," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9542–9552, 2020.

[20] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[21] R. Koenker, *Quantile Regression*. Cambridge University Press, 2005.

[22] N. Meinshausen and G. Ridgeway, "Quantile regression forests." *Journal of Machine Learning Research*, vol. 7, no. 6, 2006.

[23] C. Strange, M. Allerhand, P. Dechent, and G. Dos Reis, "Automatic method for the estimation of li-ion degradation test sample sizes required to understand cell-to-cell variability," *Energy and AI*, vol. 9, p. 100174, 2022.

[24] H. Zhang, Y. Su, F. Altaf, T. Wik, and S. Gros, "Interpretable battery cycle life range prediction using early cell degradation data," *IEEE Transactions on Transportation Electrification*, 2022.