Received: 22 May 2021 | Revised: 10 August 2021 | Accepted: 12 August 2021

DOI: 10.1002/er.7202

RESEARCH ARTICLE

INTERNATIONAL JOURNAL OF ENERGY RESEARCH WILEY

Python-based scikit-learn machine learning models for thermal and electrical performance prediction of high-capacity lithium-ion battery

Manh-Kien Tran[1] | **Satyam Panchal**[2] | **Vedang Chauhan**[3] | **Niku Brahmbhatt**[4] | **Anosh Mevawalla**[1] | **Roydon Fraser**[2] | **Michael Fowler**[1]

[1]Department of Chemical Engineering, University of Waterloo, Waterloo, Canada

[2]Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Canada

[3]Department of Mechanical Engineering, Western New England University, Springfield, Massachusetts, USA

[4]J G College of Computer Applications, ASIA Campus, Ahmedabad, India

Correspondence
Manh-Kien Tran, Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada.
Email: kmtran@uwaterloo.ca

Summary

With the increasing popularity of electric vehicles (EVs), the demands for rechargeable and high-performance batteries like lithium-ion (Li-ion) batteries have soared. Li-ion battery systems require the use of a battery management system (BMS) to perform safely and efficiently. Accurate and reliable battery modeling is important for the BMS to function properly. Currently, many BMS applications use the equivalent circuit model due to its simplicity. However, with the development of a cloud BMS, machine learning battery models can be utilized, which can potentially improve the accuracy and reliability of the BMS. This work investigates the performance of four different machine learning models used to predict the thermal (temperature) and electrical (voltage) behaviors of Li-ion battery cells. A prismatic Li-ion battery cell with a capacity of 25 Ah was cycled under a constant current profile at three different ambient temperatures, and the surface temperature and voltage of the battery were measured. The four machine learning regression models—linear regression, k-nearest neighbors, random forest, and decision tree—were developed using the scikit-learn library in Python and validated with experimental data. The results of their performance were reported and compared using the R^2 metric. The decision tree-based model, with an R^2 score of 0.99, was determined to be the best model in this case study.

KEYWORDS

battery modeling, lithium-ion battery, machine learning, multivariate multioutput regression, scikit-learn, thermal modeling

1 | INTRODUCTION

With the effects of climate change worsening, the push for electric vehicles (EVs), hybrid electric vehicles (HEVs), and plug-in hybrid electric vehicles (PHEVs) has drastically increased in recent years to reduce carbon dioxide emissions.[1] Due to the increasing popularity of EVs, the demand for rechargeable and high-performance

ACRONYMS: Ah, Ampere-hour; BMS, Battery management system; DT, Decision tree; DTR, Decision tree regressor; EV, Electric vehicle; ECM, Equivalent circuit model; HEV, Hybrid electric vehicle; KNN, k-Nearest neighbors; LR, Linear regressor; ML, Machine learning; MAE, Mean absolute error; NI, National instrument; PHEV, Plug-in hybrid electric vehicle; R^2, coefficient of determination; RFR, Random forest regressor; RVM, Relevance vector machine; SOC, State of charge; SOH, State of health; SVM, Support vector machine.

batteries has risen significantly. Lithium-ion (Li-ion) batteries have garnered the most interest out of all battery types.[2,3] This is because they have numerous advantages such as high energy density, high power density, long cycle life, low self-discharge rate, small size, lightweight, rapid charging capabilities, and wide temperature range.[4-7] The battery packs in EVs usually consist of hundreds to thousands of Li-ion cells. To extend the life of the battery, some precautions must be considered during the operation of the battery application since, for instance, exceeding voltage, current, capacity, or power limits may result in damages to the battery cells.[8-10] Sometimes, the possibility of a dangerous thermal runaway can also occur without appropriate actions being taken promptly, as there have been cases of EVs catching fire.[11-13] Therefore, Li-ion batteries must be electrically and thermally monitored and controlled to prevent these issues to ensure the performance of the vehicles and the safety of the users.[14,15] The battery management system (BMS), hence, is an important part of the battery system in EVs because it helps to control and monitor the safety and performance of the battery pack.[16,17]

The BMS acts as the brain of the battery system. Using sensors, it measures and records the voltage, current, and temperature of the battery in order to perform its functions.[18] The main functions of the BMS include battery monitoring, fault detection, cell balancing, determination of battery states such as state of health (SOH) and state of charge (SOC), energy management during operation, and thermal management.[19-21] For example, a BMS in an EV application can capture the current, voltage, power, and temperature of the battery to then make comparisons with predicted values from the battery model, allowing the BMS to detect any battery faults. Many of the BMS functions require an accurate and reliable battery model. Among various battery models, the equivalent circuit model (ECM) is the most used in real-time BMS applications due to its sufficient accuracy and low computational effort as well as its stability when used on major commercial Li-ion battery chemistries.[22,23] There are methods of estimating the ECM parameters online as the battery degrades. However, if measurement errors accumulate or BMS sensor faults occur, the ECM parameters often become more complicated and harder to estimate to predict the behavior of the battery system and help the BMS function accurately and reliably. With the rise of cloud computing and storage, as well as better BMS architecture, more complex yet more accurate battery algorithms than the ECM, including machine-learning-based battery algorithms, can be implemented in the BMS for EVs in the near future, as a complementary approach to assist with the existing battery control methods.

Machine learning (ML) is an approach within the field of artificial intelligence that teaches computers to learn and improve from data and process history using a set of algorithms so that reprogramming is not necessary.[24] In the field of energy storage, ML has emerged as a promising modeling approach recently. Feng et al.[25] used a simplified single particle model and a lumped thermal model to predict the battery temperature and terminal voltage. A neural network was incorporated to enhance the performance of the two models, formulating a complete electrochemical-thermal-neural-network (ETNN) model. Through validation experiments, the ETNN model was found to be able to estimate battery voltage and temperature accurately within an ambient temperature range of −10 to 40°C and a discharge rate of 10C. Li et al[26] developed a Li-ion battery model based on a deep learning algorithm and validated the model using data from electric buses. The result showed that the battery modeling method was able to simulate the battery characteristics accurately, and the mean absolute percentage error of the terminal voltage estimation is within 2.5%. Accurate SOC estimation plays a very important role in ensuring reliable functional operation of the BMS within a Li-ion battery system. Widodo et al[27] proposed a battery SOH estimation approach using the sample entropy (SampEn) feature of discharge voltage. SampEn can be used as a battery SOH indicator by utilizing algorithms to compute the predictability of a time series as well as to quantify the regularity of a data set. Two ML algorithms were developed, which were the support vector machine (SVM) and the relevance vector machine (RVM), where SampEn was the input and the SOH was the target vector to be estimated. It was found that both proposed algorithms performed well in terms of SOH prediction, with RVM being the superior algorithm. Hu et al[28] utilized the fuzzy C-means clustering technique on a Li-ion battery dataset of an EV drive cycle to determine the topology and antecedent parameters of the model. The results were then used to estimate the battery SOC. The proposed SOC estimation method was experimentally validated and found to be more accurate than other conventional fuzzy modeling approaches. Chemali et al.[29] proposed an SOC estimation method using deep neural networks, another type of ML algorithm. The experiments were conducted using various drive cycle profiles and ambient temperatures to obtain data to train the neural networks. In this approach, the battery voltage, current, and temperature were directly mapped to the SOC via the neural networks. This ML-based SOC estimator was then validated using different datasets, returning relatively low errors. The performance of various ML models has been studied and compared with the equivalent circuit and physics-based models.[30] The incorporation of domain knowledge provided the means of explainable "white box" predictions. ML models can result in fast and accurate battery state predictions in real time.

In this work, a comparison between four multivariate multioutput regression ML methods used to predict battery voltage and temperature is conducted using the scikit-learn library in Python. This study focuses on analyzing the feasibility of using ML in battery modeling as well as comparing the performance of different ML algorithms as battery models. The experimental data are recorded at different ambient temperatures along with other measurements. The ambient temperature, capacity, and current are used as inputs to the ML regressors to predict the voltage and temperature as the outputs. The experiments conducted on a 25-Ah GBS Li-ion battery cell performing under a different constant current discharge/charge rate of 1C at different ambient temperatures of 5°C, 22°C, and 35°C are presented. To the best of the authors' knowledge, no similar studies have been reported in the existing literature.

The structure of this article is as follows. In the first part of the research, an experimental study is presented including a detailed description of the experimental setup. Thereafter, the results obtained from the cell using four ML models are presented and discussed in the main section of this article. Potential applications and conclusions are then discussed in the last section.

2 | EXPERIMENTAL SETUP

The experiments involved a 25-Ah GBS lithium iron phosphate battery cell, used for data collection and model validation, with specifications shown in Table 1. The completed experimental setup for battery testing is shown in Figure 1. This setup consists of three components: a MACCOR battery tester, a CSZ MicroClimate thermal chamber, and a computer. The current collectors and voltage sensors from the MACCOR were connected to the cell in the thermal chamber. The computer has a software application that controls the cycling of the cell, where the current is assumed to be positive during the discharging step and negative during the charging step. A 1C-rate constant current profile was used for charging and discharging the battery. The voltage and current data were collected at a frequency of 1 Hz and stored in the computer. The thermal chamber was used to simulate different ambient temperatures for the battery testing environments. Three ambient temperatures were considered in this study, which were run in the order of 22°C, 5°C, and 35°C. The surface temperature of the battery was measured and stored in the computer using the thermocouples attached to the battery cell as shown in Figure 2.

TABLE 1 GBS 25-Ah lithium-ion battery cell specifications

Specifications	Value
Cathode material	$LiFeMnPO_4$
Anode material	Graphite
Electrolyte	Carbonate-based
Nominal capacity	25 Ah
Nominal voltage range	3.2 V
Dimensions	125 mm × 280 mm × 234 mm
Max charging current	3C
Max discharging current	3C (Continuous)/10C (pulsed)
Operating temperature	−20°C to 65°C
Weight	3.2 kg

3 | MACHINE LEARNING MODELS

Four ML models for battery behavior prediction and analysis were built using the scikit-learn library in Python.[31-33] The experimental data such as test time, capacity, current, voltage, energy, and temperature were recorded during the battery cycling at three different ambient temperatures. The temperatures were recorded from eight distinct points on the battery cell. A glimpse of sample data during the cycling at the ambient temperature of 22°C is shown in Figure 3. The goal of using ML is to develop a model that can predict the battery voltage and temperature based on the capacity, current, and surface temperature of the battery. The first step in the ML model development was to identify the type of ML

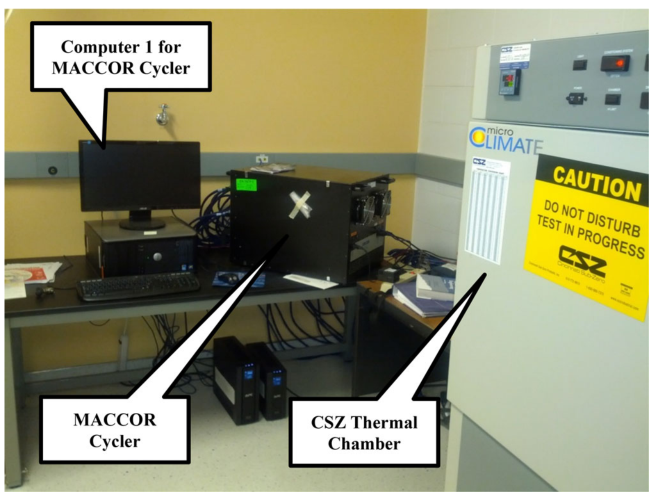

FIGURE 1 Experimental setup

domain. This battery model is the type of multivariate supervised regression learning using batch learning for training.[34] The steps for ML battery data modeling are shown in Figure 4.

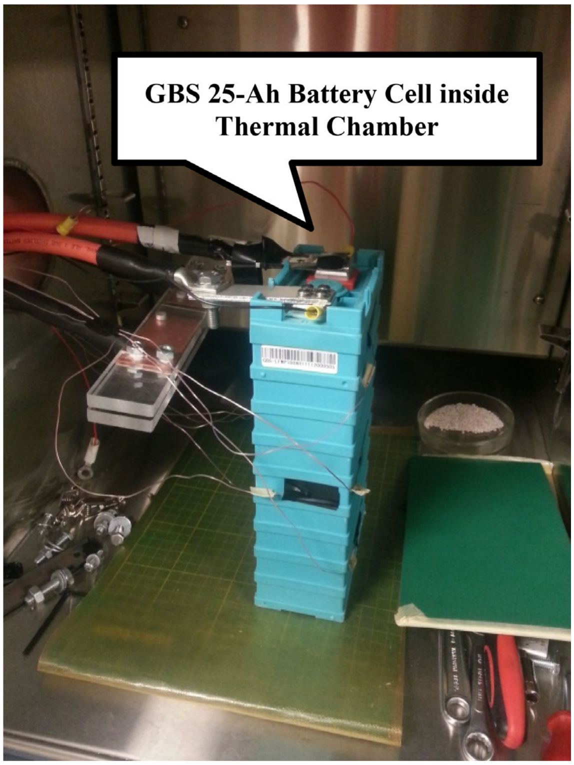

FIGURE 2 GBS 25-Ah battery cell and thermocouple locations

The first step for ML model development is to collect relevant experimental data. A total of 18 600 sample points were recorded at three operating temperatures. Jupyter Notebook was used for ML algorithm development. The data were imported into the notebook using the Pandas dataframe structure. A dataframe is a two-dimensional size-mutable data structure. The rows of the dataframe represent the samples, and the columns represent the data features. The number of features should be reduced to effectively develop the model and minimize data dimensionality. During the data preparation step, the average temperature was calculated from the eight individual temperatures. The ambient temperature, capacity, current, voltage, and average surface temperature of the battery were selected as the final features for the ML model development. Since the input features are not on the same scale, feature scaling was applied during ML model development. The final step of data preparation is to split the data into training, validation, and testing sets. Specifically, 70% of total data points were used for training and cross-validation, and the remaining 30% of the data were used for testing.

Developing ML algorithms involves a sequence of tasks including data preprocessing, features extraction, model selection, hyperparameter tuning, model fitting, and validation. A pipeline consists of a sequence of stages. There are two basic types of pipeline stages: transformers and estimators. Data visualization and preparation belong to the transformer pipeline, whereas model selection and hyperparameter tuning belong to the

test time	step tim	capacity	current	adj. current	voltage	energy	macstate	temp0	temp1	temp2	temp3	temp4	temp5	temp6	temp7	Avg
30	0.05	0	47.49	47.49	3.222	0	'D'	24.72374	25.65759	24.69261	24.66148	24.72374	25.7821	24.94164	24.66148	24.98054
30	0.21	0.004	100.00	100.00	3.115	0.012	'D'	24.72374	25.65759	24.69261	24.69261	24.72374	25.81323	24.94164	24.69261	24.99222
30	0.37	0.008	99.99	99.99	3.112	0.026	'D'	24.72374	25.65759	24.69261	24.69261	24.75486	25.84436	25.00389	24.72374	25.01167
31	0.6	0.015	100.00	100.00	3.109	0.046	'D'	24.72374	25.65759	24.72374	24.72374	24.75486	25.81323	25.00389	24.72374	25.01556
31	0.9	0.023	99.99	99.99	3.106	0.072	'D'	24.75486	25.65759	24.72374	24.72374	24.75486	25.84436	25.00389	24.72374	25.02335
31	1.3	0.034	100.00	100.00	3.104	0.106	'D'	24.75486	25.65759	24.75486	24.75486	24.75486	25.84436	25.03502	24.75486	25.03891
32	1.79	0.048	100.00	100.00	3.101	0.148	'D'	24.75486	25.65759	24.75486	24.75486	24.75486	25.84436	25.03502	24.75486	25.03891
32	2.31	0.062	99.99	99.99	3.099	0.193	'D'	24.75486	25.65759	24.75486	24.75486	24.81712	25.90661	25.03502	24.81712	25.06226
33	2.99	0.081	99.99	99.99	3.097	0.252	'D'	24.75486	25.65759	24.81712	24.75486	24.81712	25.90661	25.06615	24.81712	25.07393

FIGURE 3 Sample battery performance data at ambient temperature

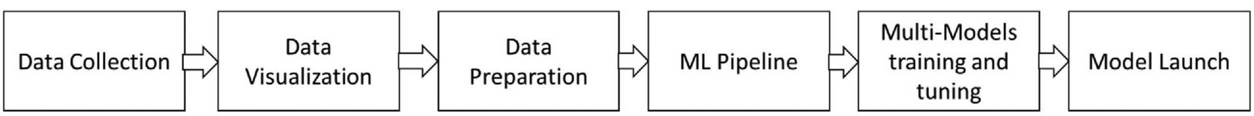

FIGURE 4 Development steps for machine learning battery modeling

estimator pipeline. Scikit-learn provides libraries to build both pipelines.

The most crucial step in data preparation is to select which models to try to fit the data. The battery problem is a multi-output regression-type ML problem. Multi-output regression refers to regression problems that involve predicting two or more numerical values given some input samples, for example, predicting the future battery voltage and temperature from inputs such as ambient temperature, capacity, and current, as well as current battery voltage and temperature. Many ML algorithms are designed for predicting a single numeric value. Some regression ML algorithms support multiple outputs directly. The following four ML algorithms have been implemented using scikit-learn:[35]

1. Linear regressor (LR)
2. k-Nearest neighbors regressor (KNNR)
3. Random forest regressor (RFR)
4. Decision tree regressor (DTR)

Linear regression is used to fit a model with coefficients by minimizing the residual sum of squares between the observed data and the predicted data. It is used to determine whether there is a linear relationship between a dependent variable and one or more independent variables. In this study, a linear regression algorithm was developed, and a model was fit using the training data. Once trained, the model was tested with both the test data and the full data. The performance of the model is discussed in the "Results and Discussion" section.

k-Nearest neighbors (KNN) is a simple algorithm that classifies new data based on a similarity measure (eg, distance functions) with the input data. KNN has been used in statistical estimation and pattern recognition. The idea behind nearest neighbor methods is to find a predetermined number of training objects that are the closest to the new point and use them to predict its label. The number of samples can be a user-defined constant (KNN learning). It can also vary depending on the density of points locally (radius-based neighbor learning). The distance can often be of any metrics, but the standard Euclidean distance is selected for the ML algorithm in this study. The KNNR is the modified implementation of KNN that calculates the average of the numerical target of the KNN. A KNNR algorithm was developed in this work, and a model was fit using the training data. Once trained, the model was tested with both the test data and the full data. The performance of the model is discussed in the "Results and Discussion" section.

Random forest or random decision forest is an ensemble learning method for classification, regression, and other tasks that functions by constructing a multitude of decision trees (DTs) at training time and outputting the class that is the mode of the classes (classification) or the mean prediction (regression) of the individual trees. The RFR scikit-learn estimator was developed, trained, and tested with the experimental battery datasets in this study.

DTs are versatile ML algorithms that can perform both classification and regression tasks and support multiple outputs. The structure of a DT model resembles an inverted tree where the nodes represent features or predictors, the links or branches between the nodes represent decisions, and the leaves represent responses or outputs. Regression DT can output continuous values that are appropriate for battery modeling. This method can perform data fitting, make predictions quickly, and require relatively low memory usage. DTR was developed using the scikit-learn library and used to train, validate, and test the experimental battery data.

These ML models were selected for this study due to their ease of implementation using the scikit-learn library in Python as well as their popularity in real-world data fitting problems. Since other physical models such as ECM are often unable to fit battery data well once the battery is deeply degraded, ML models can be more accurate by learning and adapting to data from new battery states. All four models have been trained separately with the same training data and tested on the same testing data. All models output two values for each set of input values. The performance of each algorithm and its prediction plots are discussed in the "Results and Discussion" section.

4 | RESULTS AND DISCUSSION

The data plots for the features used in the ML models are shown in Figure 5. It can be seen that the voltage plot is almost identical for each cycle at various operating temperatures, while the battery average temperature variation is significant due to the difference in the three ambient temperatures simulated during the experiment. The average temperature plot demonstrated that battery temperature nearly remained constant during its discharge cycle at an ambient temperature of 22°C. The battery temperature increased and then decreased to a lower value for the discharge cycle at an ambient temperature of 5°C, whereas at an ambient temperature of 35°C, it increased and remained high throughout the discharge cycle.

All four ML models were trained, validated, and tested with the battery dataset, and their performance is measured using R^2 (R-squared). R^2 is a statistical measure of how close the data are to the fitted regression line.

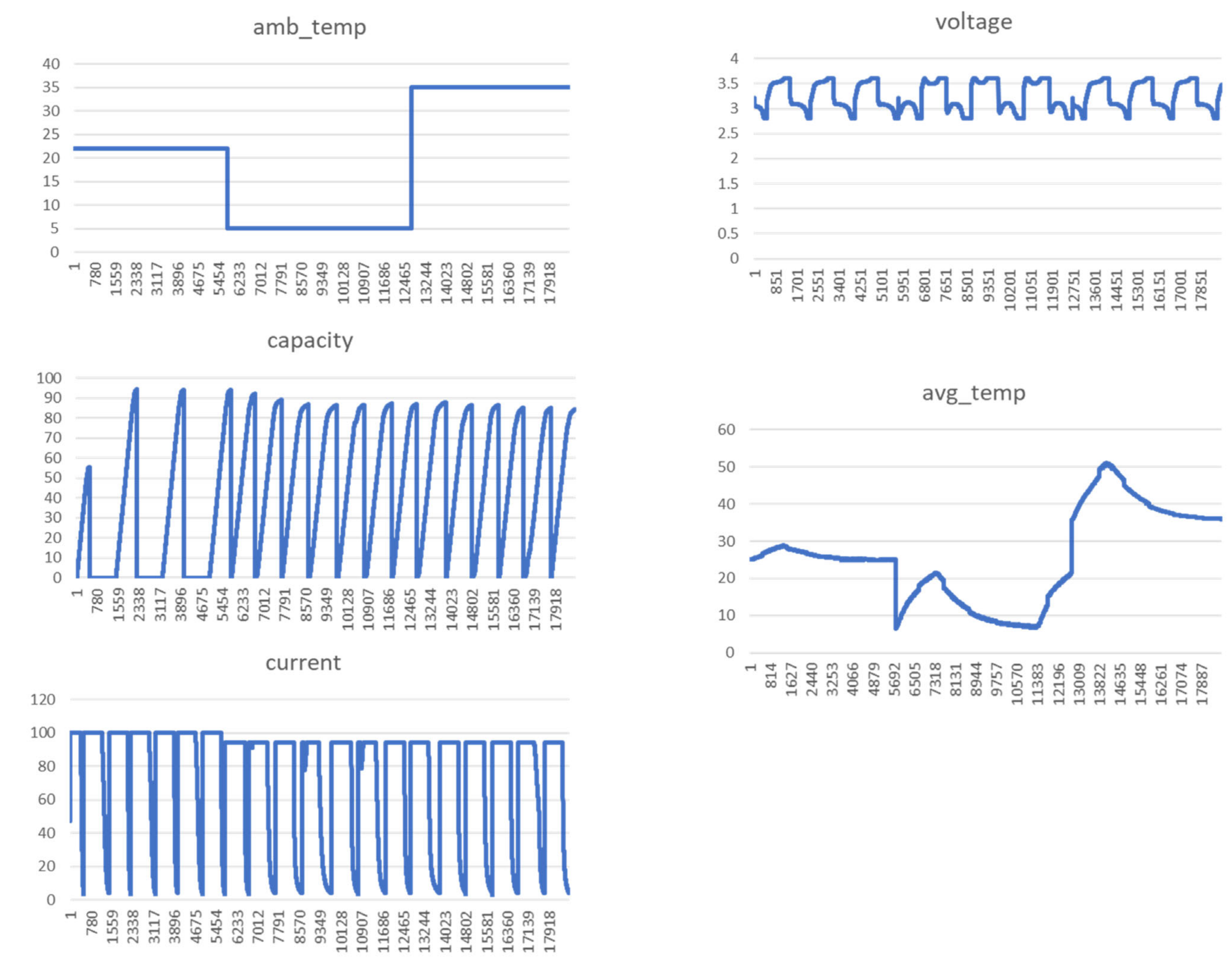

FIGURE 5 Data plots of battery machine learning model features

It is also known as the coefficient of determination. A higher coefficient is an indicator of better goodness of fit for the observations. The best possible score for R^2 is 1. A constant model that always predicts the expected value of y, disregarding the input features, would get an R^2 score of 0. The R^2 score can be negative because the model can be arbitrarily worse. Once the models were trained and cross-validated, they were also tested on the full set to obtain the plots for both voltage and temperature predictions.

The plots for all four methods are shown in Figure 6. Figure 6A,B shows voltage and temperature predictions using LR. The LR approach did not fit the data well, and it predicted a low value for both voltage and temperature. The R^2 score for LR was found to be 0.5. The plots shown in Figure 6C,D are voltage and temperature plots using the KNNR method. It performed better than linear regression. However, the results using KNNR did not fit well at high ambient temperature values. The R^2 score for KNNR was determined to be 0.75. RFR voltage and temperature prediction plots are shown in Figure 6E,F. The model fits well with the full range of data. However, it seems noisy for both voltage and temperature predictions. The R^2 score for RFR was calculated to be 0.94. The performance of the DTR approach is shown in Figure 6G,H. The DT model with a leaf size of 150 was found to have the best fit. It predicted both voltage and temperature for the entire range of data with good results. This method achieved the R^2 score of 0.99.

All four models were observed to perform better at the ambient temperature of 22°C compared to those at the ambient temperatures of 5°C or 35°C. Overall, the four ML models showed decent levels of accuracy as can be seen with the trends of the predicted data matching those of the experimental data. Even though the ML models did not perform as well as other models in existing literature, such as the ECM, due to the lack of battery data, the model errors would be improved in future studies with more data. ML battery models, once optimized, should have the ability to outperform other

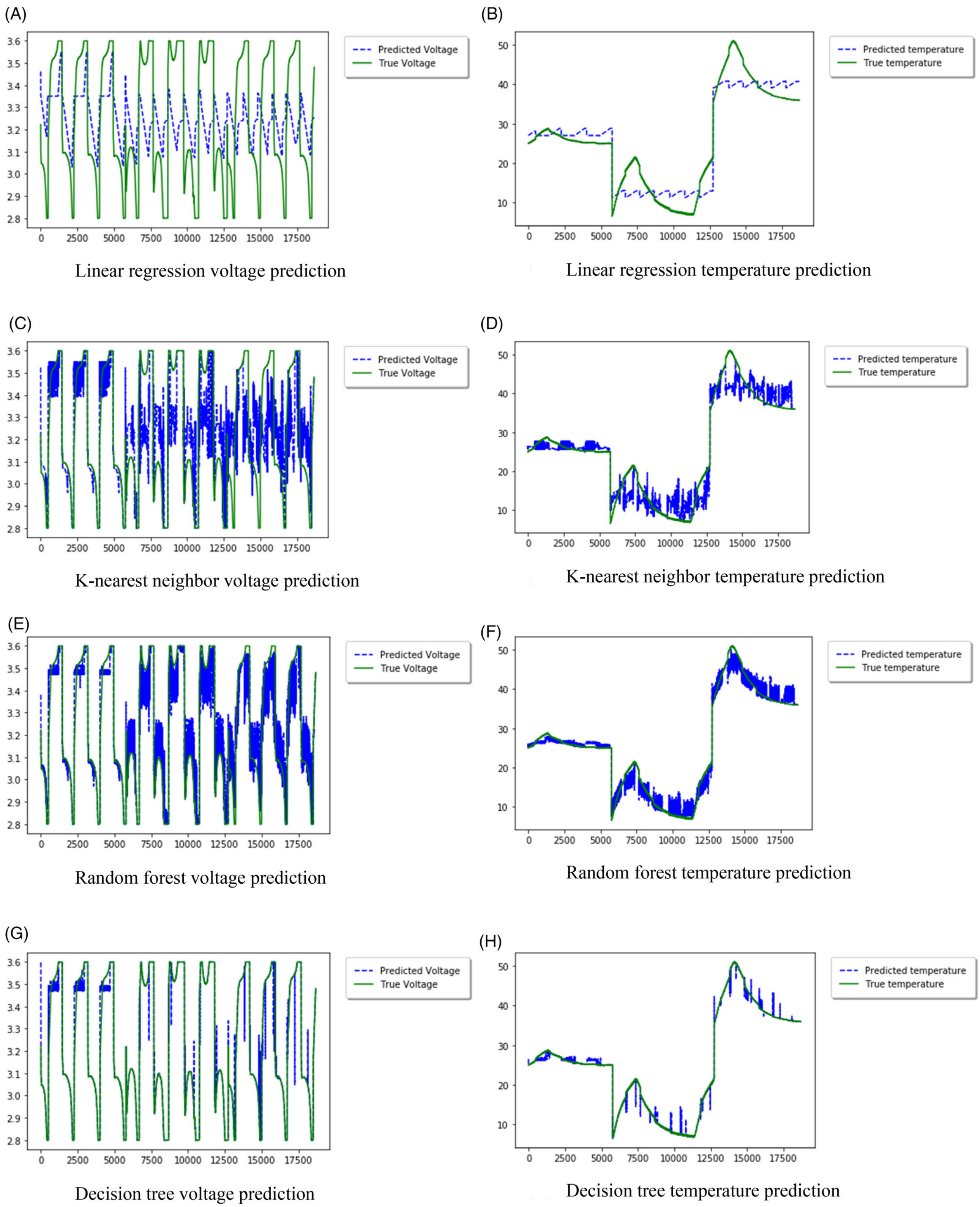

FIGURE 6 Machine learning models predictions for battery voltage and temperature

physical models especially when the battery is deeply degraded. More data from more battery cells tested can potentially improve the accuracy of the models, as ML models are often significantly dependent on data. With assistance from cloud computing and cloud storage in the future, more battery data can be obtained and analyzed

TABLE 2 Summary of model performance using R^2 score

Method	R^2 score
Linear regression	0.5
k-Nearest neighbors	0.75
Random forest	0.94
Decision tree	0.99

in real time using cloud BMS applications.[36] The increasing amount of data will allow ML models to be more accurate and reliable than the ECM, especially as the batteries start to degrade over time.[37]

Table 2 shows a summary of the performance of all ML models with their R^2 score. As can be seen, the DT-based regressor is the best for the multiple output prediction for the battery analysis problem, and it is recommended to be used for ML battery modeling in BMS applications when the BMS architecture allows for more complex algorithms. The use of ML battery models can potentially increase the accuracy and reliability of the BMS in EVs in the future to ultimately improve the safety and performance of EV applications.

5 | CONCLUSIONS

This research work has reported a comprehensive investigation and comparison of four ML models used to predict the electrical (voltage) and thermal (temperature) performance of a GBS 25-Ah Li-ion battery cell cycling at a current of 1C-rate and three different ambient temperatures. The ML models were created using the scikit-learn library in Python, and the simulated data were validated with the experimental data for battery voltage and temperature. The battery ML models were able to capture the battery's electrical and thermal behaviors over a wide range of ambient temperatures. It is observed that battery temperature nearly remained constant during its discharge cycle at the ambient temperature of 22°C, increased and decreased to a low value for discharge cycle at the ambient temperature of 5°C, and increased and remained high throughout the discharge cycle at the ambient temperature of 35°C. The ML regression models, including linear regression, k-nearest neighbors, random forest, and decision tree, were trained, validated, and tested with five features as inputs: ambient temperature, battery capacity, battery current, past battery voltage, and past battery temperature. The models predicted two outputs: future battery voltage and future battery temperature. Among the ML models, the DT-based model resulted in the best performance with an R^2 score of 0.99.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Manh-Kien Tran https://orcid.org/0000-0001-9937-1749
Satyam Panchal https://orcid.org/0000-0001-8152-0451

REFERENCES

1. Faria R, Moura P, Delgado J, de Almeida AT. A sustainability assessment of electric vehicles as a personal mobility system. *Energy Conver Manage*. 2012;61:19-30.
2. Ould Amrouche S, Rekioua D, Rekioua T, Bacha S. Overview of energy storage in renewable energy systems. *Int J Hydrogen Energy*. 2016;41(45):20914-20927.
3. Tran M-K, Sherman S, Samadani E, et al. Environmental and economic benefits of a battery electric vehicle powertrain with a zinc–air range extender in the transition to electric vehicles. *Vehicles*. 2020;2(3):398-412.
4. Ling Z, Wang F, Fang X, Gao X, Zhang Z. A hybrid thermal management system for lithium ion batteries combining phase change materials with forced-air cooling. *Appl Energy*. 2015;148:403-409.
5. Tran M-K, Akinsanya M, Panchal S, Fraser R, Fowler M. Design of a hybrid electric vehicle powertrain for performance optimization considering various powertrain components and configurations. *Vehicles*. 2021;3(1):20-32.
6. Ritchie A, Howard W. Recent developments and likely advances in lithium-ion batteries. *J Power Sources*. 2006;162:809-812.
7. Tran M-K, DaCosta A, Mevawalla A, Panchal S, Fowler M. Comparative study of equivalent circuit models performance in four common lithium-ion batteries: LFP, NMC, LMO, NCA. *Batteries*. 2021;7:(3):51.
8. Woody M, Arbabzadeh M, Lewis GM, Keoleian GA, Stefanopoulou A. Strategies to limit degradation and maximize Li-ion battery service lifetime - critical review and guidance for stakeholders. *J Energy Storage*. 2020;28:101231.
9. Panchal S, Gudlanarva K, Tran M-K, Fraser R, Fowler M. High Reynold's number turbulent model for micro-channel cold plate using reverse engineering approach for water-cooled battery in electric vehicles. *Energies*. 2020;13(7):1638.
10. Mevawalla A, Panchal S, Tran M-K, Fowler M, Fraser R. Mathematical heat transfer modeling and experimental validation of lithium-ion battery considering: tab and surface temperature, separator, electrolyte resistance, anode-cathode irreversible and reversible heat. *Batteries*. 2020;6(4):61.
11. Wang Q, Mao B, Stoliarov SI, Sun J. A review of lithium ion battery failure mechanisms and fire prevention strategies. *Prog Energy Combust Sci*. 2019;73:95-131.
12. Liu H, Wei Z, He W, Zhao J. Thermal issues about Li-ion batteries and recent progress in battery thermal management systems: a review. *Energy Conver Manage*. 2017;150:304-330.
13. Kong L, Li C, Jiang J, Pecht M. Li-ion battery fire hazards and safety strategies. *Energies*. 2018;11(9):2191.

14. Mevawalla A, Panchal S, Tran M-K, Fowler M, Fraser R. One dimensional fast computational partial differential model for heat transfer in lithium-ion batteries. *J Energy Storage*. 2021;37:102471.
15. Tran M-K, Bhatti A, Vrolyk R, et al. A review of range extenders in battery electric vehicles: current progress and future perspectives. *World Electr Veh J*. 2021;12(2):54.
16. Xing Y, Ma EWM, Tsui KL, Pecht M. Battery management systems in electric and hybrid vehicles. *Energies*. 2011;4(11):1840-1857.
17. Lu L, Han X, Li J, Hua J, Ouyang M. A review on the key issues for lithium-ion battery management in electric vehicles. *J Power Sources*. 2013;226:272-288.
18. Tran M-K, Fowler M. Sensor fault detection and isolation for degrading lithium-ion batteries in electric vehicles using parameter estimation with recursive least squares. *Batteries*. 2020;6(1):1-16.
19. Tran M-K, Fowler M. A review of lithium-ion battery fault diagnostic algorithms: current progress and future challenges. *Algorithms*. 2020;13(3):62.
20. Cheng KWE, Divakar BP, Wu H, Ding K, Ho F. Battery-management system (BMS) and SOC development for electrical vehicles. *IEEE Trans Veh Technol*. 2011;60(1):76-88.
21. Chang W-Y. The state of charge estimating methods for battery: a review. *ISRN Appl Math*. 2013;2013:1-7.
22. Gomez J, Nelson R, Kalu EE, Weatherspoon MH, Zheng JP. Equivalent circuit model parameters of a high-power Li-ion battery: thermal and state of charge effects. *J Power Sources*. 2011;196(10):4826-4831.
23. Tran M-K, Mevawala A, Panchal S, Raahemifar K, Fowler M, Fraser R. Effect of integrating the hysteresis component to the equivalent circuit model of Lithium-ion battery for dynamic and non-dynamic applications. *J Energy Storage*. 2020;32:101785.
24. Khang T, Vuong N, Tran M-K, Fowler M. Fuzzy C-means clustering algorithm with multiple fuzzification coefficients. *Algorithms*. 2020;13(7):158.
25. Feng F, Teng S, Liu K, et al. Co-estimation of lithium-ion battery state of charge and state of temperature based on a hybrid electrochemical-thermal-neural-network model. *J Power Sources*. 2020;455:227935.
26. Li S, Li J, He H, Wang H. Lithium-ion battery modeling based on big data. *Energy Procedia*. 2019;159:168-173.
27. Widodo A, Shim M-C, Caesarendra W, Yang B-S. Intelligent prognostics for battery health monitoring based on sample entropy. *Expert Syst Appl*. 2011;38:11763-11769.
28. Hu X, Li SE, Yang Y. Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles. *IEEE Trans Transp Elect*. 2016;2(2):140-149.
29. Chemali E, Kollmeyer P, Preindl M, Emadi A. State-of-charge estimation of Li-ion batteries using deep neural networks: a machine learning approach. *J Power Sources*. 2018;400:242-255.
30. Ng M-F, Zhao J, Yan Q, Conduit GJ, Seh ZW. Predicting the state of charge and health of batteries using data-driven machine learning. *Nat Mach Intell*. 2020;2:161-170.
31. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2011;12:2825-2830.
32. Python Software Foundation, Python Language Reference, Available at http://www.python.org.
33. Reback J, McKinney W, Bossche JVD, et al. Pandas-dev/pandas: Pandas 1.0.5. *Zenobo*. 2020. https://doi.org/10.5281/zenodo.3898987
34. A. J. Izenman, "*Modern Multivariate Statistical Techniques*," Springer Texts in Statistics. New York, NY: Springer; 2008.
35. Géron A. *Hands-on Machine Learning With Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, CA: O'Reilly Media, Inc.; 2017.
36. Li W, Rentemeister M, Badeda J, Jöst D, Schulte D, Sauer DU. Digital twin for battery systems: cloud battery management system with online state-of-charge and state-of-health estimation. *J Energy Storage*. 2020;30:101557.
37. Aykol M, Herring P, Anapolsky A. Machine learning for continuous innovation in battery technologies. *Nat Rev Mater*. 2020;5:725-727.

How to cite this article: Tran M-K, Panchal S, Chauhan V, et al. Python-based scikit-learn machine learning models for thermal and electrical performance prediction of high-capacity lithium-ion battery. *Int J Energy Res*. 2022;46(2):786-794. doi:10.1002/er.7202