

Time and Frequency Domain Health Indicators for Capacity Prediction of Lithium-ion Battery

Ma'd El-Dalakhmeh

School of Computing, Engineering and
Digital Technologies
Teesside University
Middlesbrough, UK
Ma'd.El-Dalakhmeh@tees.ac.uk

Prudhiv Thummarapally

School of Computing, Engineering and
Digital Technologies
Teesside University
Middlesbrough, UK
W9210355@live.tees.ac.uk

Maher Al-Greer

School of Computing, Engineering and
Digital Technologies
Teesside University
Middlesbrough, UK
M.Al-Greer@tees.ac.uk

Mo'ath El-Dalakhmeh

School of Computing, Engineering and
Digital Technologies
Teesside University
Middlesbrough, UK
Mo'ath.El-Dalakhmeh@tees.ac.uk

Abstract— Predict the capacity of lithium-ion batteries with high accuracy is crucial to the reliability and safety of the system. Due to the complex nature and the nonlinear degradation phenomena of the lithium-ion battery, monitoring the battery's capacity is a challenging task. This paper proposes a machine learning model based on time and frequency domain health indicators to predict the capacity of lithium-ion battery cycled under different operational conditions. The time and frequency domain health indicators have been extracted from the measured voltage. The extracted features have been fed into extreme learning machine model to predict the capacity. This approach has been tested on 16 lithium-ion batteries cycled at many operational conditions from NASA. The results show that the proposed method can track the degradation from the extracted health indicators from both domains (time and frequency). The extreme learning model can effectively predict the capacity with a root mean square error of 1.3%.

Keywords—Lithium-ion battery, time-domain health indicator, frequency-domain health indicator, capacity prediction, extreme learning machine

I. INTRODUCTION

Over the last 20 years, lithium-ion batteries (LIBs) have been considered a potentially promising technology in the electrification of transportation and grid network systems due to their superior energy density, long cycle lifetime, and low discharge rate [1]. However, over time the performance of LIBs is gradually decreased. This reduction in battery performance due to numerous irreversible degradation mechanisms such as formulation of solid electrolyte interphase (SEI) layer, loss of active materials (LAM), and lithium plating [2]. Therefore, one of the main functions of the battery management system (BMS) is to accurately monitor the State of Health (SoH) of LIB over its lifetime to ensure stability, durability and avoid any failures of the LIB during its operation [3,4]. The SoH is defined as the capability of the LIB to store energy corresponding to its initial value [5]. The LIB capacity and internal ohmic resistance are commonly utilised as health indicators of the LIBs. When the chosen health indicator crosses a predefined specific point, which is 200% and 80% of the initial values in internal ohmic resistance and LIB capacity, respectively, the LIB reaches its End of Life (EoL) should be replaced [6]. In this paper, the capacity of the LIB is selected as a SoH indicator to be

predicted. Nevertheless, the complex and nonlinear nature of the degradation mechanisms of LIB makes the prediction of the capacity of LIB a challenging task.

Recently, several studies have reported numerous approaches for online capacity estimation of LIBs. The model-based and data-driven techniques are the general approaches applied for the capacity estimation of the LIBs [7].

In the case of model-based approaches, electrochemical models (EMs) or equivalent circuits models (ECMs) are developed to explain the dynamic behaviour of the LIB. The capacity of the LIB is assumed as one of the model parameters and is estimated by identifying the model parameters. For example, the authors in [8] proposed the ECM combined with two extended Kalman filters (EKF) to estimate the State of Charge (SoC) and capacity of the LIB over its lifespan. Moreover, the EM along with EKF is also utilised for LIB capacity estimation [9]. However, the major drawback of these models is that their performance highly relies on selecting the model parameters [10]. Moreover, ECMs cannot describe the underlying ageing phenomena of the LIB [11]. In contrast, the number of partial differential equations used to explain the LIB ageing mechanism is large and present a vast computational complexity causing it hard to be applied for online capacity estimation of LIB.

On the other hand, data-driven approaches have earned more reputation recently because of their ability to learn the correlation between the input measured data and target output without any required physical understanding of the principle operation of the LIB [12]. Various data-driven algorithms have been applied for capacity estimation, such as Support Vector Machine (SVM), Neural Network Models (NN), and Gaussian Process Regression (GPR). However, the accuracy of the mentioned approaches is highly dependent on the selected features or health indicators that define a clear pattern as the LIB ages. For example, the authors in [13] chose five health indicators from the charge curves, which are voltage, current, and temperature that are indicative of the capacity of the LIB and then are fed into the Relevance Vector Machine (RVM) algorithm to learn the correlation between the selected health indicators and the capacity of the LIB. Likewise, the authors in [14] obtained 14 health indicators from the charge

curves and used the extracted features to train a RVM algorithm to estimate the battery's capacity.

Moreover, the charge curves are also proposed for online capacity estimation combined with Deep Neural Network (DNN) [15]. However, the previous works extracted the health indicators and estimated the capacity of the LIB at strictly predefined operation conditions such as constant current and constant voltage. They used a dataset of LIB cycled at a constant current protocol that failed to consider the real-world operation conditions of LIBs.

To this end, a few studies have attempted to predict the battery's capacity cycled under dynamic real-world operating conditions [16]. For instance, 18 health indicators extracted from the measured voltage, current, and temperature values of LIB cycled under randomised current profile; these features are then used as input data for DNN [17]. However, the above study selected only time-domain health indicators for capacity prediction of LIB. The authors in [18] proposed a time and frequency domain diagnostic features approach for online battery SoH estimation to overcome this limitation. They extracted the time, and frequency domain features from the measured voltage of LIB cycled by using Worldwide Harmonized Light Vehicles (WLTC) test. After that, the extracted features such as mean, covariance, kurtosis, and Total Harmonic Distortion (THD) ranked based on three prognostic metrics named monotonicity, prognosability, and trendability to select the most health indicators that related to the degradation process inside the LIB.

Moreover, the most critical health indicators are then fed into ensemble gradient boost tree data-driven algorithm to estimate the SoH of the LIB. However, the previous study has not discussed the quality of the selected health indicators if the LIB cycled at an entirely different current profile. In other words, are the selected features efficient enough to estimate the capacity of the LIB if the LIB cycled with a different current profile.

To answer this question, this paper investigates the quality of the extracted health indicators from time, and frequency domains for LIB cycled at a randomised current profile provided by the NASA centre. The extracted health indicators are utilised as input for extreme learning machine algorithms to predict the LIB's capacity. The remainder of this paper is organised as follow: Section II presents the methodology for capacity estimation of LIB. Section III presents the results and discussion. Section IV offers the conclusion of this study.

II. METHODOLOGY

The framework for capacity estimation contains four main steps, as shown in Fig. 1. The first step is the experimental data acquisition. The second step is data pre-processing and health indicators extraction from time and frequency domains. The extracted features are then ranked based on two prognostic metrics named monotonicity and trendability in the third step. Lastly, Extreme Learning Machine (ELM) algorithm is applied to learn the correlation between the extracted health indicators and the capacity of the LIB. The following subsections discuss these steps in detail.

A. Experiment Dataset Analysis

In this study, a public dataset provided from NASA Ames Prognostics Centre of Excellence (PCoE) [19] was applied for capacity estimation of LIB. This dataset called Randomised Battery Usage dataset which using a random current profile in

cycling the LIB to reflect the real-world operation conditions of LIB. This dataset comprises the cycling results of 28 LIBs of type 18650 Li-cobalt, and their maximum capacity is 2 Ah. Each 4 LIBs of this dataset were cycled under the same operating conditions, so the dataset is divided into seven groups, and each group cycled under various conditions.

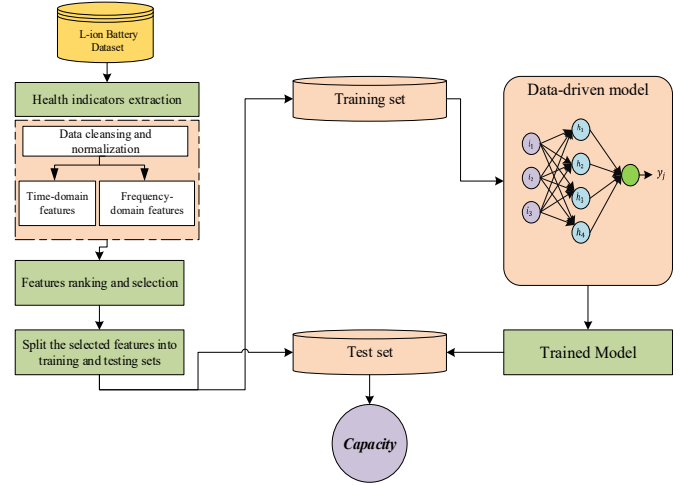


Fig. 1. The framework for the proposed capacity estimation.

Moreover, the LIBs in each group cycled under two current profiles, namely, random walk (RW) profile and reference charge/discharge profile. After every five days or 50 cycles, the capacity of LIB is measured as an indicator for the SoH of the LIB by integrating the 2 A constant current between the maximum and minimum voltage limits of the battery as express in (1). A more detailed description of this dataset can be found in [19].

$$C_{present} = \int_0^t I(t) dt \quad (1)$$

Where $C_{present}$ represent the present capacity of the LIB, and I is the discharge constant current of the LIB. Fig.2 and Fig.3 show the measured terminal voltage of the first and last ten cycles of LIB cycled at RW current profile, respectively.

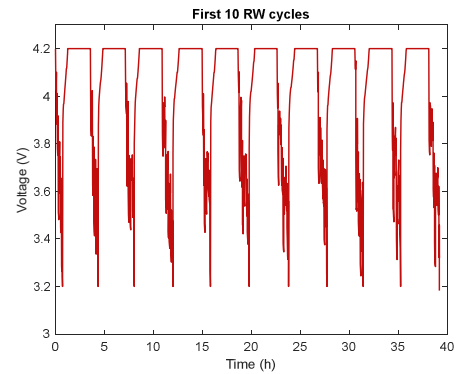


Fig. 2. Measured voltage for the first ten cycles.

As shown in Fig.2 and Fig.3, the behaviour of the terminal voltage is changed as the LIB degrades; thus, there is a clear relationship between the degradation and the measured voltage. In this paper, the measured voltage is utilised to extract the health indicators and estimate the capacity of the LIB. Moreover, Fig.4 illustrate an example of the capacity

degradation for four LIBs. As shown in Fig.4, each LIB degraded differently even if they cycled under the exact.

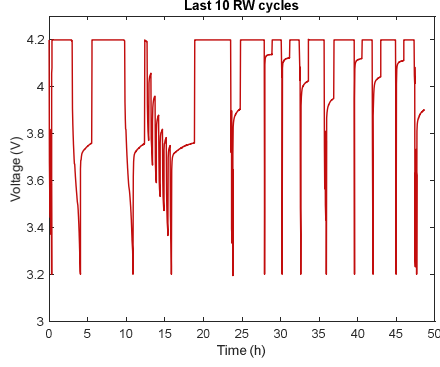


Fig. 3. Measured voltage for the last ten cycles.

Operating conditions this is due to the nonlinear behaviour of the LIB. Therefore, this study investigates the capability of time and frequency domain health indicators to deliver a robust feature during the degradation of the LIB. The results of this investigation are presented in Section III.

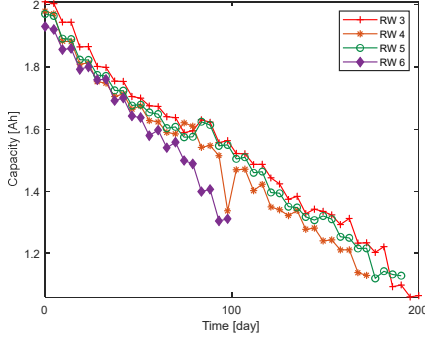


Fig. 4. The measured capacity of four LIBs.

B. Data Preprocessing and Health Indicators Extraction

In this step, before extract, the health indicators from the measured voltage, a pre-processing step for the data is required. Therefore, the measured voltage is pre-processed by removing the outliers, resampled, and normalised. After completed the data preparation, the health indicators can be extracted from both time and frequency domains. Fig.5 shows an example of the types of health indicators that can be extracted from the measured voltage in time and frequency domains [18]. More health indicators will be extracted, such as skewness and standard deviation from the time-domain and spectrum peak frequency from the frequency domain.

1) Time-Domain Health Indicators

To determine which are the most time-domain health indicators that their behaviour varies predictably during the LIB degradation, nine health indicators are selected to study their behaviour as the LIB ages. Table I illustrate these nine time-domain features corresponding with the mathematical expression of each feature. These features are selected based on recently published works related to health indicators analysis and selection in [20].

2) Frequency-Domain Health Indicators

After extracting the health indicators of the measured voltage in the time-domain, the measured voltage signal now should transform into frequency-domain to extract the health indicators in frequency-domain using power spectrum.

Therefore, In this study, Welch's technique is applied to convert the measured voltage from the time-domain to frequency-domain.

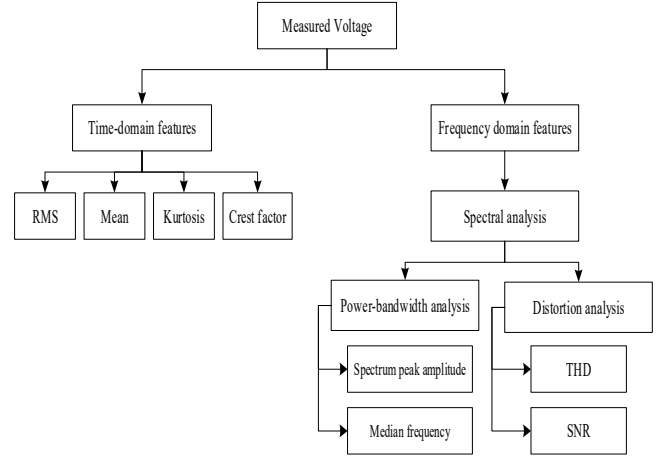


Fig. 5. The health indicators from time and frequency domains.

Table I. The mathematical expression of the time-domain health indicators. K is the total data points in the voltage, x_n is the measured voltage.

Health indicators	Mathematical expression
Mean	$\frac{1}{K} \sum_{n=1}^K x_n$
RMS	$\sqrt{\frac{1}{K} \sum_{n=1}^K x_n ^2}$
Standard Deviation (SD)	$\sqrt{\frac{\sum_{n=1}^K (x_n - \text{mean}(x))^2}{K - 1}}$
Shape Factor (SF)	$\frac{RMS}{Mean}$
Peak Value	The maximum absolute value of the measured signal
Impulse Indicator	$\frac{Peak\ value}{Mean}$
Crest Factor (CF)	$\frac{ Peak }{RMS}$
Kurtosis	$\frac{1}{K - 1} \left(\frac{\sum_{n=1}^K (x_n - \text{mean}(x))^4}{\text{Standard Deviation}^4} \right)$
Skewness	$\frac{1}{K - 1} \left(\frac{\sum_{n=1}^K (x_n - \text{mean}(x))^3}{\text{Standard Deviation}^3} \right)$

Welch's method applies the three following steps to obtain the power spectrum of any signal. Firstly, the time-domain signal is divided into N overlapping parts and each part of length L . Secondly, a window is applying to each N part and computing the periodogram for each windowed part. There are various types of window functions such as rectangular window, Hamming window, and Hann window. In this work, a rectangular window is applied [21]. Finally, the power spectrum is calculated by applying the average value of all parts as express in (2) [21]:

$$PS = \frac{1}{N} \sum_{i=1}^{N-1} \frac{1}{L} \left| \sum_{m=1}^{L-1} x_i(m) w(m) e^{-j2\pi f m} \right|^2 \quad (2)$$

Where $w(m)$ is the i^{th} windowed part, a more detailed explanation of Welch's method can be found in [21]. Once the power spectrum is obtained, the health indicators can be extracted. From Fig.5, the power spectrum amplitude, median frequency, and spectrum peak frequency health indicators are utilised from the power-bandwidth analysis. THD and signal-to-noise ratio (SNR) are also used as health indicators from the distortion analysis.

C. Prognostic Health Indicators Ranking

The selection process of the highly correlated health indicators to the capacity degradation of LIBs is crucially important. Therefore, health indicators ranking is applied to all the extracted time and frequency domain features to identify the best health indicators that can precisely track battery ageing. In this paper, two prognostic metrics are used for health indicators ranking known as monotonicity and trendability [22]. Monotonicity ranges between 0 and 1, where 0 indicates there is no monotonic between the extracted feature and the target output (capacity), and 1 means there is highly monotonic between the extracted feature and the target output. Monotonicity is expressed in (3) [22]:

$$M = \frac{u_i^+ - u_i^-}{u - 1} \quad (3)$$

Where u represent the number of observations from the beginning of life to the end-of-life trajectory of the LIB, and u_i^+ , u_i^- are the positive and negative first derivatives.

The trendability measure the similarity between trajectories of health indicators from the beginning of life to the end of life of the LIB. Trendability ranges between 0 and 1, and it is expressed in (4) [22]:

$$T = \min(|\text{corrcoef}_{ij}|), j = 1, 2, \dots, N \quad (4)$$

The essential health indicator is represented by adding M and T . In this paper, the extracted health indicator is a promising feature if $M + T$ is equal to or greater than 0.75.

D. Extreme Learning Machine (ELM)

A machine learning model called ELM is applied to learn the correlation between the extracted health indicators and capacity degradation of LIB. The ELM model was initially developed by [23]. The advantages of the ELM algorithm over other machine learning algorithms are learning very quickly and the random selection of the input weight and biases [23]. The structure of the ELM model includes three layers: input, hidden, and output layer, as shown in Fig.6. The basic operation of the ELM algorithm is described as (5) [23]:

$$\sum_{i=1}^L \beta_i \cdot h(w_i \cdot x_j + b_i) = y_j, j = 1, 2, \dots, N \quad (5)$$

Where h represent the activation function, w_i is the weight vector between the input and the hidden layers, and b_i indicates the weight vector between the hidden and the output layers.

E. Evaluation Criteria

To evaluate the accuracy performance of the proposed ELM algorithm for capacity prediction of LIB, two evaluation criteria are selected in this work which is the root mean square error (RMSE) and mean,

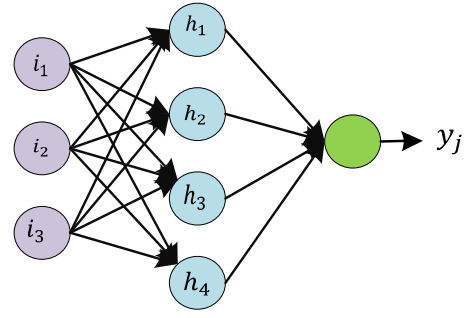


Fig. 6. The structure of the ELM model.

absolute error (MAE). These evaluation criteria are expressed as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (6)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (7)$$

Where \hat{y} and y_i represent the predicted and true capacity of LIBs, respectively, N is the number of data points.

III. RESULTS AND DISCUSSION

This section will discuss the extracted health indicators according to the monotonicity and trendability metrics to determine the best health indicators related to the LIB capacity degradation. The accuracy of the proposed ELM model is evaluated based on the selected evaluation criteria.

In this study, 16 LIBs from 4 groups cycled under different operating conditions are used for capacity estimation. One LIB from each group is kept aside for the test, and the rest 12 LIBs are utilised for model training.

A. Health Indicators Results

As mentioned earlier, the measured voltage is used to extract the health indicators from time and frequency domains. Generally, 14 health indicators are extracted from both domains. The extracted health indicators are then ranked based on the monotonicity, and trendability metrics discussed in subsection C. Fig.7 and Fig.8 show the extracted health indicators ranked based on monotonicity and trendability metrics. Fig.9 shows the most critical health indicators by adding $M + T$. As shown in Fig.9, any health indicator equal to or greater than 0.8 is considered a promising feature to track the degradation inside the LIB. From Fig.9, the "crest factor" value is the most promising health indicator while the health indicators with low importance value have not been considered sufficient health indicators to track the LIB trajectory degradation. On the other hand, the authors in [18] found that the "mean" value is the most promising health indicator to provide a predictable pattern reflect the LIB degradation during its lifetime. This difference between our findings and theirs indicates that a larger dataset with various operational conditions is crucially required to create a generalised machine learning model that can estimate the LIB's capacity at any working conditions.

B. ELM Model Results

As mentioned earlier, 16 LIBs from four different groups cycled at different current profiles are selected to train the ELM model, and 4 LIBs kept away to serve as a test set for the trained model. Each LIBs of the test set were cycled at different operating conditions, as illustrated in Table II. Fig.10 shows the capacity estimation results of LIBs RW4, RW7, RW26, and RW23, respectively. The ELM model estimated the capacity of LIBs cycled at different conditions with high accurate prediction results. Hence, that indicates that the selected health indicators from both domains can effectively track the degradation of the LIBs even if they were cycled at different conditions. Moreover, from Fig.10, it can be observed that the degradation pattern become inherently nonlinear precisely when the LIB reaches its end-of-life predefined point. However, the ELM model still provides high accurate prediction results even if the degradation pattern is extensively nonlinear, as seen in Fig.10 (a) and (b), which indicate that the proposed method can be applied for capacity estimation of the second life of LIB.

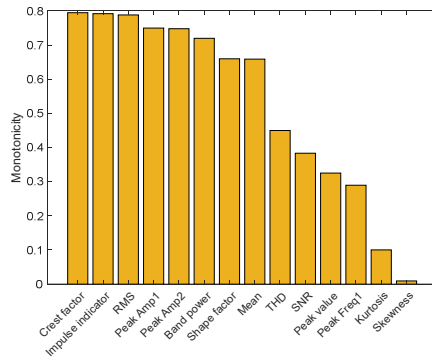


Fig. 7. Monotonicity results.

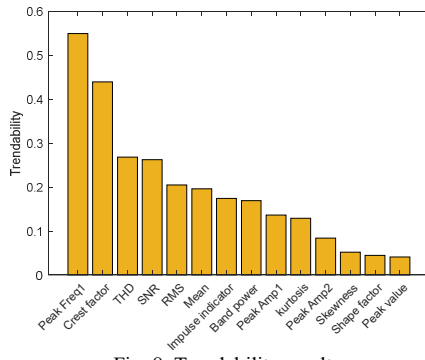


Fig. 8. Trendability results.

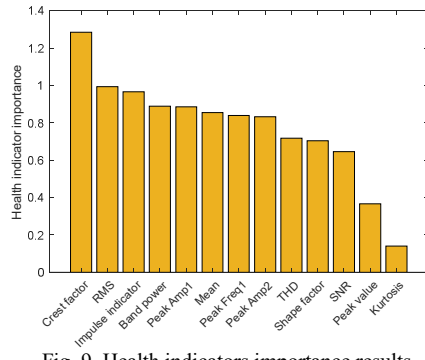


Fig. 9. Health indicators importance results.

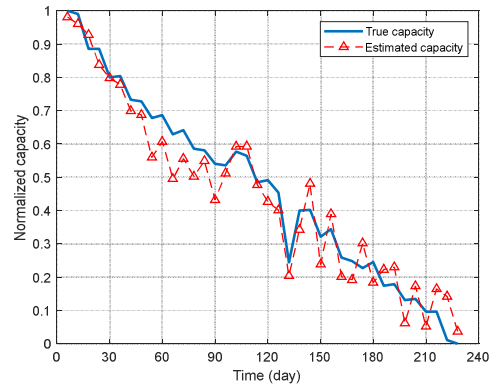
Table II. The operation condition of the test set.

Test set ID	Operation condition
RW4	Cycled under RW at 25°C.
RW7	Cycled under RW 25°C, with variable charging period.
RW26	Cycled under RW 40°C, with high rate-discharge.
RW23	Cycled under RW 40°C, with low-rate discharge.

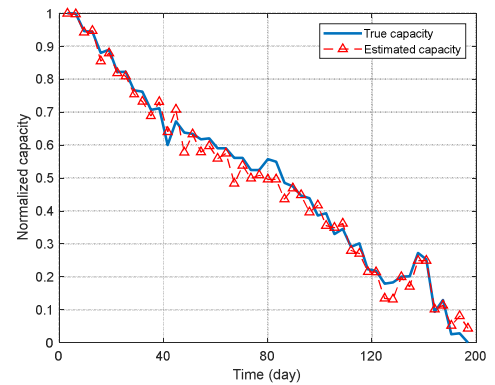
Moreover, Table III illustrates the performance evaluation results for capacity estimation in RMSE and MAE for each LIB. The maximum RMSE error is 1.3% for RW4, and the least RMSE error is 0.53% for RW26. Therefore, from the errors values, it is clear that the extracted health indicators from both domains combined with the ELM model provide high accurate prediction results for LIBs cycled under numerous operating conditions.

Table III. Performance evaluation of ELM model for capacity estimation

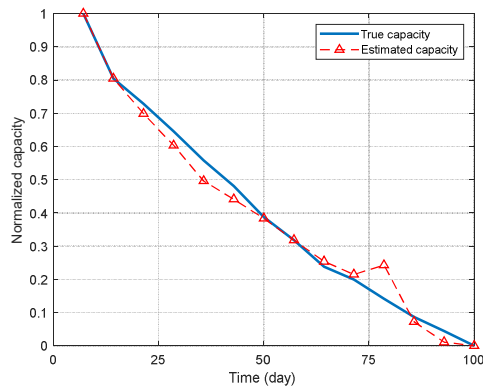
Battery ID	RMSE (%)	MAE (%)
RW4	1.3	1.1
RW7	0.91	0.86
RW26	0.53	0.61
RW23	0.62	0.76



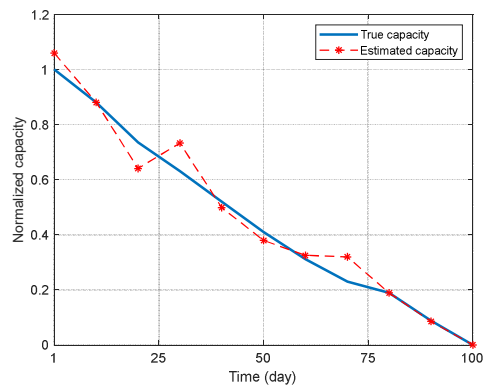
(a) RW 4



(b) RW 7



(c) RW 26



(d) RW 23

Fig. 10. Capacity prediction results of ELM model.

IV. CONCLUSION

This study investigates the performance of the time domain and frequency domain health indicators to track the capacity degradation of LIBs cycled under a randomised current profile. 14 health indicators were extracted from both domains and then ranked based on monotonicity and trendability prognostics metrics. All the health indicators greater than 0.8 were used to train a machine-learning model called ELM. The results of the trained model shown that the extracted health indicators from both domains can effectively capture the capacity degradation of LIB with a maximum of 1.3% and 1.1% RMSE and MAE errors, respectively.

REFERENCES

- [1] B. Wu, W. D. Widanage, S. Yang, and X. Liu, "Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems," *Energy and AI*, vol. 1, p. 100016, 2020.
- [2] C. R. Birkel, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, "Degradation diagnostics for lithium ion cells," *Journal of Power Sources*, vol. 341, pp. 373-386, 2017.
- [3] M. S. H. Lipu *et al.*, "A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations," *Journal of Cleaner Production*, vol. 205, pp. 115-133, 2018.
- [4] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *Journal of Power Sources*, vol. 226, pp. 272-288, 2013.
- [5] M. a. El-Dalahmeh, M. Al-Greer, M. d. El-Dalahmeh, and M. Short, "Smooth particle filter-based likelihood approximations for remaining

- useful life prediction of Lithium-ion batteries," *IET Smart Grid*, vol. 4, no. 2, pp. 151-161, 2021.
- [6] R. Xiong, L. Li, and J. Tian, "Towards a smarter battery management system: A critical review on battery state of health monitoring methods," *Journal of Power Sources*, vol. 405, pp. 18-29, 2018.
- [7] M. A. Hannan, M. M. Hoque, A. Hussain, Y. Yusof, and P. J. Ker, "State-of-the-Art and Energy Management System of Lithium-Ion Batteries in Electric Vehicle Applications: Issues and Recommendations," *IEEE Access*, vol. 6, pp. 19362-19378, 2018.
- [8] Y. Zou, X. Hu, H. Ma, and S. E. Li, "Combined State of Charge and State of Health estimation over lithium-ion battery cell cycle lifespan for electric vehicles," *Journal of Power Sources*, vol. 273, pp. 793-803, 2015.
- [9] J. Li, K. Adewuyi, N. Lotfi, R. G. Landers, and J. Park, "A single particle model with chemical/mechanical degradation physics for lithium ion battery State of Health (SOH) estimation," *Applied Energy*, vol. 212, pp. 1178-1190, 2018.
- [10] M. Einhorn, F. V. Conte, C. Kral, and J. Fleig, "Comparison, Selection, and Parameterisation of Electrical Battery Models for Automotive Applications," *IEEE Transactions on Power Electronics*, vol. 28, no. 3, pp. 1429-1437, 2013.
- [11] R. R. Thakkar, Y. S. Rao, and R. R. Sawant, "Performance Analysis of Electrical Equivalent Circuit Models of Lithium-ion Battery," in *2020 IEEE Pune Section International Conference (PuneCon)*, 16-18 Dec. 2020.
- [12] H. Tian, P. Qin, K. Li, and Z. Zhao, "A review of the state of health for lithium-ion batteries: Research status and suggestions," *Journal of Cleaner Production*, vol. 261, p. 120813, 2020.
- [13] C. Hu, G. Jain, C. Schmidt, C. Strief, and M. Sullivan, "Online estimation of lithium-ion battery capacity using sparse Bayesian learning," *Journal of Power Sources*, vol. 289, pp. 105-113, 2015.
- [14] P. Guo, Z. Cheng, and L. Yang, "A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction," *Journal of Power Sources*, vol. 412, pp. 442-450, 2019.
- [15] S. Shen, M. Sadoughi, X. Chen, M. Hong, and C. Hu, "A deep learning method for online capacity estimation of lithium-ion batteries," *Journal of Energy Storage*, vol. 25, p. 100817, 2019.
- [16] M. El-Dalahmeh, M. Al-Greer, M. a. El-Dalahmeh, and M. Short, "Time-Frequency Image Analysis and Transfer Learning for Capacity Prediction of Lithium-Ion Batteries," *Energies*, vol. 13, no. 20, p. 5447, 2020.
- [17] P. Venugopal, "State-of-Health estimation of li-ion batteries in electric vehicle using IndRNN under variable load condition," *Energies*, vol. 12, no. 22, p. 4338, 2019.
- [18] S. Khaleghi, Y. Firouz, J. Van Mierlo, and P. Van den Bossche, "Developing a real-time data-driven battery health diagnosis method, using time and frequency domain condition indicators," *Applied Energy*, vol. 255, p. 113813, 2019.
- [19] B. Bole, C. S. Kulkarni, and M. Daigle, "Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomised use," *SGT, Inc. Moffett Field United States*, 2014.
- [20] J. B. Ali and L. Saidi, "A new suitable feature selection and regression procedure for lithium-ion battery prognostics," *International Journal of Computer Applications in Technology*, vol. 58, no. 2, pp. 102-115, 2018.
- [21] R. Kunjir, V. Bhanuse, J. Kulkarni, and S. Patankar, "Determination of Deformation of Steel Plate Using Welch's Periodogram Estimate," in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 14-15 June 2018 2018, pp. 1169-1174, doi: 10.1109/ICCONS.2018.8662967..
- [22] G.B. Huang, Q.Y. Zhu, and C.K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489-501, 2006.