Battery Lifetime Prediction and Capacity Estimation Based on Entropy and Bayesian Neural Networks

2022 IEEE 5th International Conference on Electronics Technology (ICET) | 978-1-6654-8508-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICET55676.2022.9825351

Meizhu CHEN
Institute of Electronic Engineering, IEE
China Academy of Engineering Physics, CAEP
Mianyang, China
chenmeizhu1996@qq.com

Xu SONG
Institute of Electronic Engineering, IEE
China Academy of Engineering Physics, CAEP
Mianyang, China
songxuuestc@163.com

Yunhong CHE*
College of mechanical and vehicle engineering
Chongqing University
Chongqing, China
cyh@cqu.edu.cn

***Abstract*—Capacity degradation of lithium-ion batteries influences their service abilities as storage systems. Lifetime prediction and capacity estimation help prognose the health status of batteries. Accurate predictions would guide early maintenance, which ensures the battery operates in a safe range. This paper proposed a novel method to predict the lifetime and estimate the capacity based on entropy information features and Bayesian neural networks. Partial voltage curves are used for health indicators extraction, which recognizes the degradation properties and builds the capacity model as well. The feature can also use for battery lifetime prediction directly at the same time. The probabilistic Bayes neural network is proposed for both the lifetime and capacity models construction, which provides the estimated results together with probabilistic uncertainty. The results show that the predictions are accurate with only three known cycles of the testing batteries.**

Keywords—battery, lifetime prediction, capacity estimation, health indicators, bayesian neural network

I. Introduction

Transportation electrification is regarded as the global industrial trend, where batteries are used as the main energy storage systems [1]. The energy and power of the batteries affect the performance during usage, but they would suffer attenuation due to the capacity degradation [2]. The battery systems would suffer safe issues such as thermal runaway if they were overused. Therefore, it is significant to get accurate lifetime prediction and state of health estimation to help the battery management system ensure the batteries work in the safe range [3]. Generally, methods for battery lifetime prediction or capacity estimation can be divided into model-based and data-driven methods [4].

The first kind of model-based method is to fit an empirical model using testing data for the degradation representation. It is simple but has poor generalization ability. The main idea for the other two model-based methods (equivalent circuit model and physical model) is to build accurate models to fit the electrochemical behavior of batteries. Then, the battery capacity and lifetime can be predicted via optimization algorithms. But it is difficult to ensure the model accuracy and low computational burden at the same time. Compared to the model-based methods, data-driven methods show better generalization ability, better convenience application, and have satisfactory accuracy and robustness, which makes them develop rapidly in recent years.

Data-driven methods for battery lifetime prediction and capacity (or state of health, SOH) estimation mainly include data preprocessing, model training, and predictive validation [5]. In data preprocessing, the raw data are collected and cleaned up to form the input and output of the machine learning model. Then, the data-driven model is trained using the processed data. Finally, the model is used for lifetime prediction and capacity estimation when new data are obtained. The output of the model is generally the lifetime or capacity. While three main methods can be used to form the input according to the published papers. The first method utilizes the capacity sequence for model building, which uses the capacities of the former a few cycles as the input [6]. The second method directly uses the raw data of voltage, current, and temperature as the input, then uses some deep learning methods such as convolutional neural network and recurrent neural network to extract the intrinsic information automatically [7]. The last one extracts some health indicators (HIs) from measured parameters to reflect the aging status of the battery and then builds the relationship between those HIs and lifetime or capacity via machine learning [5]. The HIs based method draws more attention because online extraction can be achieved, and low computational cost is needed. However, the main challenge is the online useful HIs extraction method, which enables the extracted HIs to have high correlations with capacities and can be applied in real applications. The second key task of the data-driven battery capacity estimation method is machine learning. The neural network is one main category of algorithms to fit the regression model, which can suit any kind of linear or nonlinear mapping [8]. However, most of them just provide a specific prediction but are not probabilistic. In addition, many real data are required for the supervised model training which is hard to obtain in real applications.

Therefore, to fill the problems stated above, this paper proposes a novel Bayesian neural network-based battery lifetime and capacity estimation method. Firstly, partial voltage curve and charge quantity data are used for entropy-based HIs extraction, which can be used for degradation characteristic identification and capacity estimation at the same time. Then, the Bayesian neural network is proposed for the model training, which could provide probabilistic prediction results of the predicted lifetime and estimated capacity. Finally, the transfer learning strategy is adopted to retrain the capacity estimation model only using several checkpoints while satisfactory estimation results can be obtained. The remainder of this paper is organized as follows; the data set and the HI extraction method are introduced in section II. Then the prediction method is described in section III. Following that, the results are provided and discussed in section IV. Finally, the main conclusion is summarized in section V.

II. Battery Data Set and Health Indicator Extraction

The public data set collected from ref. [9] are used for the prediction verification in this paper, where 124 batteries are included in the total. During the aging process, each battery cell is charged in two-stage fast charging profiles to 80% state of charge and the constant current and constant voltage charging is added to make the battery fully charged. Then every battery is discharged under 4 C current.

In this paper, the partial charge quantity-voltage (Q-V) curve is used for the HIs extraction. The partial Q curve during the voltage interval 2.85-3.25V with 30 segments is used. To extract the HIs, the shannon entropy of the Q curve of each cycle (called ShanEnQ) and Q difference curve between each cycle to 10th cycle (called ShanEnDQ^2) are calculated by the following expression [10, 11]. The entropy of variable P is indicated by equation (1),

$$En = \sum_{j=1}^{n} p(j) * \log(p(j)) \tag{1}$$

The degradation curves of battery capacity, the extracted ShanEnQ and ShanEnDQ^2 are shown in Fig. 1 (a). It clearly shows that the degradations of the proposed HIs almost have the same characteristic as battery capacities.

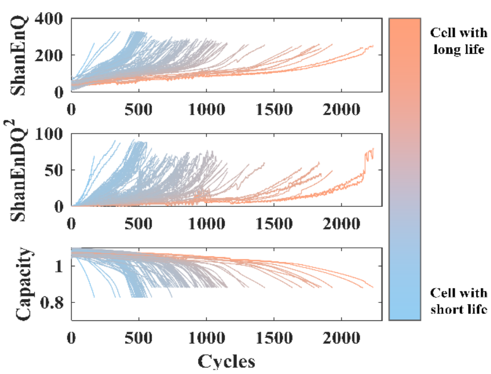
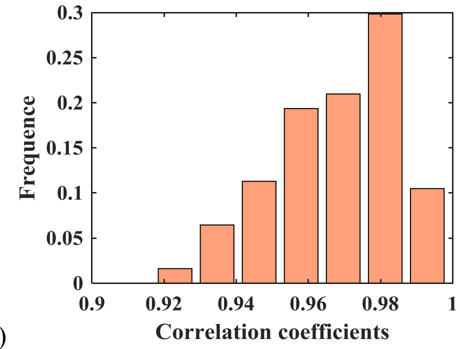


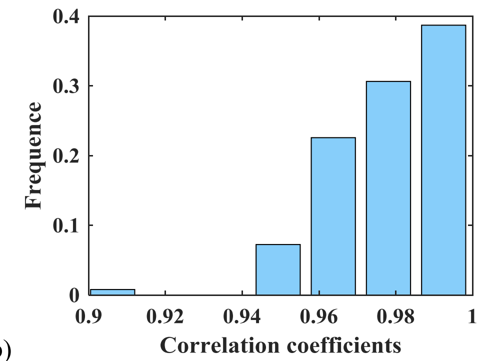
Fig. 1. Variations of extracted HIs and battery capacity.

Pearson correlation coefficient is used to analyze the correlation between the extracted ShanEnQ / ShanEnDQ^2 and battery capacity. The results are shown in Fig. 2 (a) and Fig. 2 (b). It can be seen that most of the correlation coefficients are concentrated in the interval of [0.9 1]. The statistical results of the correlation coefficients are shown in Table 1, where the mean value of ShanEnQ and ShanEnDQ^2 is 0.97 and 0.98 respectively. In both cases, the proportion that the correlation coefficients are greater than 0.9 is larger than 0.99, which means the extracted HIs have high correlation with battery capacities.

For battery lifetime prediction. The early two cycles of 10 and 100 are used for ShanEnDQ^2 extraction to fit the regression between it and lifetime. The log-log plot of ShanEnDQ^2 and the life cycle of all batteries are shown in Fig. 3. The results show that the correlation coefficient is 0.907, which also has a satisfactory linear relationship.



(a)



(b)

Fig. 2. (a) probability distribution of correlation coefficients between ShanEnQ and capacities (b) probability distribution of correlation coefficients between ShanEnDQ^2 and capacities.

TABLE I. Statistical Results of Correlation Analysis between *STD_Q*/*STD_DQ* and Battery Capacity

HI	Mean	Ratio greater than 0.9	Ratio greater than 0.95	Ratio greater than 0.98
ShanEn_Q	0.97	1	0.85	0.26
ShanEn_dQ^2	0.98	0.99	0.95	0.57

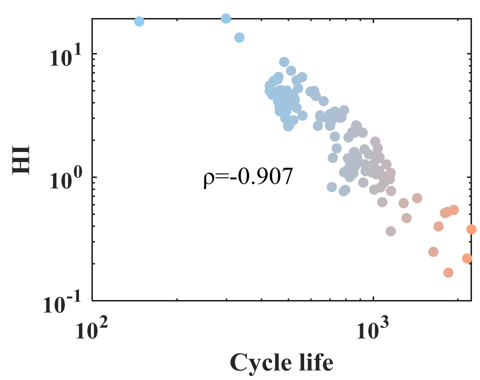


Fig. 3. Correlation analysis between ShanEnDQ^2 and battery lifetime.

III. Prediction Method

A. Data-driven Modeling

The neural network is widely used in battery state estimation, which shows great accuracy and robustness. The basic function of a neural network is [4],

$$y = activate\_fun(\sum w_i * x_i + b_i) \quad (2)$$

where x and y are input and output respectively, the *aticve_fun* is the activate function, w and b are weight and bias respectively. However, in the conventional neural network, the w and b have specific values, which make the output just have a specific value. The Bayes neural network, by contrast, sets the weight and bias as distributions. And the training process optimizes the distribution instead of one specific value. The loss function is defined as follows [12],

$$\text{negative\_loglikelihood}(y, p\_y) = -p\_y.\text{log\_prob}(y) \quad .(3)$$

The output is determined by the predicted mean and standard deviation, which makes the prediction to be probabilistic.

For battery lifetime prediction, the input and output are the ShanEnDQ^2 between 100 cycle and 10 cycle, and the battery lifetime respectively. For battery capacity estimation, the inputs are ShanEnQ and ShanEnDQ^2 and the output is the capacity for each cycle. In addition, the transfer learning strategy is used to improve the capacity estimations. The output layer is set trainable while the hidden layer is freezed after training using the training battery. The pre-trained model is migrated to the testing battery and then retrained by several checkpoints for the capacity estimation.

B. Prediction Framework

The proposed prediction framework is illustrated in Fig. 4. Firstly, the data set is divided into training data and testing data. Then, the HIs proposed above are extracted. Next, the battery capacity estimation model and lifetime prediction model are trained by Bayesian neural networks. For testing battery prediction, the lifetime is predicted by the trained model. While the capacity estimation model can be further retrained by only a few data, and then the capacity is estimated by the updated model, which makes the estimates more accurate.

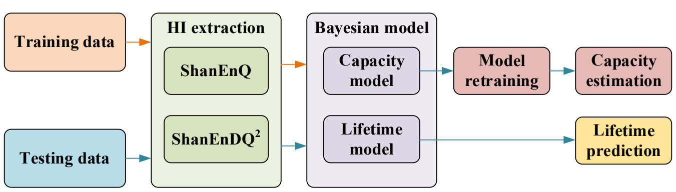


Fig. 4. Prediction framework for battery lifetime prediction and capacity estimation based on entropy feature and Bayesian neural network.

IV. Results and Discussion

In this section, the prediction results of the proposed method are provided and discussed. Firstly, the prediction of battery lifetime is verified. The whole data set is divided into 60% for model training and the remaining 40% for testing. The prediction results are shown in Fig. 5. The y axis is the log value of battery lifetime. It shows that the predictions follow the real values satisfactorily. The 95% confidence intervals (CI) cover the gaps between the predictions and real values. The mean relative error of these 50 battery cells is only 13.4%, which means accurate predictions are obtained.

Then, the capacity estimation results are provided and discussed. The base model is trained by the two batteries those have most similar battery lifetimes to the predicted lifetime of the testing battery. Three batteries are randomly selected for demonstration, which have three representative life span for the whole battery data set (Cell 42 for short, Cell 121 for medium, and Cell 124 for long). When the models trained by the selected batteries are used for the capacity estimation of testing batteries directly, the results are shown in Fig. 6, where the results of Cell 42, Cell 121, and Cell 124 are shown in Fig. 6(a), Fig. 6(b), and Fig. 6(c) respectively. The numerical results of the estimated errors are listed in Table. 2. The results show that the estimations almost follow the real values accurately although the CIs fail to cover the real values at some points. The estimated mean absolute error (MAE) and root mean square error (RMSE) are 0.40% and 0.48 % for Cell 42, 0.63% and 0.83% for Cell 121, and 0.73% and 0.92% for Cell 124, respectively.

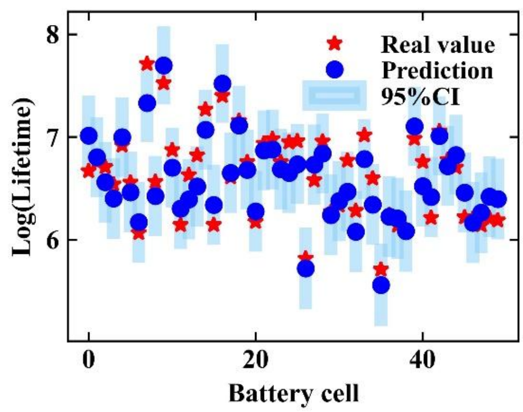


Fig. 5. Battery lifetime prediction results.

To improve the accuracy and generalization of the estimation model, the transfer learning strategy is proposed to update the model with only three checkpoints, which are 10 cycle, 100 cycle, and 50% cycle. The estimation results at this moment are shown in Fig. 7. The numerical results of the estimated errors are also listed in Table. 2. It shows that the predictions get closer to the real values and the 95% CIs cover the real values during the whole life span, which means that the estimations become more accurate and reliable with only a little retraining process. The estimated mean absolute error (MAE) and root mean square error (RMSE) are 0.23% and 0.29 % for Cell 42, 0.42% and 0.64% for Cell 121, and 0.32% and 0.39% for Cell 124, respectively. All the errors have been reduced significantly. The capacity estimation results for these three representative batteries verify that the Bayesian neural network has good performance for probabilistic estimation, and the transfer learning shows great ability to improve the estimation accuracy and reliability even using a few data.

TABLE II. Estimated MAE and RMSE (%) for the Three Batteries

HI	Cell 42		Cell 121		Cell 124	
	MAE	RMSE	MAE	RMSE	RMSE	RMSE
Proposed	0.23	0.29	0.42	0.64	0.32	0.39
Base model	0.40	0.48	0.63	0.83	0.73	0.92

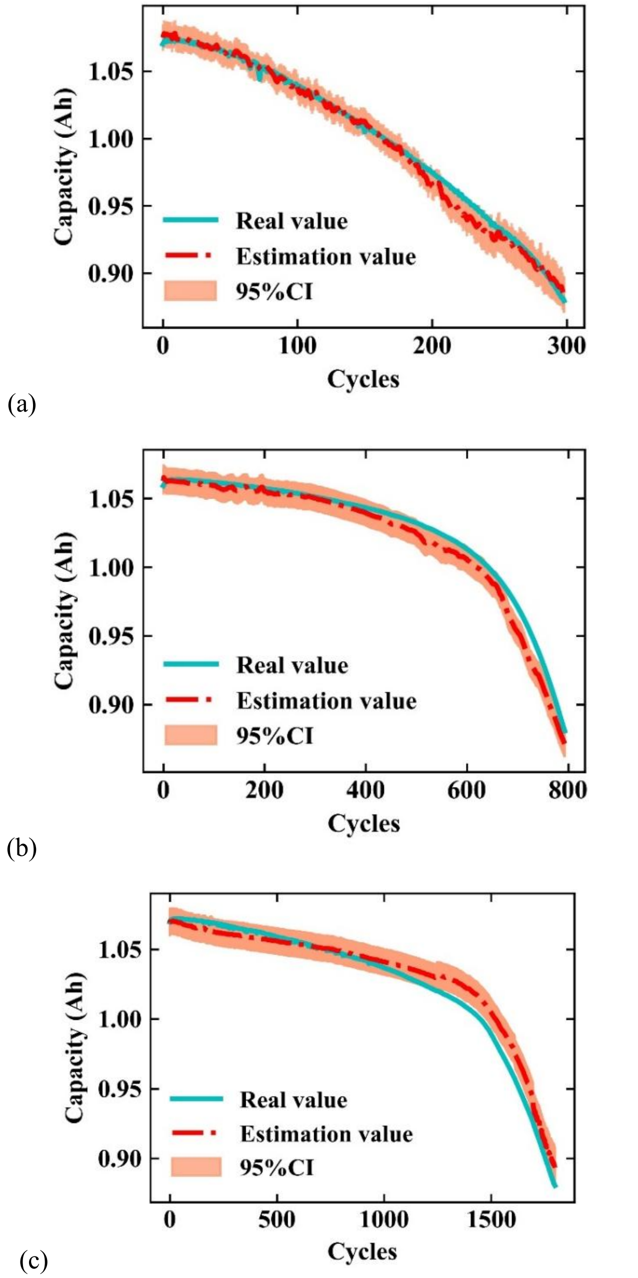


Fig. 6. Capacity estimation results for (a) Cell 42, (b) Cell 121, and (c) Cell 124.

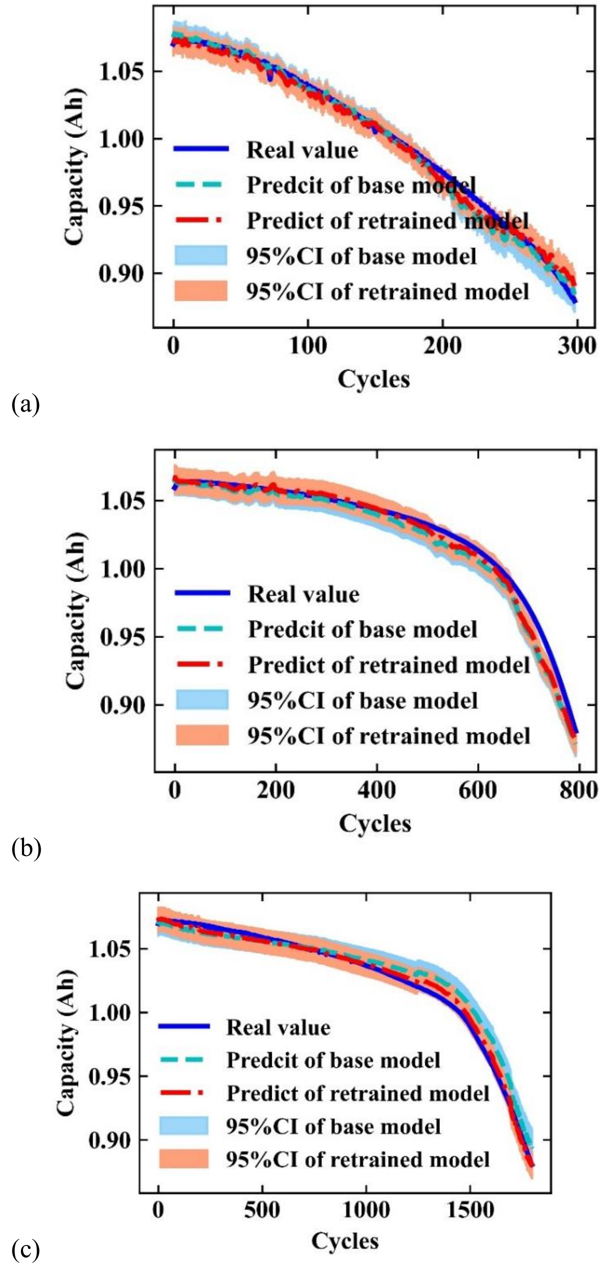


Fig. 7. Capacity estimation results after transfer learning with only three poitns for (a) Cell 42, (b) Cell 121, and (c) Cell 124.

V. Conclusion

Lifetime prediction and capacity estimations are the major scientific and industrial challenges that are regarded as the main bottlenecks limiting the development of prognostic and health management for batteries. This paper proposes a novel method for lifetime prediction and capacity estimation with the same framework. The entropy-based HIs are extracted from partial curves and used for both lifetime modeling and capacity estimation modeling. The correlation analysis results show that more than 0.99 of the correlation coefficients between HIs and capacities are larger than 0.9. The correlation coefficient between the HIs and battery lifetimes is also 0.907. The Bayesian neural network is proposed for data-driven mapping. And transfer learning strategy is proposed for the model improvement with only three checkpoints. The results show that the whole capacity curve could be accurately estimated with MAE less than 0.42% and RMSE less than 0.64%. Future work would focus on the implementation under various aging conditions and different battery types.

References

[1] L. Hu, X. Hu, Y. Che, F. Feng, X. Lin, and Z. Zhang, "Reliable state of charge estimation of battery packs using fuzzy adaptive federated filtering," *Applied Energy,* vol. 262, 2020.

[2] Y. Che *et al.*, "State of health prognostics for series battery packs: A universal deep learning method," *Energy,* vol. 238, 2022.

[3] Y. Che, Z. Deng, X. Tang, X. Lin, X. Nie, and X. Hu, "Lifetime and Aging Degradation Prognostics for Lithium-ion Battery Packs Based on a Cell to Pack Method," *Chinese Journal of Mechanical Engineering,* vol. 35, no. 1, 2022.

[4] X. Sui, S. He, S. B. Vilsen, J. Meng, R. Teodorescu, and D.-I. Stroe, "A review of non-probabilistic machine learning-based state of health estimation techniques for Lithium-ion battery," *Applied Energy,* vol. 300, 2021.

[5] X. Hu, Y. Che, X. Lin, and S. Onori, "Battery Health Prediction Using Fusion-Based Feature Selection and Machine Learning," *IEEE Transactions on Transportation Electrification,* vol. 7, no. 2, pp. 382-398, 2021.

[6] K. Liu, Y. Shang, Q. Ouyang, and W. D. Widanage, "A Data-Driven Approach With Uncertainty Quantification for Predicting Future Capacities and Remaining Useful Life of Lithium-ion Battery," *IEEE Transactions on Industrial Electronics,* vol. 68, no. 4, pp. 3170-3180, 2021.

[7] P. Li *et al.*, "State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network," *Journal of Power Sources,* vol. 459, 2020.

[8] Y. Che, Z. Deng, X. Lin, L. Hu, and X. Hu, "Predictive Battery Health Management With Transfer Learning and Online Model Correction," *IEEE Transactions on Vehicular Technology,* vol. 70, no. 2, pp. 1269-1277, 2021.

[9] K. A. Severson *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy,* vol. 4, no. 5, pp. 383-391, 2019.

[10] Y. Che, A. Foley, M. El-Gindy, X. Lin, X. Hu, and M. Pecht, "Joint Estimation of Inconsistency and State of Health for Series Battery Packs," *Automotive Innovation,* 2021.

[11] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling," *IEEE Transactions on Industrial Electronics,* vol. 63, no. 4, pp. 2645-2656, 2016.

[12] S. Pavel, S. Chris, B. Jacob, and V. D. Joshua. (2019). *Regression with Probabilistic Layers in TensorFlow Probability*. Available: https://blog.tensorflow.org/2019/03/regression-with-probabilistic-layers-in.html?hl=zh-cn