

Combining machine learning algorithms and an incremental capacity analysis on 18650 cell under different cycling temperature and SOC range

John Lai*, David Chao, Alvin Wu, Carl Wang

Underwriters Laboratories Taiwan Co. Ltd., No. 260, Daye Rd., Beitou Dist., Taipei City 112, Taiwan

Abstract. A novel way to apply machine learning algorithms on the incremental capacity analysis (dQ/dV) is developed to identify battery cycling conditions under different temperatures and working SOC ranges. Batteries are cycled under each combination of temperatures (-10°C, 25°C, 60°C) and SOC ranges (0-10%, 25-75%, 90-100%, 0-100%) up to 60 equivalent cycles. The discharge data is transformed into dQ/dV-V curve and its features of the peaks and valleys are further taken for machine learning. Both supervised and unsupervised machine learning algorithms (PCA and LDA) are applied to classify batteries in terms of temperature or SOC range. The results reveal that batteries cycled under different temperatures can be identified separately regardless of the working SOC range. When splitting 60 samples with a ratio of training set equals to 0.85, the remaining test set gives an identification accuracy of 89% in temperature and 67% in working SOC range.

1 Introduction

Rechargeable batteries have been widely used in our daily life, for example, from a stationary electrical storage system connected to an electric grid to portable devices such as smartphones and laptops. Among several types of rechargeable batteries, the lithium-ion battery is the dominant one with high energy density and working voltage. Nowadays, lithium-ion batteries can be easily found on 3C products and electric vehicles. The real battery usage in these applications suffers degradation from different temperatures and working SOC ranges due to environmental temperature, device design, and user's habit that makes life evaluation difficult [1-3]. There have been massive researches on battery aging behavior under different temperature and working SOC ranges owing to their significant impact on battery degradation mechanisms, which applies huge uncertainty and variation in battery lifespan and safety [4-8]. Thus, there has been numerous inspection methods developed to evaluate the cell aging status. One of the well-known methods is called incremental capacity (dQ/dV) analysis, which transfers charging/discharging data into dQ/dV-V curves followed by observing peak shifts or amplitude changes [9-11]. Although many tests have been conducted and shifts of peaks have been explained in previous studies, an easy, general and accurate way to identify or predict battery cycling conditions remains unrevealed. Recently, machine learning as well as deep learning have privileged in all industry including lithium-ion battery to address complex problems by processing classification or providing more insight from only data. It has been largely studied in state of health (SOH) and

remaining useful life (RUL) prediction [12-13]. Hence, this study tries to apply machine learning algorithms on features extracted from dQ/dV-V curve to investigate changing directions of these features under different cycling temperatures and working SOC ranges in pursuit of the identification of battery aging status.

2 Experimental procedure

Commercial Panasonic 18650 cell, 3350 mAh, was selected for the test. All cells were conducted 3-cycle conditioning before experiments to check cell consistency, these conditioned samples were called fresh cells hereafter. Each cell corresponded to one of combinations composed of different temperature (-10°C, 25°C, 60°C) and working SOC ranges (0-10%, 25-75%, 90-100%, 0-100%). The cells were cycled by CC discharge mode and CC-CV charge mode with 0.2C within a voltage range between 2.5V and 4.2V and a cutoff current of 0.02C. For 0-10% SOC, cells were fully discharged before charge/discharge cycles. For 25-75% SOC, cells were discharged to 25% SOC and then followed by charge/discharge cycles. For 90-100% and 0-100% SOC, cells were directly cycled from a fully charged state. To ensure all cells are cycled under proper SOC ranges, the maximum capacity at each cycling temperature was measured as a baseline to estimate charged/discharge time. The charge/discharge time and cycles for 15 equivalent cycles under different temperature and SOC ranges are summarized in Table. 1. After every 15 equivalent cycles, cells were conducted a 2-cycle retention capacity test according to battery specification,

* Corresponding author: john.lai@ul.com

0.5C charge with a cutoff current equals to 0.02C and 0.2C discharge, followed by a 24-hour rest to record OCV drop. The 2nd discharge data in the retention

capacity test was further used for incremental capacity (dQ/dV) analysis and machine learning.

Table 1. Cycling parameters for 15 equivalent cycles under different temperatures and SOC ranges.

SOC range	-10 °C		25 °C		60 °C	
	C/D time	Cycles	C/D time	Cycles	C/D time	Cycles
0-10%	1504s	150	1781s	150	1841s	150
25-75%	7520s	30	8905s	30	9205s	30
90-100%	1504s	150	1781s	150	1841s	150
0-100%	*	15	*	15	*	15

*Cells stop charge/discharge when reaching termination conditions.

3 Experimental procedure

3.1 Battery cycling data

The discharge capacity and 24-hour OCV drop of fresh cells (3-cycle conditioning) and cycled cells (15, 30, 45, 60 equivalent cycles) are summarized in Fig. 1 and Fig. 2 labeled with temperature and working SOC range separately. In Fig. 1, cells cycled under -10 °C and 60 °C show lower retention capacity but cannot be separated clearly, and cells cycled at 25 °C shows higher capacity but still partially mix with cells cycled at 60 °C. Some of the cells cycled under -10 °C and 60 °C have an obvious higher OCV drop after 24-hour rest but the remaining cells have lower OCV drop than fresh cells and are mixed with cells cycled at 25 °C. In Fig. 2, 0-10% SOC cycled cell has slightly higher retention capacity while other SOC ranges are mixed together. Similarly, most of the cells have OCV drops lower than fresh cells and are inseparable. Based on the result, the aging effect caused by cycling temperature or working SOC range factors is not easy to be identified by simply retention capacity and 24-hour OCV drop data.

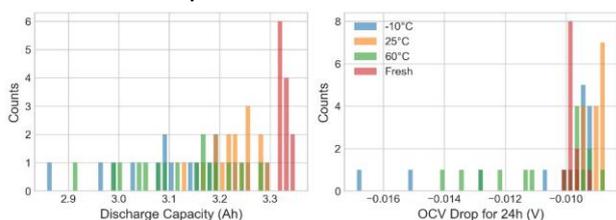


Fig. 1. The discharge capacity and 24-hour OCV drop of fresh and cycles under different cycling temperature.

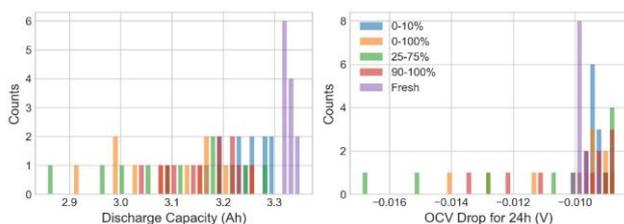


Fig. 2. The discharge capacity and 24-hour OCV drop of fresh and cycles under different cycling SOC range.

3.2 Features extracted from dQ/dV-V curve

The incremental capacity analysis and feature extraction process applied on 2nd discharge data in the retention capacity test are summarized in Fig. 3(a) and follows the steps: (i) slice original discharge data into multiple segments with a voltage interval of 0.011V from 4.2V to 2.5V; (ii) calculate the capacity difference (dQ) in each segment and divided by 0.011V (dV) to obtain dQ/dV value; (iii) calculate mean voltage of each segment; (iv) take dQ/dV as y-axis and mean voltage as x-axis to plot dQ/dV-V curve. The dQ/dV-V curves of fresh cells (red solid line) and cycled cells (black dash line) are shown in Fig. 3(b). The study of dQ/dV-V curve generally focuses on the positions of peaks and valleys because they represent a phase transformation in cathode or intercalation of lithium into a graphite anode that is highly-related to cell aging behavior. Hence, four peaks (P1, P2, P3, P4) and three valleys (V1, V2, V3) are marked in Fig. 3(b), and their x-values and y-values are used as features for algorithms to make grouping or classification later in the sections 2.4 and 2.5. For example, the x-value and y-value of peak4 (P4) are marked as VP4 and IP4 respectively. From the dQ/dV-V curves, the fresh cells with uniform curves usually indicate the cell quality is consistent and it serves as a baseline to observe how curve shifts after cells are cycled. After being aged, cells show the following trends: peak1 (P1) shifts to an upper-left place; peak2 (P2), peak3 (P3), peak4 (P4) and valley2 (V2) shift to a lower place; valley1 (V1) and valley3 (V3) shift to both an upper and a lower place. It should be noticed that the curves near peak3 (P3), valley3 (V3), and peak4 (P4) have large variation without a consistent trend, which might be caused by different cycling conditions.

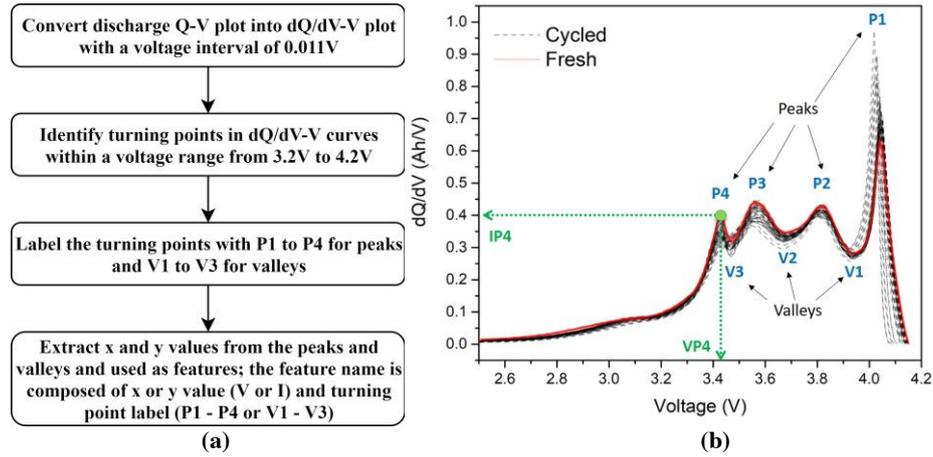


Fig. 3. (a) the flowchart of data processing from Q-V plot into dQ/dV-V plot and feature extraction from dQ/dV-V curve; (b) the dQ/dV-V curves of fresh (red solid line) and cycled cells (black dash line).

The features extracted from the dQ/dV-V curve are further standardized in order to improve algorithm results. To be more precise, the standardization rescales data with a mean of 0 and a standard deviation of 1, and it is widely used before machine learning algorithms to avoid being governed by a particular feature having a broad range of values. The normalized features labeled with cycling temperature and SOC range are shown in Figs. 4-7. For different cycling temperatures, 60 °C cycled cells have higher VV3 and IP1 than -10 °C cycled cells, but lower

IV3 than -10 °C cycled cells. Among all y-value (I) features, 25 °C cycled cells locate in the middle of the distribution. In contrast, for different cycling SOC ranges, only 0-10% SOC cycled cells show higher IV2 and lower IV3 clearly among all cells. The remaining SOC ranges don't provide any obvious trend. Overall, the trends mentioned above are unable to separate cycled cells under different temperatures or working SOC ranges. In next section, PCA will be applied to reduce the data dimension to visualize data distribution.

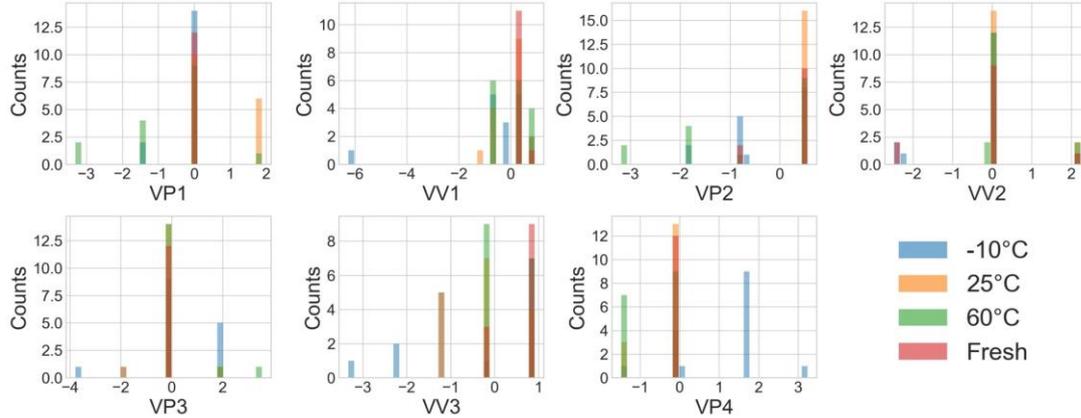


Fig. 4. Normalized data of x-value (V) of each peak (P1-P4) and valley (V1-V3) in dQ/dV-V curve labeled with cycling temperature.

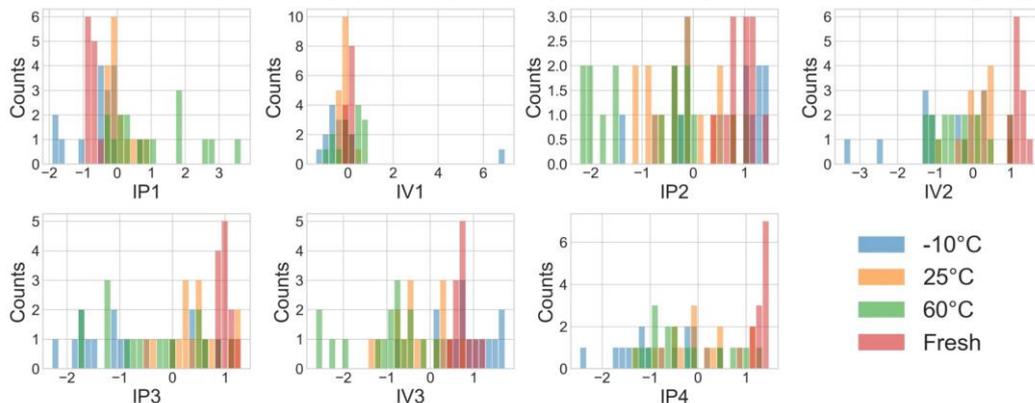


Fig. 5. Normalized data of y-value (I) of each peak (P1-P4) and valley (V1-V3) in dQ/dV-V curve labeled with cycling temperature.

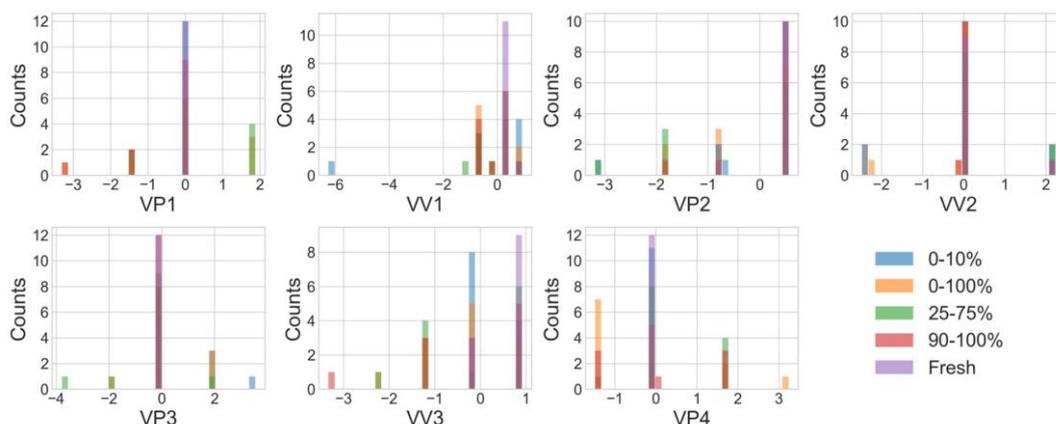


Fig. 6. Normalized data of x-value (V) of each peak (P1-P4) and valley (V1-V3) in dQ/dV-V curve labeled with cycling SOC range.

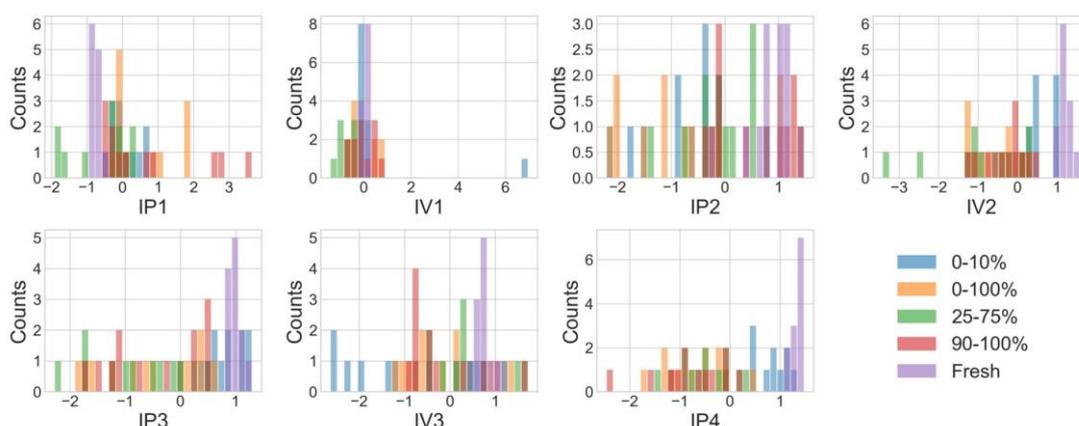


Fig. 7. Normalized data of y-value (I) of each peak (P1-P4) and valley (V1-V3) in dQ/dV-V curve labeled with cycling temperature.

3.3 Unsupervised algorithm (PCA) analysis

Principle component analysis (PCA) is an unsupervised algorithm used to keep maximum data variation when reducing data dimensions to provide a visualized data distribution. The PCA plots of cells cycled under different temperatures and working SOC ranges are shown in Fig. 8 (a) and (b) respectively. In Fig. 8 (a), the fresh cells have the most narrow distribution, while cells cycled under $-10\text{ }^{\circ}\text{C}$ and $60\text{ }^{\circ}\text{C}$ show the widest

distribution among all. It indicates that cells cycled under $-10\text{ }^{\circ}\text{C}$ and $60\text{ }^{\circ}\text{C}$ will cause much more variation in features than those cycled at $25\text{ }^{\circ}\text{C}$. Besides, a slight grouping can be observed along PC2-axis but there are still around half of the cells mixed together. In contrast, the PCA result of different SOC ranges provides not much information but wide and mixed distributions among all cycled cells. It can be inferred that the working SOC range effect may be suppressed by the temperature factor in feature variation.

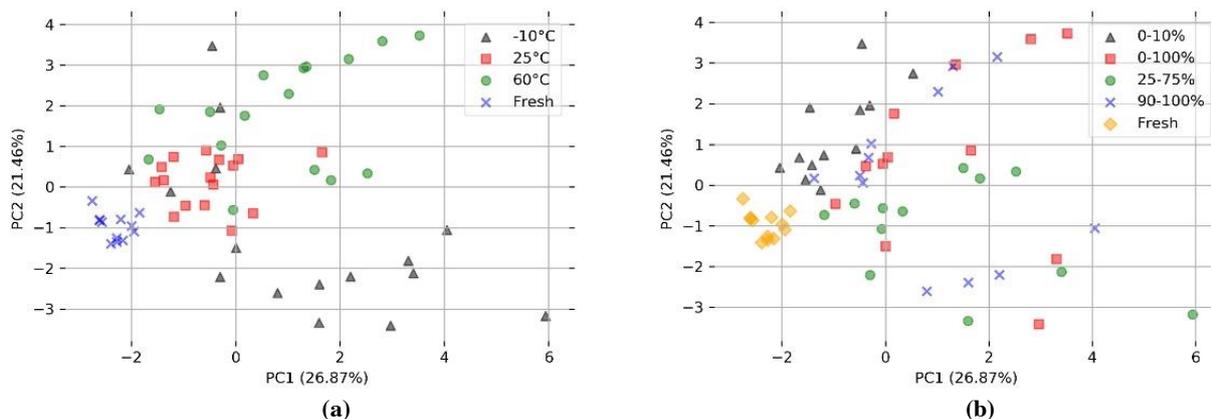


Fig. 8. PCA plots of cells cycled under different: (a) temperature (b) SOC range.

When looking into the cumulative explained variance chart shown in Fig. 9, the first two principal components only accounts for 48% explained variance. It indicates that multiple features cause similar variance so that even after rotating the principal component such as PC1 and PC2, the original variance is still unable to be effectively explained in a 2-D projected plane. Thus, a supervised algorithm is applied in the next section to obtain more explained variation and achieve the classification of cycling conditions.

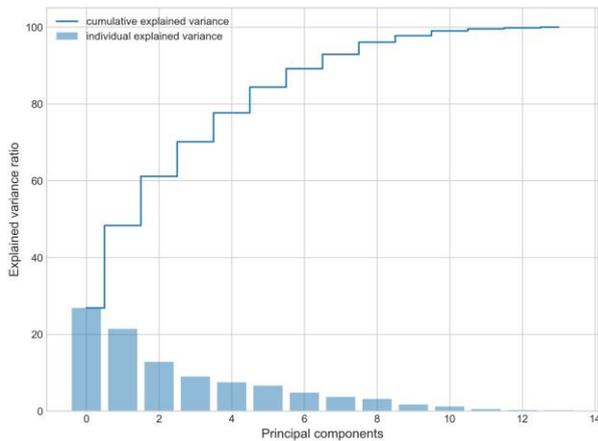


Fig. 9. PCA cumulative explained variance along with the number of principal components.

3.4 Supervised algorithm (LDA) analysis

Linear discriminant analysis (LDA) is a supervised algorithm used to maximize the gap between groups but minimize internal differences within a group. The LDA plots of different cycled temperatures and SOC ranges are given in Fig. 10 (a) and (b) respectively. According to the evaluation test in a test set ratio of 0.15, different cycling temperature gives an accuracy of 89%, while different SOC range only has an accuracy of 67%. In Fig. 10 (a), cells cycled under different temperatures can be obviously separated into three blocks along LD1-axis. It indicates that the LD1-axis is highly related to temperature and its values of 0.2 and -0.2 can be used as a simple way to identify cycling temperature. For example, cycled cells that have LD1 higher than 0.2 can be identified as 60 °C cycled cells. It should be noticed that this way of identifying cycling temperature applies to all working SOC ranges. On the other hand, cells cycled under different working SOC ranges are unable to be thoroughly categorized as four blocks. If the 90-100% SOC range is removed, the remaining SOC ranges can be identified along LD1-axis with specific values of 0.3 and -0.2. To understand the relationship between the LD1-axis and the dQ/dV-V curve, the eigenvalues are further discussed in the next section.

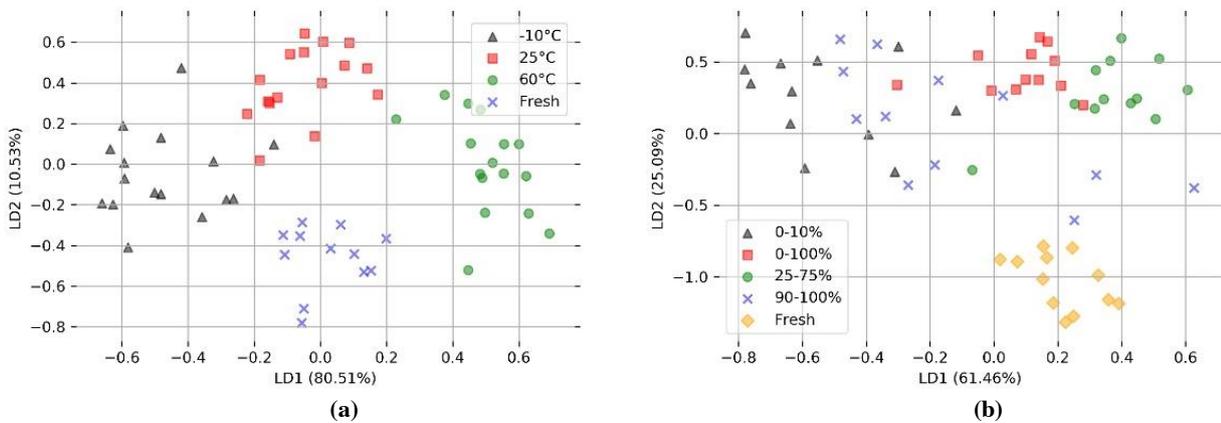


Fig. 10. LDA plots of cells cycled under different: (a) temperature (b) SOC range.

3.5 Insight Into dQ/dV-V curve from eigenvalues

In both PCA and LDA, the axes are also called eigenvectors, which are linear vectors composed of each feature multiplied by each corresponding eigenvalues. The higher the eigenvalue, the more important the corresponding feature in contributing explained variation. Generally in LDA, when a feature has a high eigenvalue, it represents the feature has more power to classify samples according to their labels. The eigenvalues of PCA and LDA (temperature and working SOC range) algorithms are summarized in Table. 2. From the LDA (temperature) result, the VP4, IP3, IV3, and IP4 features have high eigenvalues marked in bold in LD1-axis, and LD1-axis is shown capable of classifying cycling

temperature in Fig. 10 (a). Thus these features are highly related to cycling temperature. In addition, these features correspond to peak3 (P3), valley (V3), peak4 (P4) in dQ/dV-V curve, shown in Fig. 3(b), ranging from 3.4V to 3.6V. It indicates this specific voltage range in dQ/dV-V curve is highly temperature-dependent and its shifting behavior can be used to identify battery cycling temperature. On the other hand, we can observe the LD1 eigenvalues in LDA (SOC range) and those in LDA (temperature) are nearly the same. It indicates most of the explained variation in LDA (SOC range) might be caused by temperature factor. In other words, the cycling temperature has a more dominant impact than the working SOC range in the shifting behavior of dQ/dV-V curves.

Table 2. Eigenvalues of PCA and LDA algorithms.

Feature	PCA (Temperature & SOC range)		LDA (Temperature)		LDA (SOC range)	
	PC1 (26.87%)	PC2 (21.46%)	LD1 (80.51%)	LD2 (10.53%)	LD1 (61.46%)	LD2 (25.09%)
	VP1	0.20	0.20	0.16	0.13	0.19
VV1	0.12	0.16	0.04	-0.07	0.06	-0.05
VP2	0.29	0.05	-0.02	0.04	-0.09	0.10
VV2	0.11	-0.21	0.04	0.09	-0.01	0.08
VP3	0.10	-0.22	-0.06	-0.08	-0.06	0.04
VV3	0.28	-0.18	0.04	0.02	-0.13	0.17
VP4	-0.12	0.44	-0.24	0.23	-0.28	0.40
IP1	-0.13	-0.46	0.20	0.20	0.09	0.26
IV1	0.10	-0.30	0.05	-0.10	0.00	-0.05
IP2	0.14	0.36	0.16	-0.14	0.07	-0.39
IV2	0.50	-0.04	-0.17	-0.31	0.00	-0.43
IP3	0.49	-0.07	-0.54	0.75	-0.68	0.47
IV3	0.06	0.41	-0.25	0.04	0.28	0.00
IP4	0.46	0.05	0.67	-0.42	0.55	-0.37

4 Conclusion

The incremental analysis (dQ/dV) analysis is an effective way in diagnosing the battery aging behavior and a supervised algorithm is useful to classify cycling temperature and provide insight into the shifting behavior of dQ/dV-V curves. From the observation of retention capacity, 24-hour OCV drop, peak and valley positions in dQ/dV-V curve, some insignificant trends can be found but too complicated and insufficient to identify cycling temperature or working SOC range. By applying a supervised LDA algorithm, cells cycled under different temperatures (-10°C, 25°C, and 60°C) can be well separated into three blocks in a 2-D projected plane. According to an evaluation test, the identified accuracy reaches 89% in a test set ratio of 0.15. Further study on eigenvalues of the LD1-axis reveals that the effect of cycling temperature mainly reflects on a specific voltage range from 3.4V to 3.6V in the dQ/dV-V curve. On the other hand, although cells cycled under different SOC ranges cannot be well identified, the eigenvalue information shows most data variance is caused by the cycling temperature factor. In other words, the cycling temperature is more dominant than the working SOC range in the shifting behavior of dQ/dV-V curve.

References

- Xiaopeng Tang, Yujie Wang, Changfu Zou, Ke Yao, Yongxiao Xia, Furong Gao. A novel framework for Lithium-ion battery modeling considering uncertainties of temperature and aging. *Energy Convers.*, 2019; 180:162-170.
- Anthony Barré, Benjamin Deguilhem, Sébastien Grolleau, Mathias Gérard, Frédéric Suard, Delphine Riu. A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, 2013; 241:680-689.
- Christoph R. Birkl, Matthew R. Roberts, Euan McTurk, Peter G. Bruce, David A. Howey. Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, 2017; 341:373-386.
- Feng Leng, Cher Ming Tan, Michael Pecht. Effect of temperature on the aging rate of Li ion battery operating above room temperature. *Sci. Rep.*, 5, 12967; doi: 10.1038/srep12967 (2015).
- Naixing Yang, Xiongwen Zhang, BinBin Shang, Guojun Li. Unbalanced discharging and aging due to temperature differences among the cells in a lithium-ion battery pack with parallel combination. *Journal of Power Sources*, 2016; 306:733-741.
- Christoph R. Birkl, Matthew R. Roberts, Euan McTurk, Peter G. Bruce, David A. Howey. Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, 2017; 341:373-386.
- Issam Baghdadi, Olivier Briat, Jean-Yves Deléage, Philippe Gyan, Jean-Michel Vinassa. Lithium battery aging model based on Dakin's degradation approach. *Journal of Power Sources*, 2016; 325:273-285.
- Yang Gao, Jiuchun Jiang, Caiping Zhang, Weige Zhang, Yan Jiang. Aging mechanisms under different state-of-charge ranges and the multi-indicators system of state-of-health for lithium-ion battery with Li(NiMnCo)O₂ cathode. *Journal of Power Sources*, 2018; 400:641-651.
- Matthiue Dubarry, Cyril Truchot, Boryann Liaw. Mechanistic diagnosis and prognosis on Li-ion battery degradation. Presented at: 2013 IBA.
- Linjing Zhang, Jiuchun Jiang, Weige Zhang. Capacity Decay Mechanism of the LCO + NMC532/Graphite Cells Combined with Post-Mortem Technique. *Energies*. 10(8), 1147; doi: 10.3390/en10081147 (2017).

11. Tiphaine Plattard, Nathalie Barnel, Loïc Assaud, Sylvain Franger, Jean-Marc Duffault. Combining a Fatigue Model and an Incremental Capacity Analysis on a Commercial NMC/Graphite Cell under Constant Current Cycling with and without Calendar Aging. *Batteries.*, 5(1), 36; doi: 10.3390/batteries5010036 (2019).
12. Adnan Nuhic, Tarik Terzimehic, Thomas Soczka-Guth, Michael Buchholz, Klaus Dietmayer. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *Journal of Power Sources*, 2013; 239:680-688.
13. Xiaosong Hu, Shengbo Eben Li, Yalian Yang. Advanced Machine Learning Approach for Lithium-Ion Battery State Estimation in Electric Vehicles. *IEEE Transactions on Transportation Electrification*, 2016; 2:140-149.