Prabha Materials Science Letters
Vol. 5, No. 2, 323-337, 2026
https://doi.org/10.33889/PMSL.2026.5.2.018

Complex Network Analysis of Lithium-Ion Battery Life

Manoj K. Singh
School of Computer Science Engineering and Technology,
Bennett University, Greater Noida, Uttar Pradesh, India.
E-mail: manoj.kumar4@bennett.edu.in

Sangeeta Pant
Symbiosis Institute of Technology, Pune,
Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India.
Corresponding author: pant.sangeet@gmail.com

Shshank Chaube
Symbiosis Institute of Technology, Hyderabad,
Symbiosis International (Deemed University), Pune, Maharashtra, India.
E-mail: chaube.shshank@gmail.com

Harsh
Department of Computer Science & Technology,
Manav Rachna University, Faridabad, 121010, Haryana, India.
E-mail: harshyaduvanshihy@gmail.com

Anuj Kumar
School of Computer Science Engineering & Applications,
D. Y. Patil International University (DYPIU), Pune, Maharashtra, India.
E-mail: anuj4march@gmail.com

Jitendra Pal Singh
Department of Basic Sciences (Physics),
Manav Rachna University, 121010, Faridabad, Haryana, India.
E-mail: jitendra86969920@gmail.com

(Received on December 31, 2025; Revised on April 30, 2026 & June 8, 2026 & June 21, 2026; Accepted on June 23, 2026)

Abstract
Machine learning is among the most popular techniques that use existing data to learn the mechanism that drives a particular phenomenon, in particular, state-of-health (SOH). Batteries are charged at a certain ambient temperature; however, depending on the battery degradation mechanism, the temperature of the battery could vary. In addition, voltage may also not be constant during a charging process. The features extracted for voltage and temperature in a charging cycle can then be useful in predicting the SOH of a battery in conjunction with a machine-learning model. We used complex networks to form a vector of features that can be useful in predicting SOH with these features as input to long short-term memory neural networks. The comparative analysis between measured SOH and predicted SOH shows that the features derived from complex networks used in this study can be useful in predicting SOH.

Keywords- Visibility graph, LSTM, State-of-health, Battery degradation.

1. Introduction

Despite the higher energy densities of the Li-ion rechargeable batteries, the high cost and length of their life cycle have limited the expected applicability of these batteries in vehicles on a large scale. The capacity of the batteries decreases with use, and after a number of charging/discharging cycles, the battery degrades to a certain level, after which it becomes of no use (Lin et al., 2021; Shahjalal et al., 2022). This behaviour of battery is often referred to as State-of-Health (SOH), which is defined as $q(t)/q(0)$, where $q(t)$ is the capacity at time t. This means that $q(0)$ is the capacity when $t = 0$, that is, at the beginning of the charging process. SOH has been subject of numerous theoretical studies recently depicting its importance in rechargeable batteries (Wang et al., 2024; Yang et al., 2023) and is related to remaining useful life (RUL) of a battery (Zhao and Burke, 2023).

Machine learning (ML) is one of the most effective techniques for detecting and predicting SOH and RUL of batteries (Lipu et al., 2022; Priyanka et al., 2022, Rauf et al., 2022) by selecting the appropriate input feature and followed by an effective algorithm model (**Table 1**). Zuo et al. (2023) studied the factors associated with the failure of the batteries by summarizing them into three features. The study by Jiang et al. (2023), combined a self-attention mechanism with a convolutional autoencoder to find automatic features. These self-attention mechanisms are further used for mapping the features with the SOH. Pan et al. (2023) used a two-stage feature extraction method based on transformations. The features are used in an ensemble learning algorithm using a gradient-boosting decision tree to estimate the SOH. Additionally, Guo et al. (2019) used principal component analysis and rational analysis for extracting and optimizing features from the charging stages.

Table 1. Input data features for different models. Here following convention is used: *I*- Current, *V*- Voltage, SOC-State of charge, Cap- Capacity, *T*-temperature, PR- Power, *R* – Resistance, *R*/*C*- Resistance/Capacitance at time *t*. Adapted from Jin et al. (2021).

Model	**Input features**							
	$I(t)$	$V(t)$	SOC(t)	Cap(t)	$T(t)$	PR(t)	$R(t)$	$R/C(t)$
Support vector regression (SVM)	✓	✓	✓	✓	✓	×	×	×
Gaussian process regression (GPR)	✓	✓	✓	×	×	×	×	×
Deep Neural Network (ELM)	✓	✓	✓	✓	✓	✓	✓	×
Deep Neural Network (DNN)	✓	✓	×	✓	✓	×	×	×
Recurrent Neural Network (RNN)	✓	✓	✓	✓	×	×	×	✓
Long Short-Term Memory (LSTM)	✓	✓	×	✓	×	×	×	×

For predicting the SOH, a set of good quality features are required. The features extracted using complex neural networks are hard to interpret. Therefore, a feature set which explains the physics and chemistry of the aging process of the battery is required. One such set of features can be derived using visibility graphs and associated characteristics. Complex networks are a popular learning algorithm for data analysis, with a wide range of applications in fields such as time series classification (Silva et al., 2022; Singh et al., 2023). In this article, the features are extracted from the charging/discharging cycles, and they are characterized to describe the aging process, which help in better understanding of the aging process and predicting the health of a battery. By deriving multiple features from the graph formed by features from charging/discharging cycles, can be used as supporting input vectors in ML models for predicting the SOH of a battery. Models, such as Long Short-Term Memory (LSTM), are often used for this purpose, taking sequential data as input (Gong et al., 2022; Jorge et al., 2023). Thus, this work is motivated to use complex network for extracting features and to apply them as inputs to train LSTM model. The novelty lies in the use of visibility graphs, the combination of complex networks features, the physical interpretation of graph-based features related to the battery life.

2. Dataset

Oxford Capacity Degradation Dataset 1 containing measurements from battery data of eight Li-ion cells (Birkl, 2017; Birkl et al., 2017). This dataset consists of lithium-ion cells which uses cathode based on lithium and anode based on graphite. The batteries undergo multiple cycles of charge and discharge so that degradation is accelerated under similar conditions. They are charged and constant current by keeping an upper bound of roughly 4.2 V on voltage. After that a battery is discharged and a specified rate by keeping a lower bound on the voltage between 2.7-3V. The data were recorded after every 100 cycles. The environment temperature is set at 40^∘ C. The capacity can be computed by integrating the discharge current over a time interval. Using this capacity $q(t)$, in this article, the normalized capacity $q(t)/q(0)$ is used for the analysis. Early to mid-cycles were selected for the training and late cycles were used for the prediction.

The capacity of the batteries decreases smoothly with each cycle for Oxford cells (**Figure 1a**).

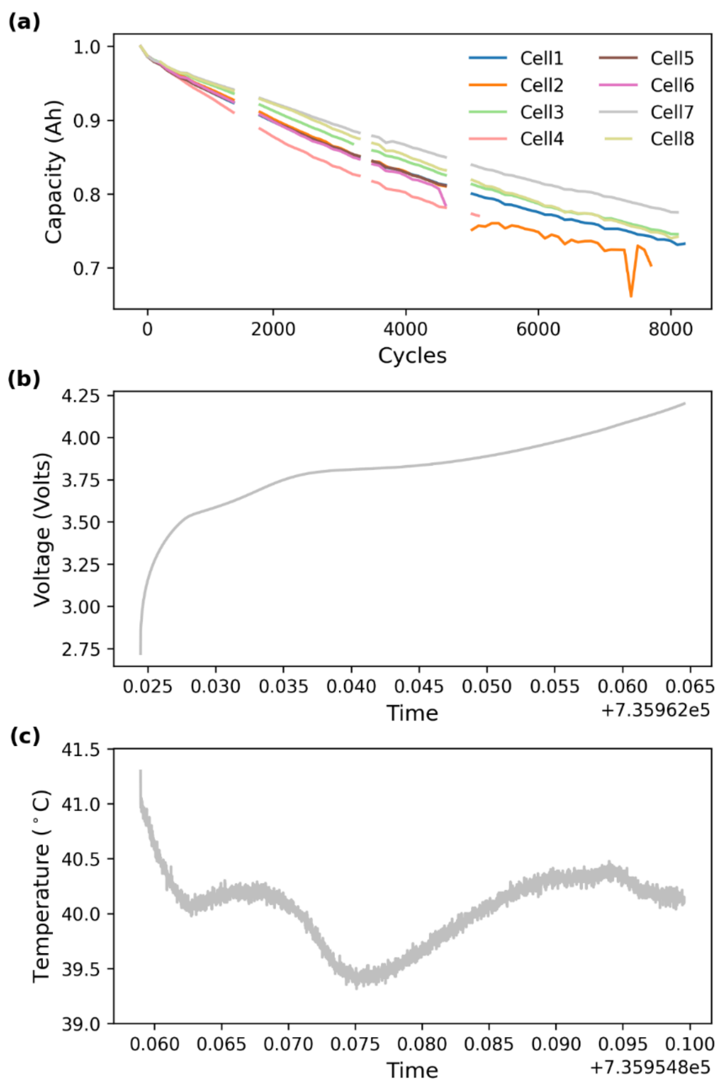

Figure 1. (a) Capacity of oxford Li-ion batteries (b) Voltage profile for 100th cycle of the seventh cell (c) Temperature profile for 100th cycle of the seventh cell.

A variation of voltage for Cell#7 during the 100th cycle over time shows an increasing pattern (**Figure 1b**) and temperature profile of the 100th cycle of the seventh cell (**Figure 1c**) shows a cyclic decreasing and increasing pattern, with an overall decreasing pattern. The dataset consists of 8 cells. Each cell contains time, voltage, charge and temperature profile at each 100 cycles until battery degrades significantly. The cells contain a total of 83, 78, 82, 52, 49, 51, 82, and 82 cycles, respectively. The dataset stores the information in nested order of cells, cycle and then voltage/temperature signals. A loop through cells and cycles are re-quired during pre-processing to extract the signals. Then SOH was extracted as ratio of capacity at time t with the initial capacity.

3. Methodology

In this work. the feature set is formed by extracting graph properties from the complex networks that are created using the voltage and temperature time series (**Figure 2**).

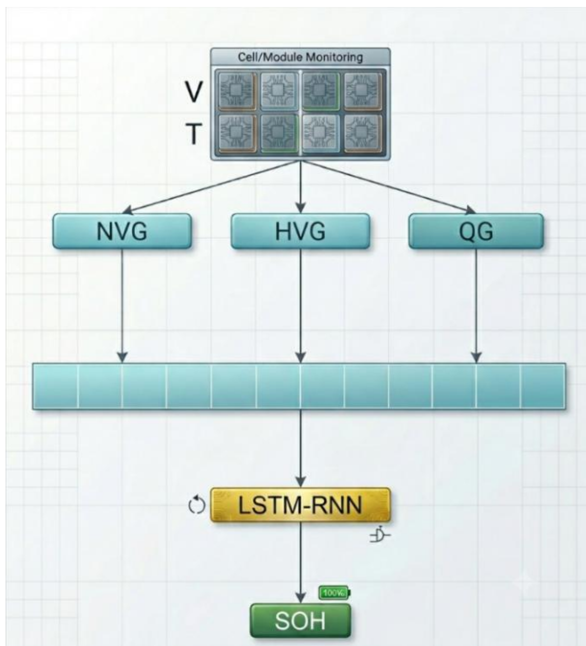

Figure 2. Flowchart for using LSTM model for estimation of battery health. Here NVG represents natural visibility graph, HVG stands for the horizontal visibility graph and QG is the quantile graph, while SOH stands for the State-of-Health of the battery.

3.1 Complex Networks

Here, three forms of complex network graph are discussed; the two forms are the natural visibility graph and the horizontal visibility graph, while the third form is the quantile graph. Visibility graphs are widely used in time series analysis, classification, and forecasting (Lacasa et al., 2008; Luque et al., 2009); visibility graphs have applications in several fields, including image processing (Ali et al., 2023). In a natural visibility graph, each node represents a time epoch of the series. Two nodes are connected through an edge only if the corresponding values of the series are visible to each other naturally (**Figure 3a**) i.e., two nodes at epoch, i and j are connected if they satisfy following equation.

$$y_k < y_j + (y_i - y_j)\frac{x_j - x_k}{x_j - x_i} \tag{1}$$

for all x_k between x_i and x_j, here y_i is the time series value at time epoch x_i.

Similarly, in a horizontal visibility graph (**Figure 3b**), two nodes are connected through an edge only if corresponding values are visible horizontally i.e, two nodes at epoch i and j are connected if they satisfy following equation.

$$y_k < \min(x_i, x_j) \tag{2}$$

for all x_k between x_i and x_j, here y_i is the time series value at time epoch x_i.

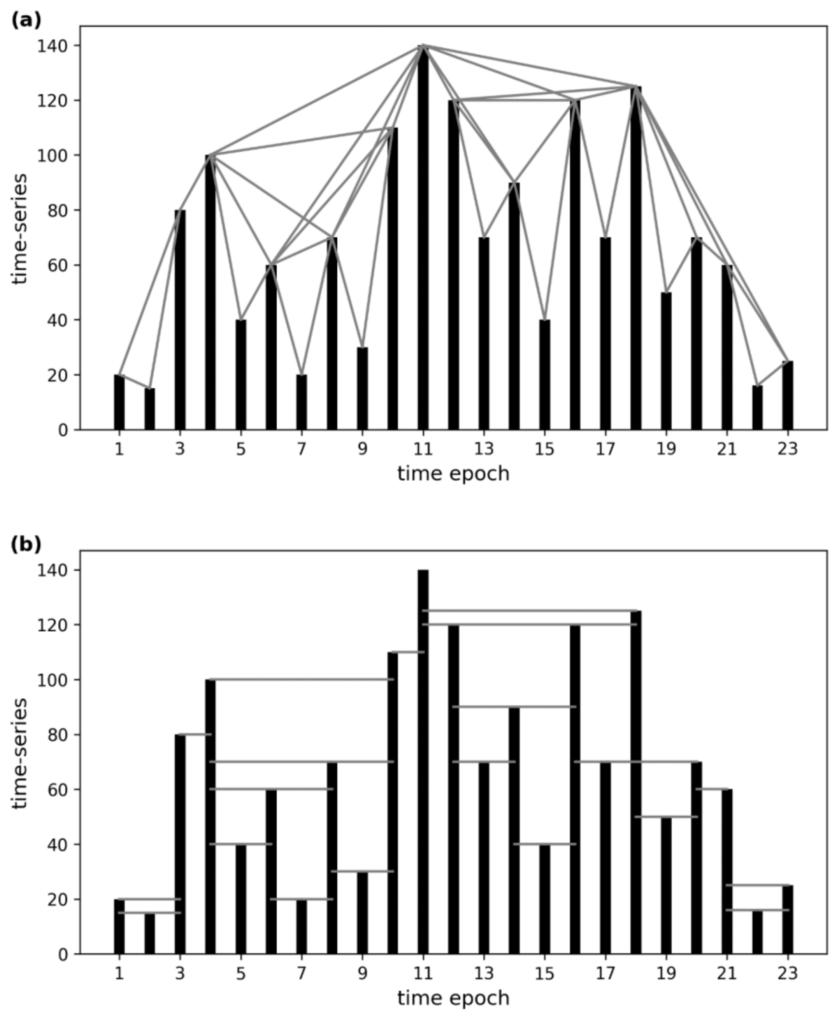

Figure 3. (a) An example of natural visibility graph with example time series of length 23. The resultant graph has 23 nodes, and two nodes are connected if they are visible to each other (b) An example of horizontal visibility graph constructed from the same data as in (a), however, two nodes are connected if they are visible horizontally.

To create a quantile graph from a time series, the values of the series are divided into n bins based on the quantile sequence of length n. Each bin is associated with a node of a graph, so the number of nodes in the graph is n. A size $n \times n$ transition matrix, $Q = [q_{ij}]$, is formed based on the number of transitions from one bin to another i.e., $q_{ij=}$ the number of transitions from ith bin to jth bin. Each element in the transition matrix is normalized by dividing the sum of elements in a row so that each row sums to one i.e., if $P = [p_{ij}]$

is the normalized transition matrix, then, $p_{ij} = \frac{q_{ij}}{\sum_{j=1}^{n} q_{ij}}$. The values in this matrix are used as edge weights to form a weighted directed graph. The selection of the number of bins is an important part in forming the graph, which can be learned as hyperparameter. However, we performed sensitivity analysis by varying bin size from 50-1000 as shown in **Figure 4**. For bin size 400 error is smallest, so we selected 400 as the bin size for forming the quantile graph.

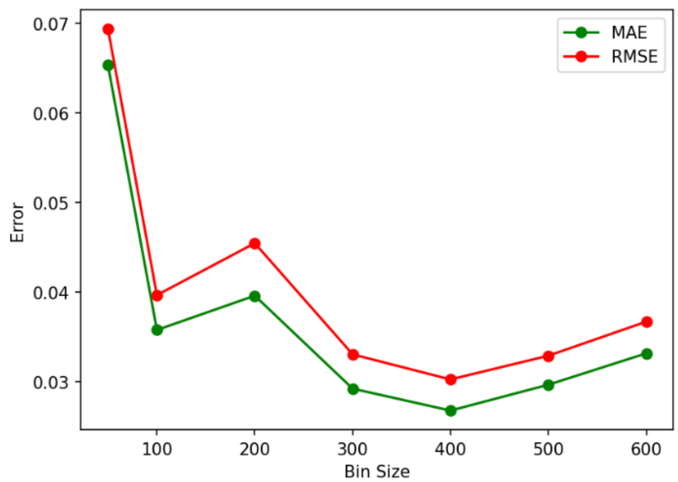

Figure 4. MAE and RMSE of prediction using varying bin size in quantile graph.

Total of 15 features were used for the prediction task, 5 features each from the three types of complex networks, natural visibility graph, horizontal visibility graph and quantile graph. The features are weighted average degree of the nodes, average path length, number of communities, average clustering, and modularity. As quality of the battery decreases the resistance in the battery could change which will lead to change in slope of the voltage and temperature change. At early stages the voltage curve has slight low value mid-way and visibility at distant points are not restricted thus high weights are given to node degrees, which means that average degree of the visibility graph is high. As battery degrades, the change in voltage gives rise to obstruction to distant nodes and hence reduction in the weighted average of the nodes. Therefore, weighted average of the nodes is a good feature to have for the prediction of health of the battery. Voltage curve is smooth and hence have many distant nodes are visible leading to the shorter path length and degradation in battery leads to a sharp bend mid-way of the curve and hence larger average path length. However, after a threshold degradation the average path length does not increase, therefore, distinction between extremely degraded and average degraded battery is not possible. In contrast the temperature curve is not smoother compared to the voltage curve, upon heating with aging the curve tends to become smooth which leads to small average path length, which helps in characterizing the aging of the battery. Mid-way during charging there is change in the voltage pattern with aging. With change in resistance the curvature of the voltage curve is more, and curve has more bends at early stages, this leads to a greater number of communities in newer batteries, with aging the number of communities decreases. The average clustering measures the likelihood of neighbors of node to be connected. The voltage curve of the relatively newer battery has less curvature and nodes are likely to connected with many nearby nodes and hence has high clustering. However, the aged voltage has a high curvature midway leaning to lower average clustering. During charging electrochemical reactions takes place and the voltage curve goes through various phases. The modularity of the graph formed measures the distinct phases within the cycle. During aging the voltage curve shows increase in the modularity. Therefore, modularity as a feature of the charging cycle has the potential for charactering the aging process. The five types of features of natural visibility graph for the voltage and temperature are shown in **Figure 5**. From each graph five types of features are computed. Therefore, NVG, HVG and QG provide 5 features each, a total of 15 from a single time series. There are two time series, voltage and temperature, leading to 15+15=30 features, by combining features from voltage and temperature time series.

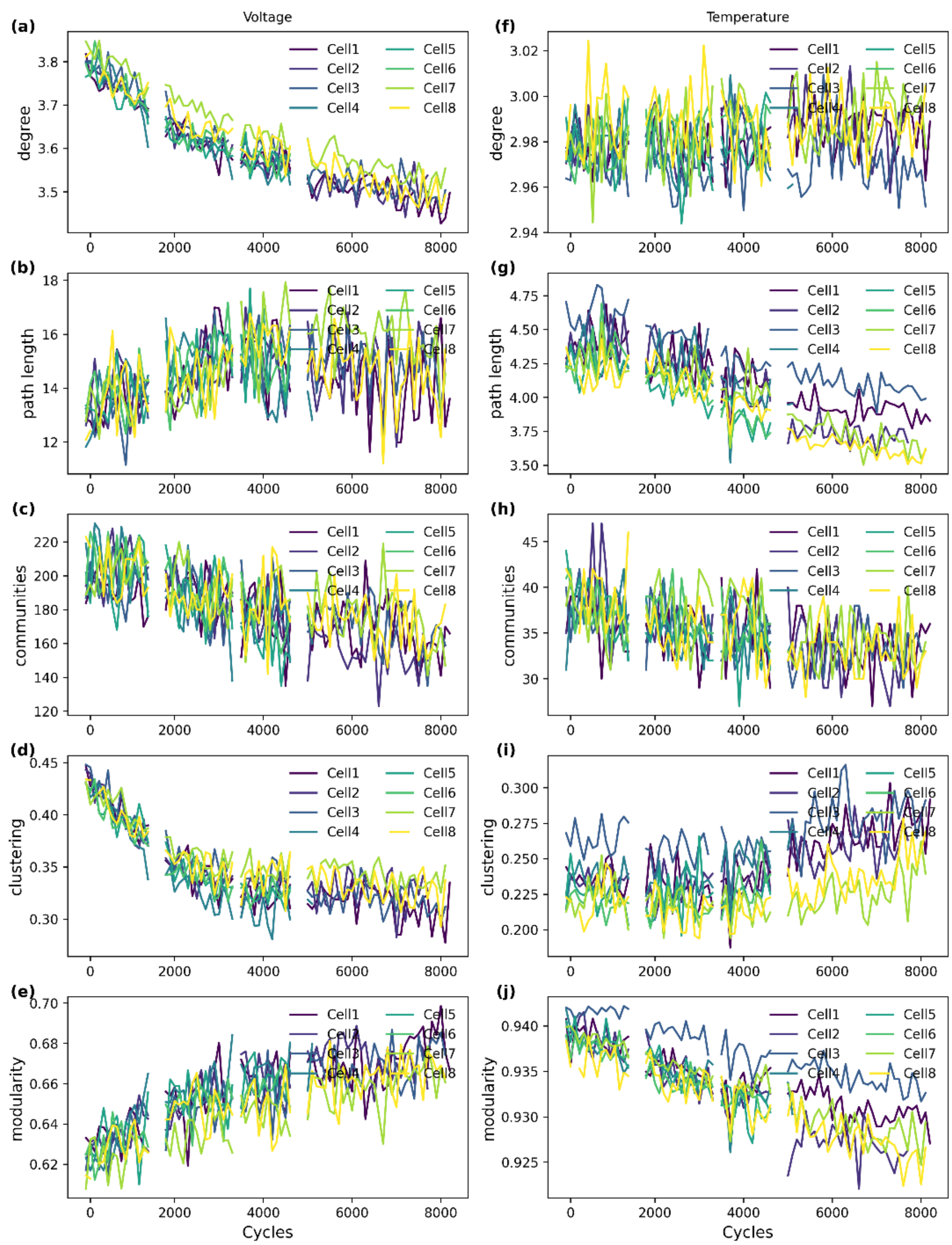

Figure 5. The five types of features extracted using natural visibility graph for voltage (a-e) and for temperature (f-j).

Algorithm 1, presented below explains the pseudocode for the extraction of features from the graph.

Algorithm 1: Formation of graph and extraction of features
1) Input: a time series of charging/discharging cycle.
2) Form an empty feature vector of length 30.
3) Form a graph from the charging/discharging cycle for voltage.
4) Compute graph features from the voltage graph using Algorithm 2.
5) Compute features from graph in step 4 and add into feature vector.
6) Compute graph features from the voltage graph using Algorithm 3.
7) Compute features from graph in step 6 and add into feature vector.
8) Compute graph features from the voltage graph using Algorithm 4.
9) Compute features from graph in step 8 and add into feature vector.
10) Repeat step 3-9 for the charging/discharging cycle for temperature.

Algorithm 2: Formation of Natural Visibility Graph
1) Input: a time series of charging/discharging cycle.
2) Form a bar chart with x-axis as time epochs and y-values as the values of the time series.
3) Create nodes of a graph one, for each time epoch.
4) Connect two nodes if bars of the corresponding nodes can be connected with a straight line without intersecting the other bars.

Algorithm 3: Formation of Horizontal Visibility Graph
1) Input: a time series of charging/discharging cycle.
2) Form a bar chart with x-axis as time epochs and y-values as the values of the time series.
3) Create nodes of a graph one, for each time epoch.
4) Connect two nodes if bars of the corresponding nodes can be connected with a horizontal straight line without intersecting the other bars.

Algorithm 4: Formation of Quantile Graph
1) Input: a time series of charging/discharging cycle.
2) Divide the data into 400 quantile levels corresponding to a percentile range.
3) Form a transition matrix A=$[a_{ij}]$ of size 400×400.
4) a_{ij} is equal to number of times data moves from ith quantile to jth quantile.
5) Transform the transition matrix into probability by dividing each element in a row by sum of the elements in the row.
6) Form a graph from this matrix with 400 nodes.

In a controlled manner, Oxford datasets embeds the characteristics of the current and the resistance into SOC (*t*), voltage and temperature. In the datasets, the current is prescribed and there is little variation in it. The measuring resistance is not straightforward and but can be modeled as voltage already contains the information carried by the resistance. Oxford dataset is created in a controlled way which makes it ideal for studying the aging mechanism. The features extracted using the complex networks are physics aware as discussed in the previous paragraph. Which depends on the shape of the curve which remains similar across charging cycle and variability is primarily due to the aging process. There for the features discussed here are able to describe the aging process and hence are generalizable. The creation of features from the natural visibility graphs is more complex than the horizontal visibility graph, for both the time and space. However, features can be extracted from the neural network as well, which could be more complex than the graph techniques.

The LSTM model consists of 2 layers, output dimension is set as 1 and the number of neurons in the hidden layer is 256. The number of epochs is 100, batch size is 16 and the optimizer is Adam. A dropout with probability 0.2 is also used. The learning rate is 0.001, the loss function used is mean squared error and the input sequence length was taken as 5. The dataset is normalized by scaling between 0 and 1. For validation forward walk method is used in this study and since this approach dominates the randomness no seed was provided to the model. For stopping a fixed epoch of 100 runs are selected. All the hyperparameter are set in advance and hence no tunning was required.

4. Results and Discussions

For this work, three types of time series during a charging process are used: temperature time series, voltage time series, and their combined time series. The combined time series is a two-dimensional time series, with the first dimension as temperature and the second dimension as voltage. To get more confidence in prediction, SOH was predicted for 10 cycles in advance with time, Initially, only the first 40% of the data is used for predicting the next 10-time epochs, as in the Oxford dataset, first 30% cell behavior are quasi-linear, and after that aging accelerates, then as we get the data, we advance by that small amount of time and predict the SOH for another 10 epochs of data. By using 40% data as initial cut off we train the model to predict aging not the battery life. The initial 40% data from each of the cells are used for the training the model separately, leading an individual parameter set for each of the cell. The trained model is then used for predicting time series of each cell individually. The degradation process in battery is time dependent and splitting the data for training and testing purely based on the cycles, e.g., early cycles were used for training but not included in the testing. Since sliding window by using cycle by cycle advancing approach were used for training and since only future cycles were used for the prediction, future information is thereby not used for learning the model. Further all the features are computed within each cycle and no future cycles were included in the feature engineering to avoid data leakage.

Figure 6 shows the actual and predicted SOH for all cells. Several predictions are made during a cycle by advancing the input dataset by one cycle. Therefore, we take an average of all the predictions to show a single value for a cycle.

Figure 6a reflects that the averaged predicted SOH follows closely the trails of measured SOH. During initial cycles, there is no prediction; only after cycle 17 prediction starts. The accuracy of the prediction at initial cycles is large but during the end of the cycles as the temperature curve starts moving with higher periods the accuracy decreases but remains low. In later cycles the battery degradation process enters a new regime, which is different from the previous cycles. Therefore, the model which trains previous cycle has found difficulty in following the SOH curve in later cycles. This is a fundamental learning problem. To measure the accuracy of prediction, we used statistics including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Small values in MAE and RMSE indicate a good prediction model. Here by prediction model, we mean the model along with the input features.

Table 2 displays the MAE and RMSE of the predictions made using the features of the complex network in conjunction with the LSTM model. The average MAE ranges from 0.0067 to 0.0156 and the RMSE ranges from 0.0082 to 0.0210. The MAE values, as well as the RMSE values, are small across all the Cells, indicating that the used features in conjunction with the machine-learning model have the potential to be used effectively in predicting the battery life after subjected to use. Estimated errors in MAE and RMSE are analogous to that obtained for other models (Ren et al., 2018; Zou et al., 2023; Zhu et al., 2022) and is shown in **Figure 7**. The cells show varying degrees of variations in the degradation patterns which significantly effects the predictions in the battery's life cycle. The irregularities in the prediction results across cells may have originated from the internal resistance and changes in aging mechanism. Further, the aging process in highly nonlinear, in initial phase of the aging mechanism the battery stablishes, then there is a linear degradation phase and then the degradation accelerates towards the end. The model trained on the initial phase adopts toward the phase. The window shift method used in the LSTM helps the model in learning these phases. At high temperature the battery degrades faster and hence inclusion of temperature features have improved the predictability of the model. Different cells have different manufacturing tolerances and structure of the electrodes the voltage and temperature features cannot include and hence the prediction variability in cells can be noted.

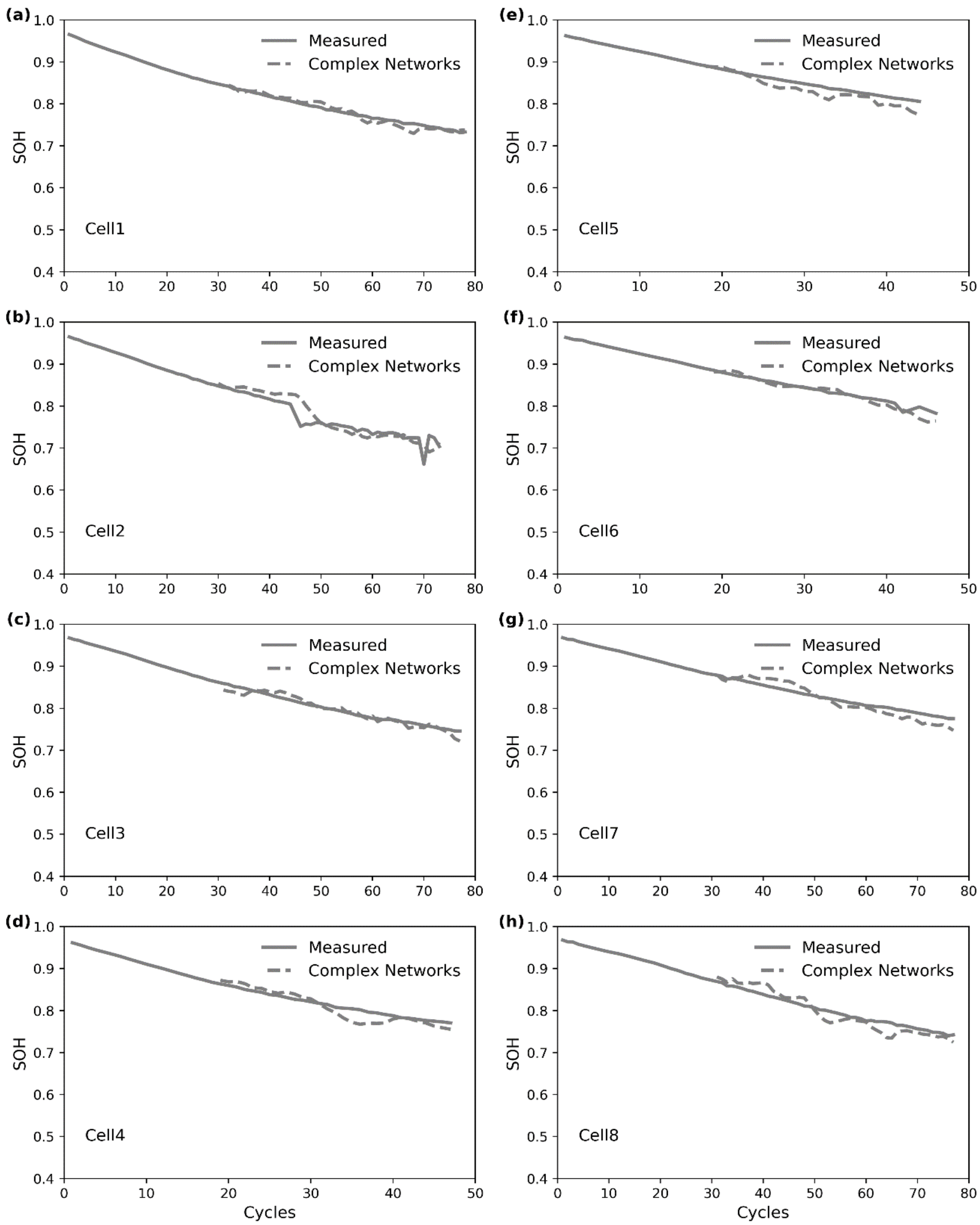

Figure 6. Measured and Estimated State-Of-Health of (a) Cell#1 (b) Cell#2, (c) Cell#3, (d) Cell#4, (e) Cell#5, (f) Cell#6, (g) Cell#7 and (h) Cell#8. Here, in the figure label by Complex Networks we mean the predicted SOH using the complex networks features.

Table 2. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0067	0.8609	0.0082	1.0711
Cell#2	0.0156	2.0522	0.0210	2.8039
Cell#3	0.0072	0.8980	0.0092	1.1501
Cell#4	0.0125	1.5412	0.0151	1.8819
Cell#5	0.0138	1.6515	0.0162	1.9460
Cell#6	0.0073	0.8989	0.0104	1.2960
Cell#7	0.0131	1.6118	0.0148	1.8285
Cell#8	0.0145	1.8114	0.0166	2.0832

Table 3 displays the MAE and RMSE of the predictions made using the features of the complex network extracted using only the horizontal visibility graph, in conjunction with the LSTM model. The average MAE ranges from 0.0173 to 0.0316 and the RMSE ranges from 0.0211 to 0.0350.

Table 3. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model with only HVG features.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0179	2.3295	0.0211	2.7763
Cell#2	0.0217	2.9007	0.0259	3.4935
Cell#3	0.0173	2.1793	0.0215	2.7185
Cell#4	0.0209	2.5984	0.0240	2.9986
Cell#5	0.0316	3.7821	0.0350	4.1995
Cell#6	0.0218	2.6527	0.0260	3.2208
Cell#7	0.0187	2.3197	0.0237	2.9711
Cell#8	0.0202	2.5870	0.0251	3.2236

Table 4 displays the MAE and RMSE of the predictions made using the features of the complex network extracted using only the natural visibility graph, in conjunction with the LSTM model. The average MAE ranges from 0.0094 to 0.0155 and the RMSE ranges from 0.0084 to 0.0207.

Table 4. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model with only NVG features.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0120	1.5303	0.0142	1.8079
Cell#2	0.0155	2.0334	0.0207	2.7550
Cell#3	0.0128	1.6141	0.0159	2.0091
Cell#4	0.0148	1.8270	0.0187	2.3114
Cell#5	0.0073	0.8595	0.0084	0.9964
Cell#6	0.0094	1.1264	0.0117	1.3951
Cell#7	0.0110	1.3023	0.0151	1.7784
Cell#8	0.0154	1.9011	0.0189	2.2982

Table 5 displays the MAE and RMSE of the predictions made using the features of the complex network extracted using only the quantile graph, in conjunction with the LSTM model. The average MAE ranges from 0.0149 to 0.0383 and the RMSE ranges from 0.0174 to 0.0463. From **Table 2-5**, it can be concluded that individual features in particular that of quantile graph have poor accuracy but their combination has better accuracy.

Table 5. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model with only QG features.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0201	2.6117	0.0245	3.1991
Cell#2	0.0412	5.4050	0.0463	6.1197
Cell#3	0.0288	3.6217	0.0331	4.1748
Cell#4	0.0305	3.7464	0.0345	4.2291
Cell#5	0.0161	1.9040	0.0185	2.1817
Cell#6	0.0383	4.6397	0.0407	4.9857
Cell#7	0.0243	2.9309	0.0270	3.2497
Cell#8	0.0149	1.8444	0.0174	2.1364

Table 6 displays the MAE and RMSE of the predictions made using the features of the complex network extracted using only the voltage features, in conjunction with the LSTM model. The average MAE ranges from 0.0110 to 0.0253 and the RMSE ranges from 0.0135 to 0.0271.

Table 6. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model with only voltage.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0131	1.6840	0.0143	1.8426
Cell#2	0.0133	1.7662	0.0169	2.2833
Cell#3	0.0110	1.4006	0.0135	1.7359
Cell#4	0.0237	2.9579	0.0257	3.2294
Cell#5	0.0253	2.9992	0.0271	3.2066
Cell#6	0.0148	1.8183	0.0198	2.4690
Cell#7	0.0130	1.6176	0.0164	2.0718
Cell#8	0.0128	1.6268	0.0151	1.9368

Table 7 displays the MAE and RMSE of the predictions made using the features of the complex network extracted using only the voltage features, in conjunction with the LSTM model. The average MAE ranges from 0.0105 to 0.0309. and the RMSE ranges from 0.0124 to 0.0338. **Table 6**-**7** shows that the accuracy of individual components of voltage and temperature have inferior performance compared to their combination. However, the voltage has far more contribution towards the accuracy of overall feature space for the prediction.

Table 7. Comparison between measured SOH and estimated SOH using MAE and RMSE based on complex networks features and LSTM model with only temperature.

Cell	MAE	MAE%	RMSE	RMSE%
Cell#1	0.0213	2.7228	0.0239	3.0623
Cell#2	0.0202	2.6573	0.0294	3.8823
Cell#3	0.0142	1.7905	0.0160	2.0186
Cell#4	0.0309	3.7860	0.0338	4.1442
Cell#5	0.0105	1.2417	0.0129	1.5136
Cell#6	0.0110	1.3261	0.0124	1.5055
Cell#7	0.0167	2.0278	0.0191	2.323
Cell#8	0.0197	2.4306	0.0234	2.8683

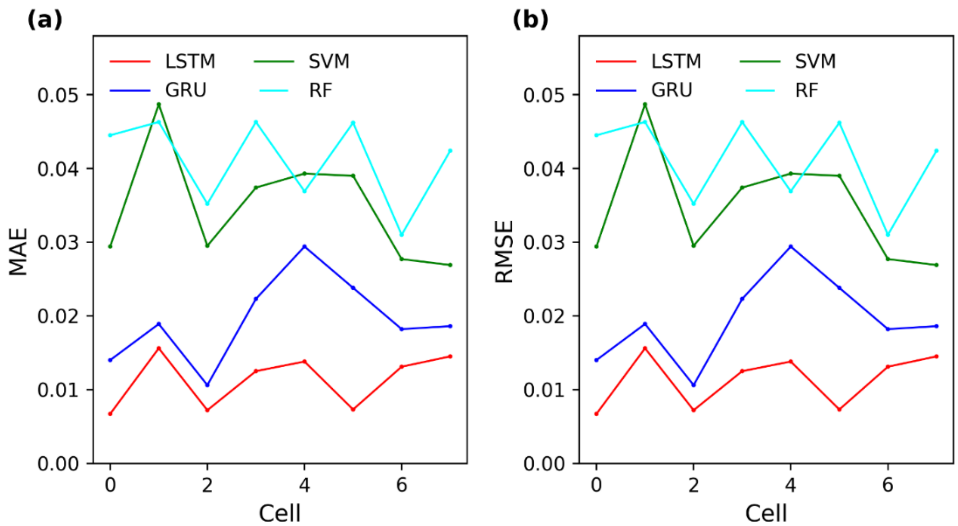

Figure 7. Comparison of the LSTM model with Gated Recurrent Units (GRU), Support Vector Machine (SVM) and Random Forest (RF) on the same data.

The proposed features were also analyzed using other models including Gated Recurrent Units (GRU), Support Vector Machine (SVM) and Random Forest (RF) on the same data on all the cells from 1-8. The **Figure 7** shows their comparative analysis based on MAE and the RMSE. Both the error metrics show that the error in the LSTM is smallest compared to GRU, SVM and RF. Similar kind of studies are also presented in various studies. The study in presented the comparison of various models using MAE and RMSE (Kadem and Kadem, 2026). Their MAE was 0.78-0.96 and RMSE was 1.03-1.07. The RMSE and MAE reported by Sun et al. (2025a) are 1.62 and 1.28, respectively. The RMSE and MAE are 1.85 and 1.41, respectively as reported by Sun et al. (2025b). RMSE is observed 1.4 by Zhang et al. (2025).

5. Conclusions

Here, machine-learning approach is successfully used to identify the characteristics that signal the degradation stages of a battery. This helps to explore the State-of-Health and Remaining Useful Life of a battery. An approach based on feature extraction from complex networks achieves high accuracy in battery health prediction, as indicated by low mean absolute error (MAE) and root mean squared error (RMSE). Therefore, these features can be useful in identifying the degradation process in battery life cycles. The voltage features are dominating compared to temperature features and visibility graphs presents themselves as a better graph network for representing the battery aging cycle compared to the quantile graphs. This study can be extended to remove several of limitations including the small number of cells, validation on a single dataset, and the need for testing across different battery chemistries and operating conditions.

Conflicts of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgments

Authors are thankful to Christoph R. Birkl for creating and maintaining the Oxford battery degradation dataset 1. JPS acknowledges the financial support received from Anusnadhan National Research Foundation (ANRF), New Delhi vis project RJF2021/000115.

AI Disclosure

The author(s) declare that no assistance is taken from generative AI to write this article.

References

Ali, M., Kumar, S., Pal, R., Singh, M.K., & Saini, D. (2023). Graph-and machine-learning-based texture classification. *Electronics*, *12*(22), 4626.

Birkl, C. (2017). Oxford battery degradation dataset 1. https://ora.ox.ac.uk/objects/uuid:03ba4b01-cfed-46d3-9b1a-7d4a7bdf6fac

Birkl, C.R., Roberts, M.R., McTurk, E., Bruce, P.G., & Howey, D.A. (2017). Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, *341*, 373-386.

Gong, Y., Zhang, X., Gao, D., Li, H., Yan, L., Peng, J., & Huang, Z. (2022). State-of-health estimation of lithium-ion batteries based on improved long short-term memory algorithm. *Journal of Energy Storage*, *53*, 105046.

Guo, P., Cheng, Z., & Yang, L. (2019). A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *Journal of Power Sources*, *412*, 442-450.

Jiang, Y., Chen, Y., Yang, F., & Peng, W. (2023). State of health estimation of lithium-ion battery with automatic feature extraction and self-attention learning mechanism. *Journal of Power Sources*, *556*, 232466.

Jin, S., Sui, X., Huang, X., Wang, S., Teodorescu, R., & Stroe, D.I. (2021). Overview of machine learning methods for lithium-ion battery remaining useful lifetime prediction. *Electronics*, *10*(24), 3126.

Jorge, I., Mesbahi, T., Samet, A., & Boné, R. (2023). Time series feature extraction for lithium-ion batteries state-of-health prediction. *Journal of Energy Storage*, *59*, 106436.

Kadem, H.C., & Kadem, O. (2026). Lithium-ion battery state of health estimation using automated machine learning (AutoML): performance comparison. *PeerJ Computer Science*, *12*, e3497. https://doi.org/10.7717/peerj-cs.3497

Lacasa, L., Luque, B., Ballesteros, F., Luque, J., & Nuno, J.C. (2008). From time series to complex networks: the visibility graph. *Proceedings of the National Academy of Sciences*, *105*(13), 4972-4975.

Lin, X., Khosravinia, K., Hu, X., Li, J., & Lu, W. (2021). Lithium plating mechanism, detection, and mitigation in lithium-ion batteries. *Progress in Energy and Combustion Science*, *87*, 100953.

Lipu, M.H., Ansari, S., Miah, M.S., Meraj, S.T., Hasan, K., Shihavuddin, A.S.M., Hannan, M.A., Muttaqi, K.M., & Hussain, A. (2022). Deep learning enabled state of charge, state of health and remaining useful life estimation for smart battery management system: methods, implementations, issues and prospects. *Journal of Energy Storage*, *55*, 105752.

Luque, B., Lacasa, L., Ballesteros, F., & Luque, J. (2009). Horizontal visibility graphs: Exact results for random time series. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, *80*(4), 046103.

Pan, R., Liu, T., Huang, W., Wang, Y., Yang, D., & Chen, J. (2023). State of health estimation for lithium-ion batteries based on two-stage features extraction and gradient boosting decision tree. *Energy*, *285*, 129460.

Priyanka, T., Singh, M., & Kumar, A. (2022). Deep learning for satellite-based data analysis. *Meta-heuristic Optimization Techniques: Applications in Engineering*, *10*, 173.

Rauf, H., Khalid, M., & Arshad, N. (2022). Machine learning in state of health and remaining useful life estimation: theoretical and technological development in battery degradation modelling. *Renewable and Sustainable Energy Reviews*, *156*, 111903.

Ren, L., Zhao, L., Hong, S., Zhao, S., Wang, H., & Zhang, L. (2018). Remaining useful life prediction for lithium-ion battery: a deep learning approach. *IEEE Access*, *6*, 50587-50598.

Shahjalal, M., Roy, P.K., Shams, T., Fly, A., Chowdhury, J.I., Ahmed, M.R., & Liu, K. (2022). A review on second-life of Li-ion batteries: prospects, challenges, and issues. *Energy*, *241*, 122881.

Silva, V.F., Silva, M.E., Ribeiro, P., & Silva, F. (2022). Novel features for time series analysis: a complex networks approach. *Data Mining and Knowledge Discovery*, *36*(3), 1062-1101.

Singh, M.K., Chaube, S., Pant, S., Singh, S.K., & Kumar, A. (2023). An integrated image visibility graph and topological data analysis for extracting time series features. *Decision Analytics Journal*, *8*, 100253.

Sun, J., Liu, X., Li, X., Chen, S., Xing, S., & Guo, Y. (2025a). State of health estimation of lithium-ion battery based on constant current charging time feature extraction and internal resistance compensation. *Energy*, *315*, 134472.

Sun, R., Chen, J., Li, B., & Piao, C. (2025b). State of health estimation for Lithium-ion batteries based on novel feature extraction and BiGRU-attention model. *Energy*, *319*, 134756.

Wang, J., Yang, H., Wang, Z., Zhou, Y., Liu, P., & Hong, J. (2024). State of health analysis of batteries at different stages based on real-world vehicle data and machine learning. *Journal of Energy Storage*, *88*, 111616.

Yang, K., Zhang, L., Zhang, Z., Yu, H., Wang, W., Ouyang, M., Zhang, C., Sun, Q., Yan, X., Yang, S., & Liu, X. (2023). Battery state of health estimate strategies: from data analysis to end-cloud collaborative framework. *Batteries*, *9*(7), 351.

Zhang, W., Ahmed, R., & Habibi, S. (2025). State-of-health estimation for lithium-ion batteries based on electrochemical impedance spectroscopy measurements combined with unscented Kalman filter. *Journal of Power Sources*, *625*, 235450.

Zhao, J., & Burke, A.F. (2023). Battery prognostics and health management for electric vehicles under industry 4.0. *Journal of Energy Chemistry*, *84*, 30-33. https://doi.org/10.1016/j.jechem.2023.04.042

Zhu, Z., Yang, Q., Liu, X., & Gao, D. (2022). Attention-based CNN-BiLSTM for SOH and RUL estimation of lithium-ion batteries. *Journal of Algorithms & Computational Technology*, *16*. https://doi.org/10.1177/17483026221130598

Zou, B., Wang, H., Zhang, T., Xiong, M., Xiong, C., Sun, Q., Wang, W., Zhang, L., Zhang, C., & Ruan, H. (2023). A deep learning approach for state-of-health estimation of lithium-ion batteries based on differential thermal voltammetry and attention mechanism. *Frontiers in Energy Research*, *11*, 1178151.

Zuo, H., Liang, J., Zhang, B., Wei, K., Zhu, H., & Tan, J. (2023). Intelligent estimation on state of health of lithium-ion power batteries based on failure feature extraction. *Energy*, *282*, 128794.