

# Early Cycle Life Prediction of Lithium-Metal Batteries with the Aid of Machine Learning

Mojtaba Eftekharnia, Tri Nguyen, Maria Forsyth, Truyen Tran, Patrick C. Howlett, and Robert Kerr\*

Battery cell manufacturing comprises numerous steps requiring co-optimization, making the development process time consuming and expensive. Lithium-metal batteries with ionic liquid electrolytes are a promising next-generation technology for applications demanding high specific energy and safety but currently suffer from limited cycle stability. Optimizing the manufacturing process can improve performance, and early cycle life prediction can accelerate this process, reducing cost and time. However, correlating early-stage behavior with long-term stability is challenging. Machine learning (ML) can assist in building these correlations, but feature extraction remains a key hurdle. A set of features

manually extracted from the first cycle with high correlation with the battery cycle life is presented here. These features are then used as inputs to a ML model based on linear regression. While the dataset contains batteries that have reached end-of-life through two different mechanisms, the model can predict the cycle life with an error of 15.3%. The error decreases to 9.6% when the cells are first sorted by end-of-life mechanism. This work highlights the importance of the early charge–discharge behavior of lithium-metal batteries and how this data can be used to inform on the battery cycle life with a view to greatly reducing experimental workload.

## 1. Introduction

Battery energy storage systems based on lithium have become an inseparable part of our society as they offer high energy density and long lifetime.<sup>[1]</sup> Although many outstanding lithium-ion battery technologies have been developed,<sup>[2,3]</sup> novel technologies must be further developed to meet the growing demand for higher energy density systems.<sup>[4]</sup> Battery technologies based on lithium metal offer the highest energy density due to the high specific capacity of 3860 mAh g<sup>−1</sup> and low electrochemical potential of lithium, −3.04 V versus the standard hydrogen electrode.<sup>[5]</sup> However, developing reliable, advanced battery technologies requires many considerations from materials development to manufacturing. This process involves optimization of a great number of variables each of which has an impact on the performance of the final cell.<sup>[6–9]</sup> Although a concerted effort involving advanced characterization techniques has enabled researchers to disentangle the effects of

the individual process steps, long-term battery cycling experiments are still required for validation purposes.<sup>[10]</sup>

Recent reports have demonstrated that time-consuming and capital-intensive battery cycling experiments can be largely avoided with the aid of machine learning (ML) through correlating the battery cycle life to the early cycle behavior.<sup>[11–13]</sup> Early cycle life prediction accelerates the overall development process and allows researchers to examine the effect of various parameters within a short period of time.<sup>[12]</sup> Therefore, developing techniques and models that can accurately predict the battery cycle life will be valuable. Owing to the fact that battery degradation processes are typically nonlinear and their prevalence is minimal during early cycles, developing accurate early prediction models is challenging.<sup>[11]</sup> Severson et al. trained an ML model based on linear regression to predict the cycle life of Graphite||Lithium iron phosphate (LFP) cells using data from the first 100 cycles and obtained a test error of 9.1%. The researchers manually identified a predictive feature from the discharge profiles that had a strong correlation coefficient of −0.93 to the cycle life.<sup>[11]</sup> Following this report, several groups applied different ML algorithms on the same data to automatically extract predictive features,<sup>[14,15]</sup> either to enhance the accuracy of early cycle life prediction models<sup>[16]</sup> or to reduce the number of early cycles required for accurate life prediction.<sup>[17,18]</sup> Saxena et al. designed a convolutional neural network to automatically identify the local patterns in discharge profiles of these graphite||LFP cells and extract predictive features from the entire discharge profiles of the first 100 cycles for prediction of the capacity fade curve.<sup>[14]</sup> However, they did not achieve lower prediction error compared to the previous study. Gong et al. combined manual and automatic feature selection techniques and applied them to the same data to achieve higher life prediction accuracy. They identified 20 degradation-related features from the first 100 cycles and then applied different

M. Eftekharnia, M. Forsyth, P. C. Howlett, R. Kerr  
Institute for Frontier Materials  
Deakin University  
221 Burwood Highway, Burwood, VIC 3125, Australia  
E-mail: Robert.kerr@deakin.edu.au

T. Nguyen, T. Tran  
Applied Artificial Intelligence Institute  
Deakin University  
75 Pigdons Road, Waurin Ponds, VIC 3216, Australia

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/batt.202500261>

© 2025 The Author(s). Batteries & Supercaps published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

feature selection techniques to find the best feature subsets and used these feature subsets to train seven different ML models. The best test error they obtained was 5.21%.<sup>[16]</sup> Fermin-Cueto et al. applied different feature selection techniques and selected 100 predictors by which they trained ML models based on the relevance vector machine technique. They used only the data from the first 50 cycles to predict cycle life and achieved a test error of 9.4%. This is superior to the model by Severson et al. in terms of required early cycle data for cycle life prediction.<sup>[17]</sup>

While there have been significant advancements in the application of ML for early prediction of the cycle life of Li-ion batteries,<sup>[11,14–17]</sup> there are no reports for this application on Li-metal batteries. Li-metal batteries offer more than 30% higher specific energy compared to the state-of-the-art Li-ion batteries, making them attractive for energy-demanding applications.<sup>[1,8,19]</sup> The degradation mechanisms of Li-metal batteries are not the same as those for Li-ion;<sup>[20,21]</sup> therefore, predictive models developed in the literature for Li-ion batteries are not necessarily applicable to Li-metal batteries. Given the extensive research being undertaken on Li-metal batteries,<sup>[9,22,23]</sup> developing ML models that can predict cycle life of Li-metal batteries is of great importance. This is particularly valuable in cases where a Li-metal battery technology has shown promising performance, but the manufacturing process has not been fully optimized. For example, Pathirana et al. showed that the applied current density during the formation process of Li-metal batteries using ionic liquid electrolytes plays a significant role in influencing the cycle life of these cells.<sup>[24]</sup> However, optimizing this variable alone would require numerous long-term experiments. In another study, we investigated the role of external compression on the cycle life of cells using a similar technology.<sup>[25]</sup> While a range of external compression levels were reported under which the Li-metal batteries performed well, the optimized value must be determined through further time-consuming studies. These studies illustrate that novel approaches like ML should be adopted to expedite the advancement of new battery technologies.

In this study, we present a feature-based ML model that predicts the cycle life of high-capacity Li-metal batteries with high level of accuracy. The features used in the model are all extracted from the voltage profiles of the first cycle after formation as well as the electrochemical impedance spectroscopy (EIS) data at the top of first charge. The dataset contains 43 coin cells which were comprised of two different types of ionic liquid electrolytes, two types of separators, 15 formation protocols, and exhibited two different end-of-life mechanisms. The results presented in this study highlight the importance of the initial properties of Li-metal batteries and their long-term impact on the battery performance.

## 2. Li-Metal Battery Cell Chemistry and Design

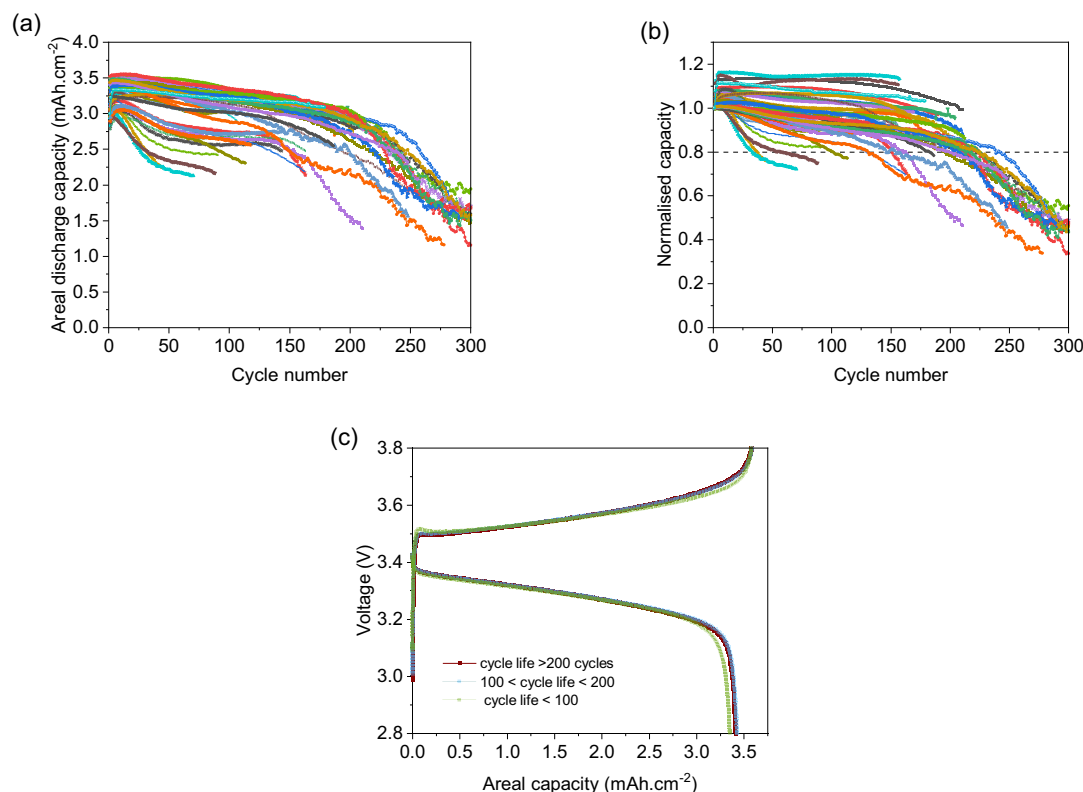
Researchers in our group have shown that ionic liquid electrolytes based on either *N*-propyl-*N*-methylpyrrolidinium bis(fluorosulfonyl)imide ( $C_3\text{mpyrFSI}$ ) or triethylmethylphosphonium bis(fluorosulfonyl)imide ( $P_{1222}\text{FSI}$ ) with high salt concentration can support efficient Li-metal full cell cycling at high areal capacities and high current densities.<sup>[26,27]</sup> These cells were able to show

stable cycling for more than 200 cycles with a capacity retention of 87%, which is one of the top results in the field. We have also explored the role of the separator in the long-term cycling performance of Li-metal batteries using these electrolytes, where it was found that the applied compression and evolution of cell pressure during cycling greatly impacts the cycling performance.<sup>[28,29]</sup> Additionally, the significant effects of the formation cycling protocol on the properties of the solid electrolyte interphase (SEI) in Li-metal batteries with these electrolyte systems have been studied.<sup>[24,30]</sup> These studies show that there is great scope to further improve the performance of these Li-metal batteries. However, optimizing the various parameters that affect the performance of batteries requires numerous experiments which are not only time consuming but also expensive. To build our ML model that can predict cell cycle life with minimal cycling data, we first generated a set of long-term cycling data. This approach potentially reduces the overall cell cycling time and expedites the process of optimization in the long run.

To generate the data, Li||LFP coin cells with high areal capacity of  $3.5 \text{ mAh cm}^{-2}$  were assembled using two different electrolytes, namely  $C_3\text{mpyrFSI}:\text{LiFSI}$  (1:1 molar ratio) and  $P_{1222}\text{FSI}:\text{LiFSI}$  (1:1 molar ratio) and two different types of surfactant-coated polypropylene Celgard separators, namely Celgard 3501 and 3401. The coin cells were conditioned under 15 different formation protocols (Table S1, Supporting Information) prior to long-term charge and discharge cycling at C/2 at 50 °C. The current, voltage, time, capacity, energy, and impedance at the top of charge were all recorded for further analysis.

The cell cycle life ranged from 59 to 242 cycles. The cycle life is defined as the number of cycles at which the cell reached its end-of-life either due to short circuit or reached 80% of its initial capacity. The dataset comprises 43 coin cells—20 of which were terminated due to short circuit and the rest reached 80% of their initial capacity. It should be noted that the capacity of the cells which were terminated due to short circuit had not reached 80% of their initial capacity at the time of termination. Hereafter, the terms ‘Complete’, ‘SC’, and ‘R80%’ refer to the complete dataset, the short-circuit dataset, and the dataset which reached 80%, respectively. Four out of 43 cells show high prediction error that we have attributed to defects in the cell construction. These specific cells exhibited performance characteristics and failure modes that were significantly divergent from the main cohort of cells. We attribute these deviations to stochastic manufacturing defects or assembly issues rather than the systemic degradation pathways that our model aims to capture through early cycle electrochemical signatures. These cells are excluded from the data set used to build the model. The prediction performance of the model including these high-error cells is presented in Table S2, Supporting Information.

Figure 1a,b shows the areal discharge capacity and nominal capacity of the cells as a function of the cycle number. Figure 1c shows a comparison of the voltage profiles of the first cycle at C/2 for three selected cells which had either low, medium, or long cycle life. The only manufacturing difference between these three cells was the formation protocol. While the formation protocol had a profound impact on the cell cycle life, it has not affected the 1st cycle voltage profiles notably. Despite this, we have been



**Figure 1.** a) Areal discharge capacity and b) normalized discharge capacity of the Li||LFP coin cells with ionic liquid electrolyte, and c) charge and discharge profiles of the first cycle at C/2.

able to construct a model which gives a strong correlation between the first cycle behavior and the cell cycle life.

### 3. Feature Extraction

We use a prediction framework based on elastic net regression to predict the end-of-life Li-metal batteries. Elastic net is a regularized regression method. It is a common choice to use for ML problems with a large number of features which may have some predictive power. Elastic net's penalty term is the combination of the LASSO (L1) and Ridge (L2) regularization term. The L1 regularization term can shrink the coefficient of unimportant features to zero, and thus, is useful for feature selection. The L2 regularization term prevents the coefficient from becoming large to prevent the model overfitting.<sup>[31]</sup>

In this work, we manually extracted 16 features from early cycles based on battery domain knowledge and their Pearson correlation<sup>[32]</sup> with cycle life across the dataset. These features were extracted from the early cycles and do not contain degradation information and reflect only the internal properties of the cells like the composition and structure of the SEI, the lithium deposit morphology, the state of the separator, and the state of the cathode, all of which can have an impact on the long-term cell performance. Importantly, this feature selection was performed once, prior to model training, and was not modified or reoptimized during cross validation. For model evaluation, we used leave-one-out cross validation (LOOCV), where the model

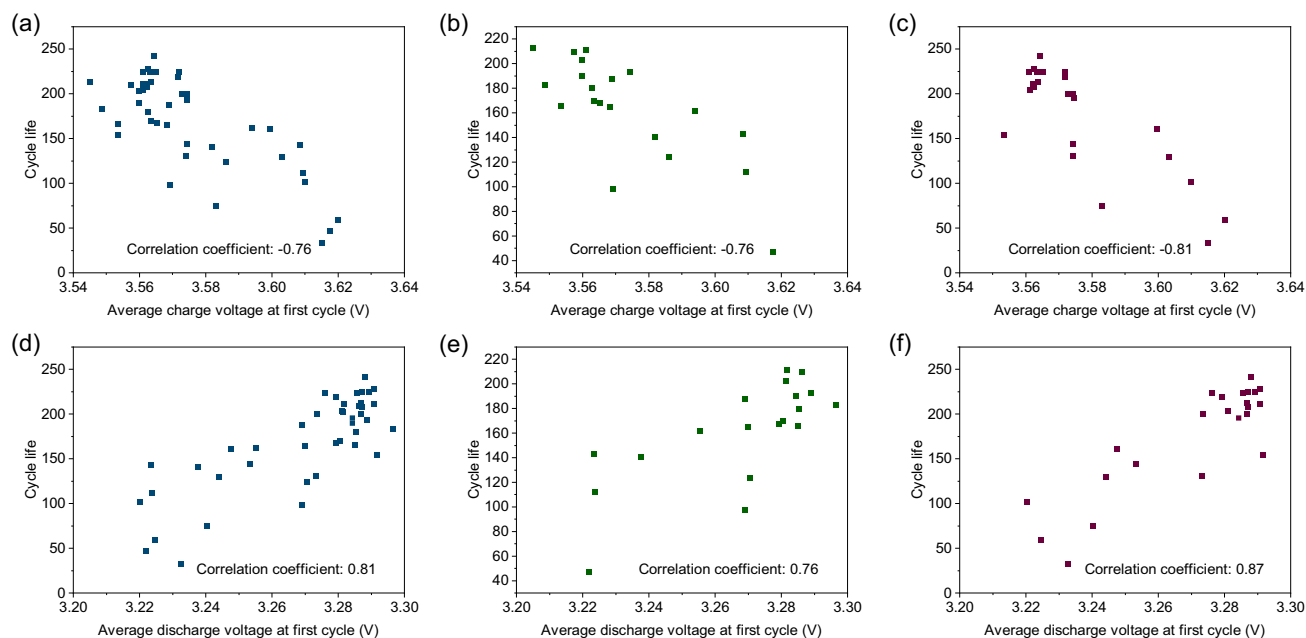
was trained on  $N - 1$  cells using the fixed 16 features and tested on the held-out cell. This ensured that the test data was not used in either feature selection or model fitting, avoiding data leakage.

By using the energy and capacity data, we calculated the average voltage during charge and discharge. To accurately calculate the average charge/discharge voltage, the cell energy at the end of the charge/discharge steps is divided by the charge/discharge capacity (below).

$$\text{Average voltage (V)} = \frac{\text{energy (Wh)}}{\text{capacity (Ah)}} \quad (1)$$

These two features are highly correlated to the long-term cell performance. **Figure 2** shows the strong correlation of the average charge and discharge voltages at the first cycle after formation with the cell cycle life. The correlation between first cycle average charge voltage and average discharge voltage with cycle life of the complete dataset is  $-0.76$  and  $0.81$ , respectively (Figure 2a,d). Although the end-of-life mechanisms of the cells are completely different, these features have an excellent predictive capability for both datasets.

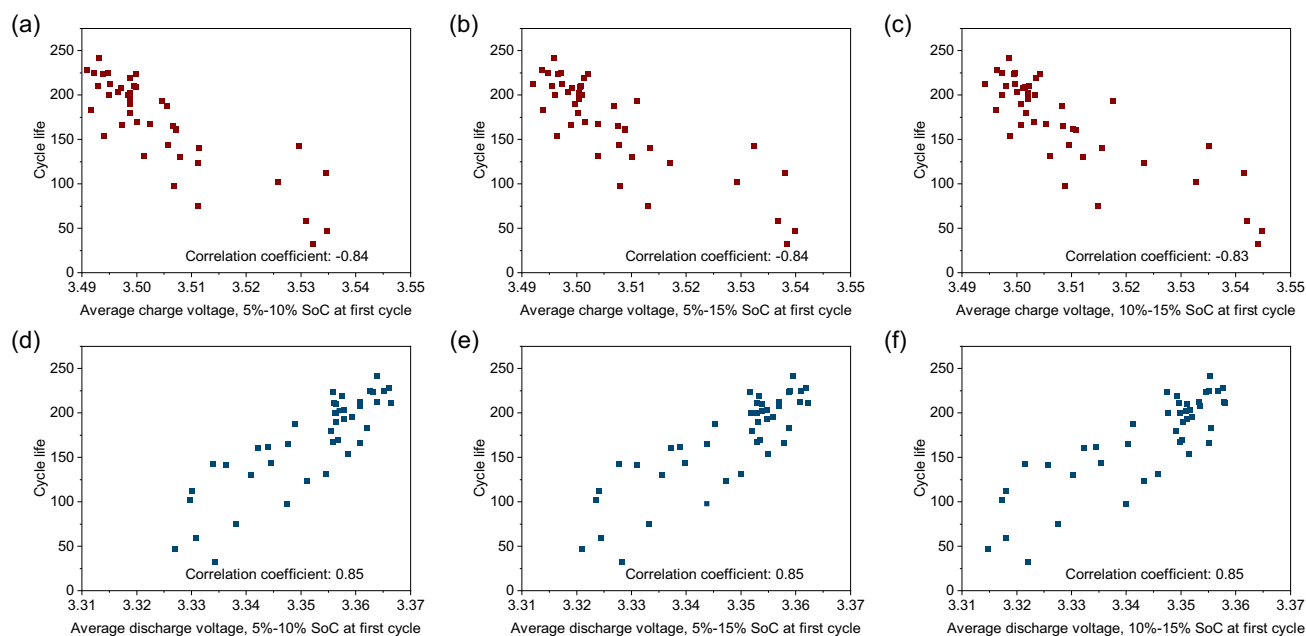
Variations in these two features reflect differences in the cell internal resistance over the entire charge and discharge steps. However, Figure 1b shows that the voltage profiles during the early phase of charge meaningfully differ for the three selected cells. Charge/discharge profiles are indicative of the kinetics of the electrochemical reactions in general.<sup>[33]</sup> However, careful analysis must be done to deconvolute the main mechanisms that



**Figure 2.** Cycle life as a function of a,b,c) the average charge voltage and d,e,f) average discharge voltage. (a,d) The complete dataset, (b,e) SC dataset, and (c,f) R80% dataset.

affect the voltage profiles. In the case of Li-metal batteries, voltage profiles provide useful information about Li plating and stripping behavior on and from the Li anode.<sup>[34]</sup> The kinetics of Li plating/stripping in early cycles could potentially be correlated to the long-term performance of Li-metal cells. We believe that the state of lithium-metal electrodes, like SEI composition and structure as well as lithium deposition morphology at the beginning and end of a cycle, would impact the longevity of these batteries. The state of lithium-metal electrodes is highly affected by

the formation cycling protocols which reflects themselves in the initial charge and discharge cycles as shown by Pathirana et al.<sup>[24,30]</sup> Therefore, extracting features that are directly related to the electrochemical processes would provide useful insights about the long-term behavior of the cells. More features based on the average charge voltage from the first 5%, 10% and 15% state of charge (SoC) as well as the average discharge voltage over the first 5%, 10%, and 15% (from 100% SoC) were extracted (Figure 3 and Table S3, Supporting Information).



**Figure 3.** Cycle life as a function of a) average charge voltage from 5% SoC to 10% SoC, b) average charge voltage from 5% SoC to 15% SoC, c) average charge voltage from 10% SoC to 15% SoC, d) average discharge voltage from 5% SoC to 10% SoC, e) average discharge voltage from 5% SoC to 15% SoC, and f) average discharge voltage from 10% SoC to 15% SoC of the complete dataset.

These features, which represent the cell behavior within the first 18 min of the first charge/discharge steps, exhibit a superior correlation coefficient compared to the features extracted from over the entire step. This shows the significance of the early behavior of Li plating/stripping that dictates the long-term performance of the Li-metal cells.

## 4. Results and Discussion

### 4.1. Model Performance

To evaluate the prediction performance of the models, we used LOOCV approach. Due to the small number of cells in this dataset, LOOCV is the most suitable evaluation method. Additionally, the low computational cost of the linear regression favored the use of this approach. However, to assess the susceptibility of the model to overfitting and underfitting, we also used K-fold cross validation where the dataset was divided into five parts, four of which were used for model training and one used for performance evaluation. All the K-fold cross validation results are presented in the Supporting Information. The prediction performance of the first model, which uses only the average charge and discharge voltage of the first cycle, is shown in Table 1. The K-fold cross validation results are presented in Table S4, Supporting Information. Using only these two features result in a 17% mean absolute percentage error (MAPE), which is excellent given the fact that the dataset includes cells with different end-of-life mechanisms. Surprisingly, these features work best when applied to the SC dataset with a MAPE of 10.9%.

EIS data in the frequency range of 1 Hz–10 kHz (10 mV amplitude) at the first cycle top of charge is also shown to be a strong predictor of the cell cycle life. The model performs well using the EIS data at the top of charge of the first cycle for both SC and R80% datasets (MAPE of 11.3% and 15.0%, respectively, Table 1). However, the performance of this feature for complete dataset is not as strong (MAPE of 19.3%). This suggests that there is a possible link between the end-of-life mode and the first cycle EIS at the top of charge.

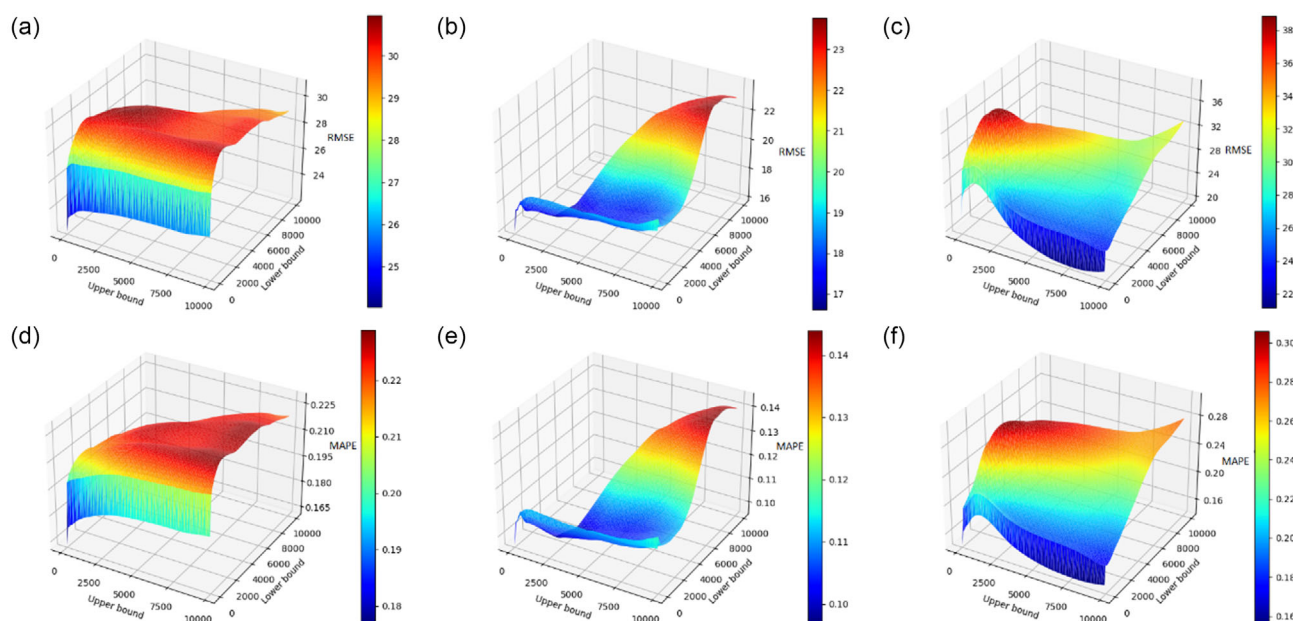
Zhang et al. have reported that the EIS data of LR2032 LiCoO<sub>2</sub>||graphite lithium-ion batteries at certain frequencies, namely 17.8 and 2.16 Hz, have higher ability to predict the battery remaining useful life compared to the entire spectrum.<sup>[35]</sup> Si et al. have also shown that the features extracted from the low-frequency region (between 0.2 and 13 Hz) of the EIS spectrum are more important to predict the capacity of Li-metal||NMC<sub>811</sub> batteries although they applied their ML model on a small data

set of 16 batteries.<sup>[36]</sup> EIS data in specific frequency ranges correspond to distinct electrochemical phenomena: high frequencies reflect bulk and interfacial resistances (e.g., SEI and ohmic resistance), mid frequencies are associated with charge transfer at the anode and cathode, and low frequencies are indicative of ion diffusion within the electrode materials. Very low frequencies may also reveal long-term degradation processes.<sup>[37,38]</sup> Identifying the phenomena that have the highest impact on the battery performance can enhance the process of technology development. In addition, less data with more predictive information is more desirable as it results in less expensive and more efficient data acquisition. Therefore, we analyzed the EIS data and extracted the frequency range that provided lower prediction error. Figure 4 shows the root mean square error (RMSE) and MAPE of the model when the data from various frequency ranges were used. The optimal frequency range for the complete and SC datasets is 1–2 Hz while this range is 1–7800 Hz for the R80% dataset. Using the EIS data within the optimal frequency range improves the MAPE of the model from 11.3% to 9.6% for the SC dataset (Table 1), which is a significant improvement. However, limiting the frequency range does not have the same magnitude of positive impact on the R80% dataset (15.0% to 14.5%). The low-frequency range of 1–2 Hz can be attributed to blended effects between charge transfer resistance and ion diffusion at the interface of Li-metal anode and electrolyte while wide frequency range of ~7800 Hz captures multiple effects such as ion diffusion, charge transfer resistance, and SEI resistance. In Li-metal batteries, the morphology and uniformity of Li-metal deposits on anode play a significant role in their longevity.<sup>[39]</sup> Initial lithium nucleation on the surface Li-metal anode, which is influenced by formation cycling,<sup>[30,39]</sup> affects battery longevity and possibly is reflected in the EIS spectrum within the 1–2 Hz frequency range.

Using extra features, extracted from the first cycle, lowers the MAPE particularly for the R80% dataset to 11.7% (Table 2). These features, listed in Table S3, Supporting Information, are all based on the early behavior of the cells in the first charge and discharge. Combining these features with the EIS data in the optimal frequency range results in excellent prediction performance of the model with a MAPE of 15.3%, 11.2%, and 11.4% on the complete, SC, and R80% datasets, respectively (Table 2). Figure 5 shows that the Optimized model can capture the degradation pattern easily in R80% cells while showing some challenges in SC cells. The K-fold cross validation results are presented in Table S5, Supporting Information, that do not show significant difference with LOOCV results. Further characterization of the Li-metal anode,

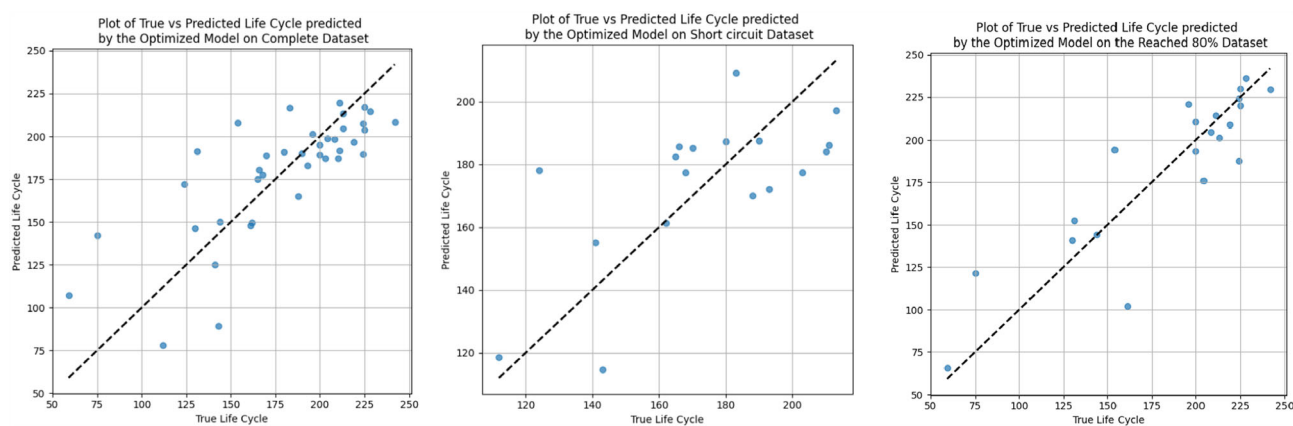
**Table 1.** Prediction performance of the models using the average charge and discharge voltages and EIS at the first cycle.

Models	Features	Complete dataset		SC dataset		R80% dataset	
		RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
"Average voltage"	1 and 8	24.14 (17.47)	17.0 (23.0)	17.64 (9.96)	10.9 (8.0)	24.7 (22.00)	19.2 (26.7)
"Impedance"	15 and 16	25.155 (23.306)	19.3 (30.1)	19.015 (14.099)	11.3 (9.9)	20.682 (14.931)	15.0 (17.2)
"Optimal impedance"	15 and 16 within optimal frequency range	22.048 (17.913)	16.2 (21.6)	15.855 (12.973)	9.6 (9.6)	19.57 (17.309)	14.5 (18.9)



**Figure 4.** a–c) RMSE and d–f) MAPE of the model using the EIS data in various frequency ranges for (a,d) complete dataset, (b,e) SC dataset, and (c,f) R80% dataset.

		Complete dataset		SC dataset		R80% dataset	
Models	Features	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
"First cycle"	1–16	22.355 (18.770)	16.8 (23.0)	20.927 (14.246)	12.6 (10.4)	17.537 (15.522)	11.7 (14.2)
"Optimized"	1–14 as well as 15 and 16 within optimal frequency range	21.002 (17.036)	15.3 (19.7)	18.484 (11.893)	11.2 (9.1)	16.687 (16.048)	11.4 (14.4)



**Figure 5.** The plot of predicted life cycle of the Optimized model and the true life cycle for complete dataset, SC dataset, and R80% dataset.

the composition and structure of SEI layer, the state of separator, and the state of the cathode is required to identify the critical characteristics of the cell components in early life that can lead to improved long-term stability and performance.

We further analyzed the cycling data over the first ten cycles and extracted more features which contain more historic information (Table S6, Supporting Information). However, these features did not improve the prediction performance of the

model (Table S7, Supporting Information). This suggests that the degradation behavior of the cells over the first ten cycles is not as informative as the initial properties of the cells.

#### 4.2. The Effect of Formation Protocols

Li plating/stripping behavior is dictated by many factors, one of which is the SEI layer characteristics.<sup>[5,21]</sup> The properties of the SEI

layer have been shown to be influenced by various parameters, including the type of electrolyte, temperature, current density, and formation cycling protocol.<sup>[24]</sup> The SEI layer is mostly formed during the initial formation cycles; therefore, the impact of these formation cycles should manifest in the early cycles. Moreover, the formation cycles affect the established morphology of the Li deposits, which directly impacts the short circuit susceptibility of Li-metal batteries.<sup>[30]</sup>

The formation protocols applied in this study can be categorized into three main groups: low current density, high current density, and hybrid (Table S1, Supporting Information). The cells with low current formation regime sustained  $180 \pm 30$  cycles, which is above the average cycle life of the complete dataset ( $170 \pm 52$  cycles). However, 92% of these cells suffered a short circuit, which is not desirable. Short circuits not only result in premature cell failure but also pose serious safety risks; hence, strategies to predict and avoid short circuits are very important. These results show that the occurrence of short circuits, which has its root cause in the morphology of the Li deposits, uniformity of the current distribution, and the integrity of the separator, is predominantly affected by the formation process and electrolyte composition. There is no conclusive relationship between the high current formation protocols and either the end-of-life mechanism or the cycle life of the cells. The average cycle life of these cells is  $155 \pm 56$ , which is below the average, and their failure rate due to short circuit is 36%, which is an improvement compared to the cells formed at low current density. This behavior suggests that tuning the cell properties by high current density formation regimes is not consistent and perhaps there are other factors, like the initial state of cell components, that contribute to variability in this behavior. However, combining the low current and high current protocols into a hybrid formation process results in interesting cycle life behavior. 100% of the cells on which hybrid formation protocols were applied consistently lasted the longest ( $225 \pm 10$  cycles) without any occurrence of short circuit.

## 5. Conclusion

Models based on ML provide an excellent tool to reduce the number and duration of the experiments required for optimizing the various parameters in the process of new battery technology development. In this study, we analyzed the voltage profiles of the first cycle after formation of 43 Li-metal coin cells, which had different electrolyte and separator types and had undergone different formation cycling, and identified predictive features with high correlation coefficients to the achievable cycle life. These features are mostly based on the average charge/discharge voltage within the first 15% of the charge/discharge steps. By employing these features in a regression model to predict the cycle life, in the best case we achieved an error of 9.6% which is excellent as only the initial 1% of the whole cycling dataset was used. The extracted features are useful for developing a predictive model and also provide interesting insights into the critical initial internal properties of Li-metal batteries with ionic liquid electrolyte systems. We also showed that the initial formation cycling regimes have a profound impact on the cell cycle life. While formation protocols with low

current density generally resulted in higher cycle life and higher rate of short circuit, the ones with high current density resulted in lower cycle life but with a lower incidence of short circuits. Combining these two protocols into a hybrid version benefited the cells with extended cycle life and zero short circuit failure. There is scope to improve cell performance by optimizing the formation protocols using ML. While ML has not been used here for optimization purposes, there is a need for further work to generate datasets and build models for formation protocols. By adapting an early cycle life prediction approach previously established for Li-ion battery chemistries, it has been shown here that it can—through careful design of feature extraction—also diminish the need for long-term cell cycling for lithium-metal batteries and in doing so accelerate the process of optimization and reduce the cost of research.

## 6. Experimental Section

### Material Preparation, Coin Cell Fabrication, and Cycling

Electrolyte was prepared by adding LiFSI (>99.5%, Nippon Shokubai) to either *N*-methyl-*N*-propylpyrrolidinium bis(fluorosulfonyl)imide (C<sub>3</sub>mpyrFSI, Boron Molecular) or triethylmethylphosphonium bis(fluorosulfonyl)imide (P<sub>1222</sub>FSI, Boron Molecular) with an IL:LiFSI molar ratio of 1:1. Ionic liquids and LiFSI were vacuum dried at 50 °C for 48 h in a Schlenk line with pressure of <1 mbar prior to mixing in a glovebox (Korea Kiyon).

Li||LFP coin cells were prepared in an argon-filled glovebox (Korea Kiyon) with oxygen and moisture levels of less than 1 ppm. The 100-μm Li-metal (Gelon) electrodes were brushed and rinsed with dried cyclohexane (Merck) prior to assembly. LFP electrodes (Customcells) with 3.5 mAh cm<sup>-2</sup> areal capacity and 89% active material content were cut and dried under vacuum at 110 °C for 16 h prior to transferring to glovebox and coin cell fabrication.

15 different formation protocols (Table S1, Supporting Information) were applied prior to long-term cycling at C/2 between a voltage range of 2.8–3.8 V. The cells were rested for 30 min after each charge and discharge step. Potentiostatic EIS was performed at open circuit voltage at the top of the charge (after 30 min rest) in the frequency range of 10 kHz–1 Hz with an amplitude of 10 mV using a BCS-810 (Biologic) battery cycler.

### ML Approach: Manual Feature Engineering and Preselection

Initially, a set of features was engineered based on electrochemical domain knowledge and insights from prior battery research. These features, derived from first-cycle voltage profiles and EIS data, were then preselected based on their individual Pearson correlation coefficients<sup>[32]</sup> with the cell cycle life, as shown in Figure 2 and 3 and detailed in Table S3, Supporting Information. This step constituted our primary feature reduction and selection phase.

### ElasticNet Regularized Regression

The ElasticNet<sup>[31]</sup> model was subsequently applied to this preselected and fixed set of features. While the L1 component of ElasticNet inherently possesses the capability to drive coefficients of irrelevant or redundant features to exactly zero, thereby performing feature selection, in the case of our 'Optimized' model (which utilized all features listed in Table S3, Supporting Information), none of the coefficients were shrunk to absolute zero (as shown in Table S8, Supporting

Information). This outcome suggests that all the manually preselected features retained some level of predictive contribution within the ElasticNet framework when considered collectively. Therefore, the primary role of ElasticNet in this study was to build a robust regression model that mitigates overfitting through L1 and L2 regularization and effectively handles potential multicollinearity among the chosen features, rather than to perform an automated selection from a much larger, unrefined pool of initial candidate features. The L2 component, for instance, helps in stabilizing the model when features are correlated.

## Acknowledgements

This research was supported by the Australian Research Council (ARC) through the ARC Training Centre for Future Energy Storage Technologies (IC180100049). This research was undertaken at the Deakin University Battery Research and Innovation Hub, Australia.

Open access publishing facilitated by Deakin University, as part of the Wiley - Deakin University agreement via the Council of Australian University Librarians.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are openly available in a GitHub repository at <https://github.com/ngminhtri0394/BatteryLifeCycle/tree/master>.

**Keywords:** cycle life • degradation • electrochemistry • lithium • machine learning

- [1] T. Placke, R. Kloepsch, S. Dühnen, M. Winter, *J. Solid State Electrochem.* **2017**, *21*, 1939.
- [2] T. Kim, W. Song, D.-Y. Son, L. K. Ono, Y. Qi, *J. Mater. Chem. A* **2019**, *7*, 2942.
- [3] A. Manthiram, *ACS Cent. Sci.* **2017**, *3*, 1063.
- [4] M. Li, J. Lu, Z. Chen, K. Amine, *Adv. Mater.* **2018**, *30*, 1800561.
- [5] D. Lin, Y. Liu, Y. Cui, *Nat. nanotechnol.* **2017**, *12*, 194.
- [6] J. B. Goodenough, Y. Kim, *Chem. Mater.* **2010**, *22*, 587.
- [7] J.-M. Tarascon, M. Armand *Materials for sustainable energy: A collection of peer-reviewed research and review articles from Nature Publishing Group*, World Scientific Singapore **2011**, pp. 171–179.
- [8] R. Schmich, R. Wagner, G. Hörpel, T. Placke, M. Winter, *Nature Energy* **2018**, *3*, 267.
- [9] J. Liu, Z. Bao, Y. Cui, E. J. Dufek, J. B. Goodenough, P. Khalifah, Q. Li, B. Y. Liaw, P. Liu, A. Manthiram, *Nat. Energy* **2019**, *4*, 180.
- [10] J. S. Edge, S. O'Kane, R. Prosser, N. D. Kirkaldy, A. N. Patel, A. Hales, A. Ghosh, W. Ai, J. Chen, J. Yang, *Phys. Chem. Chem. Phys.* **2021**, *23*, 8200.
- [11] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggadakis, *Nat. Energy* **2019**, *4*, 383.
- [12] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, *Nature* **2020**, *578*, 397.
- [13] Z. Tong, J. Miao, S. Tong, Y. Lu, *J. Cleaner Prod.* **2021**, *317*, 128265.
- [14] S. Saxena, L. Ward, J. Kubal, W. Lu, S. Babinec, N. Paulson, *J. Power Sources* **2022**, *542*, 231736.
- [15] N. H. Paulson, J. Kubal, L. Ward, S. Saxena, W. Lu, S. J. Babinec, *J. Power Sources* **2022**, *527*, 231127.
- [16] D. Gong, Y. Gao, Y. Kou, Y. Wang, *J. Energy Storage* **2022**, *51*, 104376.
- [17] P. Fermin-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M. F. Anjos, J. Sylvester, G. Dos Reis, *Energy AI* **2020**, *1*, 100006.
- [18] C. Strange, G. Dos Reis, *Energy AI* **2021**, *5*, 100097.
- [19] C. Niu, H. Lee, S. Chen, Q. Li, J. Du, W. Xu, J.-G. Zhang, M. S. Whittingham, J. Xiao, J. Liu, *Nat. Energy* **2019**, *4*, 551.
- [20] K.-H. Chen, K. N. Wood, E. Kazyak, W. S. LePage, A. L. Davis, A. J. Sanchez, N. P. Dasgupta, *J. Mater. Chem. A* **2017**, *5*, 11671.
- [21] W. Xu, J. Wang, F. Ding, X. Chen, E. Nasybulin, Y. Zhang, J.-G. Zhang, *Energy Environ. Sci.* **2014**, *7*, 513.
- [22] H. Wang, Z. Yu, X. Kong, S. C. Kim, D. T. Boyle, J. Qin, Z. Bao, Y. Cui, *Joule* **2022**, *6*, 588.
- [23] F. Duffner, N. Kronemeyer, J. Tübke, J. Leker, M. Winter, R. Schmich, *Nat. Energy* **2021**, *6*, 123.
- [24] T. Pathirana, D. A. Rakov, F. Chen, M. Forsyth, R. Kerr, P. C. Howlett, *ACS Appl. Energy Mater.* **2021**, *4*, 6399.
- [25] M. Eftekharnia, R. Kerr, M. Forsyth, P. C. Howlett, *J. Power Sources* **2023**, *559*, 232650.
- [26] H. Yoon, P. Howlett, A. Best, M. Forsyth, D. MacFarlane, *J. Electrochem. Soc.* **2013**, *160*, A1629.
- [27] R. Kerr, N. Singh, T. S. Arthur, T. Pathirana, F. Mizuno, K. Takechi, M. Forsyth, P. C. Howlett, *Sustainable Energy Fuels* **2018**, *2*, 2276.
- [28] M. Eftekharnia, M. Hasanpoor, M. Forsyth, R. Kerr, P. C. Howlett, *ACS Applied Energy Mater.* **2019**, *2*, 6655.
- [29] M. Hasanpoor, M. Eftekharnia, T. Pathirana, U. Pal, R. Kerr, M. Forsyth, P. C. Howlett, *ACS Appl. Energy Mater.* **2021**, *4*, 6310.
- [30] T. Pathirana, R. Kerr, M. Forsyth, P. C. Howlett, *J. Electrochem. Soc.* **2020**, *167*, 120526.
- [31] H. Zou, T. Hastie, *J. R. Stat. Soc. Ser. B Stat. Method* **2005**, *67*, 301.
- [32] Pearson's Correlation Coefficient, *Encyclopedia Of Public Health* (Ed: W. Kirch) Springer Netherlands **2008**, pp. 1090–1091
- [33] R. Huggins, *Advanced batteries: Materials science aspects*, Springer Science & Business Media USA **2008**.
- [34] K. Borzutzki, J. R. Nair, M. Winter, G. Brunklaus, *ACS Appl. Mater. Interfaces* **2022**, *14*, 5211.
- [35] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, A. A. Lee, *Nat. Commun.* **2020**, *11*, 1706.
- [36] Q. Si, S. Matsuda, Y. Ando, T. Momma, Y. Tateyama, *Adv. Sci.* **2025**, 2502336.
- [37] S. T. Bakenhaster, H. D. Dewald, *J. Appl. Electrochem.* **2025**, *55*, 1657.
- [38] N. Meddings, M. Heinrich, F. Overney, J.-S. Lee, V. Ruiz, E. Napolitano, S. Seitz, G. Hinds, R. Raccichini, M. Gaberšček, *J. Power Sources* **2020**, *480*, 228742.
- [39] A. Pei, G. Zheng, F. Shi, Y. Li, Y. Cui, *Nano lett.* **2017**, *17*, 1132.

Manuscript received: April 7, 2025

Revised manuscript received: July 8, 2025

Version of record online: