Scholar Publishing Group
Academic Journal of Energy
https://doi.org/10.38007/RE.2025.040101
ISSN 2790-2714 Vol. 4, Issue 1: 1-10

# *Analysis of Li-ion Batteries Health Estimation Technology Route Based on Patent Big Data*

**Xiaofen Fang[1,a*], Qin Fang[2,b], Qichang Zhuo[3,c], Tao Li[4,d]**

*[1]Quzhou Quantitative Intelligence Technology Co., Ltd., Zhejiang, China*
*[2]Nantong Jingxiang Technology Co. Ltd., Jiangsu, China*
*[3]Zhejiang Huayou Cobalt Co., Ltd., Zhejiang, China*
*[4]Jiangshan Hechuang Rapid Manufacturing Technology Co., Ltd., Zhejiang, China*
*[a]fangxiaofen@ieee.org, [b]Jfly@Jfly-tech.com, [c]zhuonianer@163.com, [d]interlt@163.com*
*[*]Corresponding author*

***Keywords:*** Li-ion batteries; patent analysis; big data; health estimation; technology route

***Abstract:*** With the continuous advancement of artificial intelligence technology, its application in lithium battery health prediction has been increasingly widespread. From the vantage point of the variations in the number of patent applications and focal points, and by leveraging the IPC classification, an analysis is carried out on the macro-distribution of patent applications and the burgeoning growth of relevant applications. Through the text data within the patents, the branches and technical trajectories of artificial intelligence based lithium battery health prediction technology are gradually becoming more distinct. This approach enables the prediction of technological development trends and challenges in this field, thereby assisting research institutions and enterprises in closely tracking the technological progress in a specific domain.

## 1. Introduction

Lithium-ion batteries have become a staple of numerous applications due to their advantageous properties, including high energy density, long service life, high power tolerance, low self-discharge rate, and green environmental protection. These batteries are utilized in a variety of fields, including electric vehicles, electronic products, power tools, photovoltaic power stations, wind power stations, home energy storage, backup power supplies, and more[1]. However, it is important to note that after extended periods of use and repeated charge and discharge cycles, the capacity and power of lithium-ion batteries may degrade to varying extents. Overcharging of the negative electrode, overcharging of the positive electrode, and the influence of ambient temperature on the exterior of the lithium battery can cause irreversible changes inside the battery. Decay of the lithium battery can lead to a decline in charge and discharge performance, as well as the potential for safety hazards such as thermal runaway. Real-time monitoring of critical battery parameters, including voltage, current, temperature, charge, discharge efficiency, and SOH, is paramount in averting the risk of overcharging, over-discharging, and overheating, thereby mitigating the potential for combustion or explosion. Among these parameters, SOH serves as a pivotal indicator of the battery's health status. SOH monitoring enables the prevention of battery overcharging, over-discharging, and other scenarios that could compromise its integrity, thus facilitating informed decision-making by users regarding the utilization and maintenance of the battery. This, in turn, ensures the safety and reliability of the battery [2].

There are three main types of SOH estimation methods for lithium-ion batteries: (1) empirical estimation methods, mainly based on mathematical statistics, such as the cycle number method and the ampere-hour method [3]. (2) Model-based methods are based on the model of the internal chemical process of the battery. Based on Ohm's law, Kirchhoff's voltage-current law, electrochemical reaction process (Butler-Volmer law), etc., the influence of aging process and stress factors on state variables are analyzed [4]. For example, the impedance spectrum curve is measured at different stages of the battery cycle life, and the battery equivalent circuit model is obtained based on the impedance spectrum curve. Then, the influence of parameters such as solution resistance, charge transfer resistance, and Warburg impedance in the cycle number and equivalent circuit model is analyzed. (3) Data-driven estimation methods, including support vector machine (SVM), autoregressive moving average (ARMA) [5], particle filter (PF), and neural network (NN). With the rapid development of artificial intelligence technology, the application of deep learning algorithms to lithium battery SOH prediction has gradually become popular. Among them, representative algorithms include CNN [6], RNN, LSTM, GAN [7], and Attention Mechanism [8].

Data-driven prediction does not require knowledge of the mechanism of the object system [9]. Based on the collected data, it mines the implicit information through various data analysis and learning methods to make predictions, thus avoiding the complexity of model acquisition. It is a more practical prediction method. The number of patent applications for lithium-ion battery SOH prediction methods based on data-driven has increased sharply year by year, generating a large amount of patent text data. The data of the patent application content is analyzed and the law is mined to identify the direction of technological evolution. By analyzing these patent application text data and the trends of patent applications, the future evolutionary route of the data-driven battery SOH prediction method is predicted. For example, new algorithm trends, design improvements of battery management systems, and new strategies for optimizing battery performance may be discovered. It is very valuable for battery manufacturers, automakers, and decision makers in various fields involved in battery technology. Through analysis based on patent big data, related industries can better understand market trends and adjust their R&D directions and business strategies to remain competitive.

## 2. Analysis Methods Based On Patent Big Data

Analysis based on patent big data is an important means of technological research, market decision-making and innovation planning [10]. Its core lies in mining the implicit rules in multidimensional data such as patent texts, applicants, and technology classifications through systematic methods.

### 2.1. Data Sources

Mainstream patent databases (e.g., WIPO, USPTO, CNIPA, Derwent Innovation, Patentics), commercial databases (e.g., Wisdom Buds, incoPat), and open-source tools (e.g., Google Patents

API), this study selects incoPat as the core data source. It contains 120 million+ patents in the world's mainstream patent databases (including the OCR text of the full-text specification of Chinese patents), integrates patent data from 120+ countries around the world, and supports advanced processing functions such as IPC/CPC intelligent classification and applicant normalization. Eliminate duplicate patents, and non-related patents (e.g., designs), standardize classification symbols (e.g., IPC, CPC), and applicant names (unify different naming of the same subject) by extracting key fields (title, abstract, claims, priority date, field of art, inventor, legal status) and performing data cleansing.

## 2.2. Keywords and IPC Codes

Keyword combination ("lithium-ion battery*" OR "Li-ion battery*") AND ("SOH" OR "state of health" OR "remaining useful life" OR "RUL") AND ("data-driven" OR "machine learning" OR "deep learning" OR "AI"). Lithium battery degradation prediction technology is mainly concentrated in battery monitoring technology (G01R31/36) and battery management systems (H01M10/48). Such patents are also distributed in machine learning applications (G06N3/00), electric vehicle applications (B60L58/10), and energy storage systems (H02J7/00). IPC classification code combination (IPC: G01R31/36 OR IPC: H01M10/48 OR IPC: G06N3/00) AND (IPC: Y02E60/32 OR IPC: B60L58/10). At the same time, combined with semantic retrieval methods, technical terms in patent titles/abstracts ("capacity trajectory prediction" and "ultra-early life prediction"), are supplemented with synonyms (such as "health estimation" and "degradation modeling").

## 2.3. Data samples and statistical analysis

Through the incoPat patent database, a comprehensive search was conducted on March 20, 2025, and after integrating relevant information such as the same family applications, a total of 1,072 patent data related to the lithium-ion battery health status (Li-ion Batteries Health Estimation) evaluation method and device were obtained. From a time dimension, patent applications in this field have shown phased changes [11][12]. In 2017, the first related patent application appeared, marking the beginning of attention in this field, as shown in Figure 1. Since then, the number of applications has gradually increased, and by 2024, the number of patent applications has shown a rapid growth trend. Different countries show different levels of activity in this field, with China, the United States, Japan, and South Korea being particularly active, as shown in Figure 2.

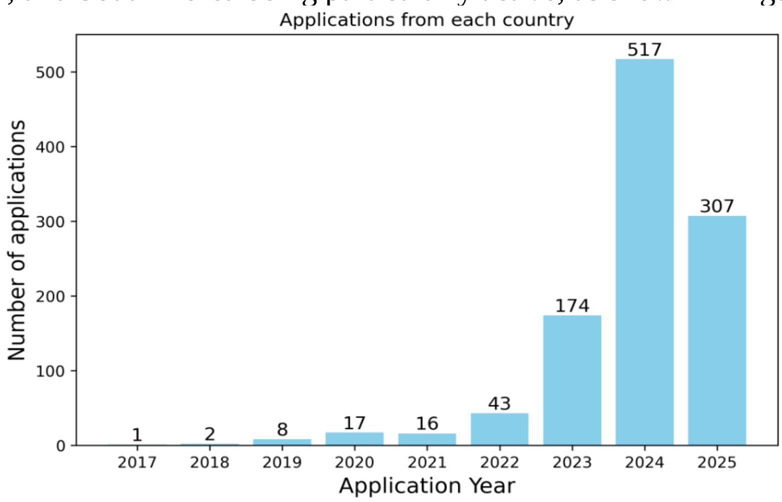


*Figure 1. Changes in patent applications year by year*

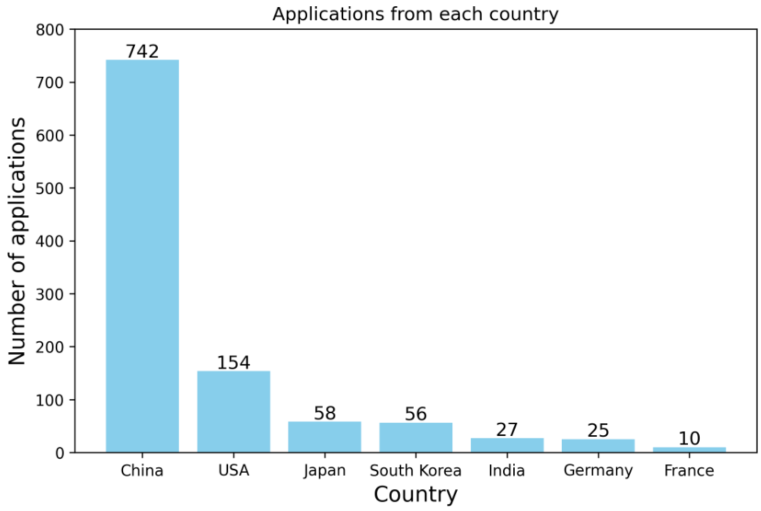


*Figure 2. Top 7 countries in terms of number of applications*

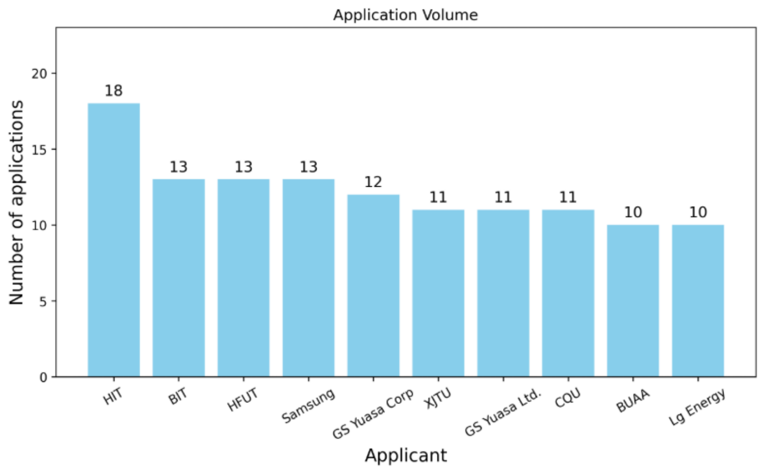


Figure 3. Top 10 applicants in terms of number of applications

From the statistical results, it is evident that with the continuous progress of lithium-ion battery technology, new research findings and innovations are constantly emerging. This has spurred companies and research institutions to proactively file patent applications to safeguard their intellectual property rights, thereby driving the growth in the number of patent applications. Among the top 10 applicants, colleges and research institutes are the main entities. Harbin Institute of Technology(HIT) Beijing Institute of Technology(BIT), Hefei University of Technology(HFUT), and Samsung have become the major applicants in this field, as shown in Figure 3. This indicates that there is still a long way to go before this technological innovation can be widely applied in enterprises.

## 3. Analysis of technology route evolution and patent layout

### 3.1. Technical Branch Division

The technical branches of patents related to the SOH of lithium-ion batteries encompass data-driven SOH prediction, artificial intelligence algorithms, battery management systems, expansion of battery application scenarios, and capacity attenuation models. These branches are interconnected and jointly drive the technological development in this field.

*TABLE I. Technical branches in patents*

| ***Keywords*** | ***IPC No.*** | ***Technology Association*** |
|---|---|---|
| Data-driven SOH prediction | G01R31/36<br>H01M10/48 | Data analysis of battery monitoring and management systems |
| AI algorithms | G06N3/00<br>G06N20/00 | Application of machine learning models in battery data |
| Battery Management System (BMS) | H01M10/48<br>G01R31/36<br>7 | BMS hardware design and health status integration |
| Battery application scenario expansion | B60L58/10<br>Y02E60/32 | On-board battery health management and clean energy technology |
| Capacity attenuation model | G01R31/36<br>G06F17/18 | Statistical capacity prediction and life modeling |

### 3.2. Patent layout hot spots and competition landscape

Based on the patent statistics and analysis results, within the realm of patents associated with the SOH of lithium-ion batteries, the battery monitoring technology (G01R31/36) stands out as the most active category. It constitutes approximately 40% of the total patent count, encompassing a diverse array of SOH evaluation algorithms. The machine learning applications (G06N3/00) follow closely, with a proportion of around 30%. In this category, numerous innovative algorithms, such as the long-short-term memory networks (LSTM) and Transformer models, are constantly emerging. These algorithms significantly contribute to enhancing the accuracy and efficiency of battery health monitoring. Evidently, the patent applications for lithium-battery SOH health-monitoring methods integrating AI algorithms currently top the list, making it the most popular research area in this field.

#### 3.2.1. Based on machine learning algorithms

The approach based on machine learning algorithms is designed to uncover the latent relationship between data features and the battery's health status through in-depth learning of extensive battery data. For instance, the support vector machine (SVM) algorithm makes full use of the charge-discharge data from multiple battery cycles by constructing and optimizing the SVM model. It extracts the inherent modal features from this data, thereby enabling accurate battery fault diagnosis and effective health status evaluation. Moreover, some research efforts employ clustering algorithms to categorize batteries. For batteries with different degradation patterns, customized improved convolutional neural networks (CNNs) are designed. This prediction method, which integrates manually extracted and self-extracted features, remarkably enhances the accuracy of battery cycle life prediction.

#### 3.2.2. Based on deep learning algorithms

Deep learning algorithms, owing to their robust feature-learning capabilities, are capable of conducting more profound analyses of battery data. Long Short-Term Memory networks (LSTMs) and their variants find extensive applications in the realm of battery health status prediction. These models can efficiently capture long-term dependencies within time-series data. When an LSTM-based model is constructed and key data like voltage and current during the battery's charging and discharging processes are inputted, the model can learn the temporal evolution patterns of battery aging characteristics, thus enabling accurate estimation of the battery's health status. Furthermore, the GCN-LSTM model, formed by integrating Graph Convolutional Networks (GCN) with LSTM, capitalizes on the GCN's proficiency in capturing local features and topological relationships during battery aging and the LSTM's aptitude for extracting temporal evolution patterns. This not only elevates the quality of feature data but also bolsters the model's generalization ability across diverse fast-charging strategies.

#### 3.2.3. Multi-source data fusion

The multi-source data fusion strategy aims to comprehensively utilize data from various sources to conduct a more thorough assessment of battery health status. Typically, the integrated data encompasses not only conventional parameters such as battery voltage, current, and temperature but also acoustic emission parameter data and cycle aging parameter data. Through gated neural networks, these multi-source data are assigned appropriate weights. Subsequently, by leveraging models like convolutional neural networks and long-short-term memory networks for training and prediction, it becomes feasible to fully uncover the valuable information embedded in different data sets, thereby effectively enhancing the prediction accuracy. For instance, acquiring the battery's state of charge and electrochemical impedance spectroscopy (EIS) data at different temperatures and utilizing these data to predict the remaining useful life of lithium batteries have yielded relatively satisfactory results.

#### 3.2.4. Model Optimization and Enhancement

To further enhance the performance of data-driven models, researchers continuously strive to optimize and refine these models. On one hand, optimization algorithms like the Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) are employed to fine-tune model parameters. By identifying the optimal parameter combination, the prediction accuracy of the model can be significantly improved. On the other hand, novel model architectures have been proposed, such as transformer-based models. These models leverage the self-attention mechanism, enabling them to handle long-sequence data more effectively. This mechanism allows the model to capture global dependencies within the data, which has led to excellent performance in battery health status prediction. Simultaneously, ensemble learning methods have gained widespread application. These methods involve fusing multiple models. For example, CatBoost can be integrated as the base learner within the NGBoost algorithm. By combining the strengths of both algorithms, the estimation accuracy of the SOH of lithium-ion batteries can be further enhanced.

### 3.3. Technology Route Integration Trend

In the realm of lithium-ion battery SOH estimation and prediction, the trend of technology route integration is becoming increasingly prominent, emerging as a pivotal driving force behind the development of this field. This integration endeavors to consolidate the merits of diverse technologies and surmount the constraints of single-technology approaches, thus enabling more precise and efficient SOH assessments.

From the algorithmic perspective, multi-algorithm fusion has emerged as a significant trend. Traditional model-based approaches, like electrochemical models, are capable of accurately mirroring the intricate physical and chemical processes within the battery. However, they are confronted with issues such as high computational costs and challenges in parameter acquisition. Conversely, data-driven methods, exemplified by long short-term memory networks (LSTM) grounded in deep learning, excel in processing large-scale data and uncovering latent data features.

Nevertheless, they lack in-depth comprehension of the battery's internal mechanisms.

Regarding data fusion, the trend of multi-source data integration is quite evident. During the operation of a battery, a wide variety of data are generated. This includes conventional data such as voltage, current, and temperature, along with acoustic emission parameter data, cycle aging parameter data, and electrochemical impedance spectroscopy (EIS) data. From the vantage point of model architecture, integrated learning, and multi-model fusion architectures have become focal points of research. To address the limitations of single models, the integration of multiple models for SOH estimation has garnered extensive attention.

With the advancement of technologies like the Internet of Things and cloud computing, the integration of technical routes is also manifested in the convergence of cross-domain technologies. Cloud-based multi-source data-driven battery health management is gradually gaining traction. By uploading the operational data of batteries from diverse scenarios to the cloud, leveraging the cloud's robust computing and storage capabilities, integrating multi-source data, and applying a plethora of algorithms for analysis and processing, it becomes possible to achieve real-time and precise monitoring and management of battery health status.

## 4. Technology Trend Forecast and Challenges

Combining industry development trends, market demands, and the technology R & D cycle, the technological roadmap of lithium-battery State-of-Health (SOH) monitoring technology is predicted using patent big-data analysis. As artificial intelligence technology continues to advance, over the next 5-10 years, the AI-based lithium-battery SOH monitoring technology will mature further and integrate more deeply with other technologies. This will not only enhance the prediction accuracy of relevant models but also boost efficiency and simplify model complexity, as shown in Figure 4.

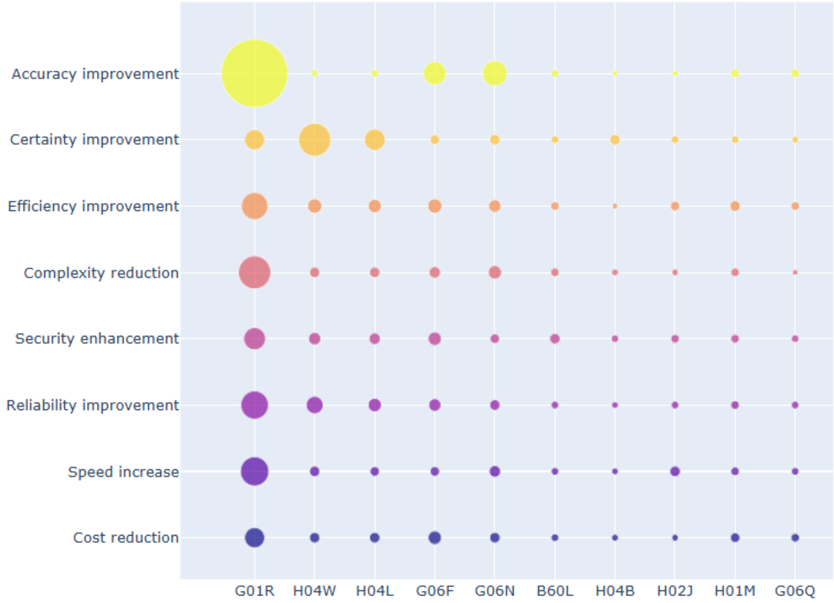


*Figure 4. Technical Efficacy Focus*

### 4.1. Accuracy

The internal dynamics of a lithium battery are highly complex. The stochastic nature of these internal changes presents significant obstacles to accurate prediction. For example, the diffusion of lithium-ions within the electrolyte, the chemical reactions taking place on the electrode surfaces, and the degradation of active materials all occur randomly. This randomness makes it arduous to precisely anticipate the battery's state. Nevertheless, AI approaches can effectively surmount these challenges. AI algorithms, such as neural networks, possess the capacity to glean complex patterns from extensive datasets. They are capable of analyzing the relationships among diverse battery parameters, including voltage, current, temperature, and historical usage data, to discern the underlying trends in the battery's internal changes. Even when confronted with the unpredictability of internal reactions, AI can detect latent patterns and generate more precise predictions.

Furthermore, the enhancement of prediction accuracy is intricately linked to the quality of the data fed into the system. High-quality data encompasses comprehensive and exact information regarding the battery's operation. It encompasses a broad spectrum of data sourced from various operating conditions, charging-discharging cycles, and environmental factors. In the case of poor-quality data, which may contain missing values, noise, or inaccurate measurements, the AI model may struggle to learn the correct patterns. This, in turn, can result in inaccurate predictions. Conversely, high-quality data enables the AI model to be trained more efficiently, leading to improved generalization capabilities and heightened prediction accuracy. Thus, to fully harness the potential of AI in lithium-battery SOH monitoring, ensuring the quality of the input data is of utmost importance.

### 4.2. Efficiency

Improving the prediction efficiency of the model while maintaining accuracy is very critical in practical applications. To improve the efficiency of the AI-based battery health state (SOH) prediction method, the first is to optimize data processing, use technologies such as median filters to remove noise, use PCA to compress large-scale data and use GANs to generate synthetic data to enhance training effects. The second is algorithm optimization, selecting appropriate models according to data and tasks, such as using LSTM to process sequence data and GNNs to process complex structure data, pruning large models and using knowledge distillation technology, and using Bayesian optimization algorithms instead of traditional grid searches for hyperparameter tuning. The third is computing facility and deployment optimization, using GPUs and TPUs for hardware acceleration, migrating training and reasoning tasks to this; performing edge computing at the data source end to reduce data transmission delays and bandwidth consumption and improve prediction efficiency.

### 4.3. Simplify Model Complexity

Simplify the model structure by choosing an appropriate basic model by the data characteristics and task requirements. For instance, when dealing with data that exhibits time-series features, a straightforward autoregressive (AR) or moving average (MA) model can efficiently process the data. This not only helps avoid overfitting issues but also reduces computational costs. Regarding complex models, pruning and compressing them is essential. We can utilize L1 and L2 regularization techniques to identify and eliminate unimportant connection weights. Additionally, model quantization can be carried out to reduce memory usage and computational load. Another viable option is to adopt an integrated learning strategy, like a random forest which is composed of multiple decision trees. By aggregating the results of these individual trees, the random forest enhances the overall prediction ability and offers strong interpretability, making it more conducive to practical applications. Select an adaptive learning rate algorithm, such as Adagrad or Adam, for training optimization. These algorithms can automatically adjust the learning rate based on the gradient changes during the training process, thus accelerating the model's convergence. For example, in the training of lithium-battery SOH prediction models, the Adam algorithm can significantly improve the training efficiency. Moreover, an early termination strategy should be employed. When the loss function value of the validation set shows no further decrease within a specific number of rounds, the training process is terminated. This effectively prevents overfitting and saves computational resources and time.

## 5. Conclusion

Based on the IPC classification number and specific search formula, we precisely retrieved the relevant patents regarding Li-ion Batteries Health Estimation Technology from the patent database. Subsequently, we conducted an in-depth analysis of the macro-level distribution of patent applications. Moreover, by leveraging the patent data, we systematically sorted out the technological branches and probed into their evolutionary trajectories. Finally, through these patent data, we forecasted the future technological development trends in this field and analyzed the potential challenges. These research findings are of crucial importance to relevant research institutions. In the future, we plan to analyze this field from the perspective of patent citation relationships and by integrating non-patent data.

## Acknowledgment

The present work was supported by Quzhou Science and Technology Key Research Project (No.2023K252, 2023K248).

## References

*[1] Wang, Yuxing, et al. "Lithium and lithium-ion batteries for applications in microelectronic devices: A review." Journal of Power Sources 286 (2015): 330-345.*
*[2] Patraboy, Timothy J., Michael D. Farrington, and George J. Donaldson. "Reliability of lithium batteries in search and rescue beacons." Journal of power sources 65.1-2 (1997): 93-99.*
*[3] Zheng, Yuejiu, et al. "A capacity prediction framework for lithium-ion batteries using fusion prediction of empirical model and data-driven method." Energy 237 (2021): 121556.*
*[4] Sun, Fengchun, et al. "Model-based dynamic multi-parameter method for peak power estimation of lithium-ion batteries." Applied Energy 96 (2012): 378-386.*
*[5] Ma, Bo, Wei Guo, and Zhaojun Steven Li. "Prediction of lithium-ion battery capacity based on the arma method." 2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai). IEEE, 2020.*
*[6] Xu, Huanwei, et al. "An improved CNN-LSTM model-based state-of-health estimation approach for lithium-ion batteries." Energy 276 (2023): 127585.*
*[7] Gu, Xinyu, et al. "A time-series Wasserstein GAN method for state-of-charge estimation of lithium-ion batteries." Journal of Power Sources 581 (2023): 233472.*
*[8] Fang X, Chen Z, Li J Q. Optimized multi-head self-attention mechanism for SOH prediction in lithium-ion battery energy storage systems within microgrids[J]. Electric Power Systems Research, 2025, 238: 111182.*
*[9] Xu, Liqianyun, et al. "Data-driven-aided strategies in battery lifecycle management: prediction, monitoring, and optimization." Energy Storage Materials 59 (2023): 102785.*
*[10] Aaldering, Lukas Jan, and Chie Hoon Song. "Tracing the technological development trajectory in post-lithium-ion battery technologies: A patent-based approach." Journal of Cleaner Production 241 (2019): 118343.*

*[11] Wang, Huilong, et al. "Understanding technological innovation and evolution of energy storage in China: Spatial differentiation of innovations in lithium-ion battery industry." Journal of Energy Storage 66 (2023): 107307.*
*[12] Xiong, Qian, and Tang, Wenzhe. "Text mining-based patent analysis on informatization for construction management of hydraulic engineering projects." Journal of Tsinghua University (Science and Technology) 63.2 (2023): 223-232.*