Celal Bayar University Journal of Science

Prediction Of The Remaining Useful Life Of Lithium-Ion Batteries Based On An Empirical Mode Approach With Artificial Neural Networks

Ozancan Bayrı[1], Sıtkı Akkaya[2]*

[1] Institute Of Graduate Studies Defence Technology Programme, Sivas University of Science and Technology, Sivas, Türkiye
[2] Department of Electrical and Electronics Engineering, Sivas University of Science and Technology, Sivas, Türkiye
*sakkaya@sivas.edu.tr
* Orcid No: 0000-0002-3257-7838

Received: 31 January 2024
Accepted: 27 March 2024
DOI: 10.18466/cbayarfbe.1429043

Abstract

Forecasting future capacities and estimating the remaining useful life, while incorporating uncertainty quantification, poses a crucial yet formidable challenge in the realm of battery health diagnosis and management. In this study, a data-driven model based on artificial neural networks (ANN) and signal decomposition techniques including Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), and Empirical Wavelet Transform (EWT) is presented to predict the capacity value of lithium-ion batteries. Signal decomposition was performed using the discharge voltage values for four different batteries. A total of 22 features were obtained. The features of the signal decomposition methods were evaluated separately as well as hybrid approaches. Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) performance metrics are used in the proposed method and the values obtained are 3.67×10-6, 0.001351 and 0.002311, respectively. According to the findings, the hybrid model proposed demonstrated positive results in terms of accuracy, adaptability, and robustness.

Keywords: Lithium-ion batteries, Signal decomposition, Artificial neural networks, Prediction

1. Introduction

Lithium-ion batteries stand out among the secondary batteries known for their remarkable energy density, and they find widespread application in today's energy storage landscape [1]. Their prominence extends to diverse sectors, including transportation, communication, aviation, and military defense, owing to their advantageous attributes such as lightweight design, robust safety features, and elevated voltage levels in comparison to alternative battery technologies [2]. Notably, as the electric vehicle market continues to expand and the prevalence of mobile phones and personal mobility devices rises, the prudent management of energy resources and the establishment of dependable, enduring battery solutions have assumed heightened significance [3].

Against this backdrop, the imperative to assess battery health becomes paramount. A pivotal threshold in this evaluation lies in the battery's capacity, where a decline below 80% signifies the completion of its operational lifespan. Critical factors that impinge upon battery longevity encompass charge-discharge cycles, overcharging, discharging under strenuous conditions, and exposure to high currents. Anticipating battery longevity carries profound implications, underpinning efficient resource allocation, stable management practices, and the uninterrupted operation of devices.

Estimation techniques for assessing the Remaining Useful Life (RUL) of lithium-ion batteries can be broadly categorized into model-based, data-driven, and hybrid methods that combine aspects of both. Among the model-based approaches, two prominent methods are the utilization of electrochemical models and equivalent circuit models. However, within the realm of model-based techniques, the challenge lies in acquiring a suitable model that can effectively encapsulate the intricate dynamic behaviors of the system. This challenge arises from the intricate physical and chemical intricacies

inherent to batteries, coupled with the presence of numerous parameters that exert influence upon battery performance parameters such as charge/discharge currents, voltage, temperature, and internal impedance [4].

Forecasting the remaining lifespan of lithium-ion batteries through data-driven techniques proves to be notably more straightforward compared to model-based approaches. This avenue circumvents the necessity for specialized expertise, offering enhanced flexibility and practicality. Within data-driven methodologies, the demand for explicit mathematical formulations to articulate the deterioration trajectory of batteries over time, as seen in model-based methodologies, is obviated. Instead, this approach hinges on the foundation of historical degradation data, harnessing machine learning algorithms like Support Vector Machines (SVM), Logistic Regression, and Artificial Neural Networks (ANN) for prognosticating the RUL of lithium-ion batteries [5]. To illustrate, exemplar studies conducted with data-driven strategies further underscore the viability of this approach.

The data-driven approach disregards the intricacies of the battery's internal reactions and parameters, instead focusing on extracting vital insights from the battery's historical data to assess its capacity degradation pattern [6]. Unlike traditional approaches that rely on specific physical models, data-driven methods leverage extensive historical battery measurement data to construct pertinent models through the use of machine learning techniques. This inherent flexibility and versatility make these methods highly promising for various applications [7]. The crux of data-driven methods lies in the careful selection and processing of data to establish the intricate relationships between parameters and State of Health (SOH). Commonly employed data-driven techniques encompass Neural Networks (NNs) [8], Relevance Vector Machines [9], Autoregressive Moving Average models [10], Support Vector Regression (SVR) [11], and SVM [12]. For instance, Li et al. [13] employed SVR to predict SOH for two batteries under different aging conditions, implementing filtering technology to refine capacity curves and enhance model accuracy. Meanwhile, Fei et al. [14] meticulously crafted six machine learning models to forecast battery life, conducting a thorough comparison of each model's distinctive characteristics. Notably, battery prediction models founded on neural networks have recently garnered increasing interest within the research community.

However, there is room for improvement in current forecasting models. Signal decomposition methods are typically applied in isolation, and the preference for a hybrid approach is underrepresented in the literature. In this paper, we introduce a novel RUL forecasting approach based on neural networks, which leverages hybrid decomposition methods.

This study introduces a data-driven model that employs artificial neural networks, built upon the foundations of Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), and Empirical Wavelet Transform (EWT) signal processing techniques. The objective is to anticipate the capacity levels of lithium-ion batteries. The efficacy of this approach was evaluated using a dataset sourced from the NASA Ames Prognostics Center of Excellence (PCoE), focusing on lithium-ion batteries [15]. Contributions of this study are as follows.

- Expanded feature set with 22 features in total obtained from residues of EMD, EEMD, EWT, and different combinations of them.
- Then the hybrid models are composed of Multilayer Perceptron (MLP) and different combinations for different scenarios.
- Performance Comparison for different mentioned hybrid models.
- Performances Comparison for different batteries comparatively.
- The best results in terms of mean values for the EMD-EEMD model
- More effective performance than similar studies

The rest of the paper is organized as follows. Section 2 describes in detail the EMD, EEMD, EWT, and MLP methods used in the proposed hybrid forecasting model. In Section 3, battery capacity forecasting is performed with the proposed hybrid model and the prediction results and model performance are presented. The last section summarizes the results of the study.

2. Materials and Methods

2.1. Dataset

The data for this study originates from the NASA Prognostic Center of Excellence (PCoE). Detail of the data acquisition is given in Table 1 as follows. The raw dataset from NASA has a range of datasets for different battery types with different time ranges. Also, most of the time series (T.S.) capacity signals have some missing data. For these causes, similar to the other studies on lithium-ion batteries, B5, B6, B7, and B18 T.S. capacity signals which are more stable than the others are selected. We employed four distinct sets of NASA Lithium-Ion Batteries denoted as B0005, B0006, B0007, and B0018, each boasting a 2 *Ah* rated capacity. These batteries underwent a meticulous testing regimen conducted at room temperature. Initially, the charging process involved a constant current of 1.5 *A*, followed by charging under a constant voltage of 4.2 *V*. Subsequently, discharging commenced with a consistent 2 *A* current until the voltage levels for B0005, B0006, B0007, and B0018 reached 2.7 *V*, 2.5 *V*, 2.2 *V*, and 2.5 *V*,

respectively. The experiment was concluded once the actual battery capacity fell below a predetermined usage threshold [15, 16]. Visualization of the capacity degradation curves for B0005, B0006, B0007, and B0018 is encapsulated in Figure 1.

Table 1. Data Acquisition details

Number of Lithium-ion batteries	18,650
Manufacture	LG Chem
Chemistry	Lithium cobalt
Number of cells	28
Nominal capacity	2.10 Ah
Capacity range	2.10 Ah → 0.80 Ah
Voltage range	4.2–3.2 V
Cycling protocols	7
Sampling Frequency	10 Hz

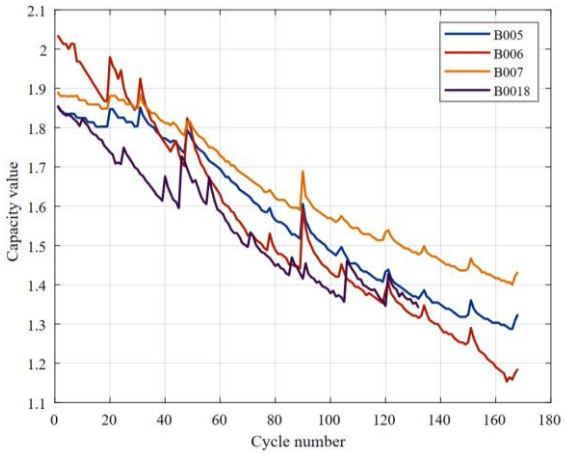

Figure 1. NASA Lithium-ion battery capacity degradation curves

2.2. Data pre-processing

Signal preprocessing refers to an important stage in the analysis of signals, where raw data undergoes a series of computational techniques aimed at improving its quality, extracting relevant features, and reducing unwanted artifacts or noise. This preparatory phase is fundamental in various fields such as signal processing, communications, and data analysis, as it forms the basis for subsequent analytical tasks. Data interpolation is one of the common preprocessing methods. Therefore, the interpolation method is applied to each T. S. signal to equalize each other in each cycle and to analyze with signal processing methods easily.

2.3. Empirical mode decomposition (EMD)

Empirical Mode Decomposition (EMD) stands as a versatile multiresolution technique, purpose-built to dissect data and unveil its intricate constituents. EMD operates as a powerful tool for scrutinizing non-linear and non-stationary signals, unraveling their complexities by fragmenting them into distinct layers of varying resolutions. The core tenet of EMD involves the meticulous breakdown of temporal sequences into a compact set of intrinsic modes, each accompanied by a residual remainder. These distinctive modes, aptly dubbed Intrinsic Mode Functions (IMFs), encapsulate the essence of the signal's dynamic behavior [17].

The EMD methodology follows a systematic course: initially, it identifies local extrema peppered throughout the temporal sequence. Correspondingly, it constructs upper and lower envelopes, employing interpolation techniques like cubic splines. Subsequently, the average of these envelopes is differentially subtracted from the original signal, birthing a local intrinsic mode function. This twofold process iterates until certain conditions converge:
The mean of the upper and lower envelopes converges toward zero. The count of extremal points and zero crossings diverge by no more than one. This iterative protocol, known as the "elimination" procedure, culminates in the generation of a finite array of IMFs, complemented by a residual component, illustrated in Equation (1):

$$x(t) = \sum_{i=1}^{N} IMF_i(t) + R_N(t) \quad \textbf{(1)}$$

Where $IMF_i(t)$ represents all potential IMFs, N stands for the count of generated IMFs, and $R_N(t)$ signifies the residual arising from the decomposition.

2.4. Ensemble empirical mode decomposition (EEMD)

The EEMD method is a noise-assisted technique that aims to reduce mode mixing and improve the separation of oscillatory modes. By adding white noise to the input signal, the EEMD algorithm generates a set of IMFs that have statistical properties that are closer to the true intrinsic modes of the signal. The resulting ensemble of IMFs obtained from multiple noise realizations improves the signal-to-noise ratio and helps in the extraction of the underlying signal components [18].

The intricacies of EEMD are outlined in a sequential procedure: the EEMD steps commence with the addition of white noise possessing specific amplitude and frequency attributes to the input signal. Following this, the noisily augmented signal undergoes decomposition through the EMD algorithm, yielding IMFs. EMD, a data-centric methodology, dissects signals into oscillatory IMFs, each characterized by a well-defined frequency range. This process of addition and decomposition is iteratively undertaken numerous times

with distinct noise realizations. Subsequently, the IMFs derived from each iteration are amalgamated and averaged to form an ensemble. This ensemble then undergoes further decomposition via EMD to extract the ultimate IMFs, which faithfully portray the underlying authentic signal. Finally, the residual of the original input signal is computed by subtracting these final IMFs from the initial signal [19].

2.5. Empirical wavelet transform (EWT)

The Empirical Wavelet Transform emerges as a synergy of the wavelet transform and EMD techniques, presenting an instrumental signal decomposition approach that effectively disentangles a provided signal into distinct modes. From a Fourier perspective, this framework aligns with the architecture of a bandpass filter setup, ultimately grounded in the formulation of an adaptive wavelet filter bank [20]. This method achieves its functionality through the creation of adaptive wavelets, dynamically conforming to the informational essence embedded within the signal.
The initial phase of EWT entails an estimation of the signal's frequency components, subsequently leading to the computation of boundaries which, in turn, facilitates the extraction of diverse signal modes based on these determined thresholds. When engaging the EWT method, the initial step involves the reception of signals followed by the computation of the Hurst Exponent value as a measure of long-term memory of time series. Through the application of specified thresholds, signal frequencies are meticulously derived via mathematical computations within the Fourier spectrum, thereby establishing precise boundaries crucial for subsequent phases. With these identified boundaries at hand, the procedure advances to the creation of N wavelet filter banks, followed by the implementation of pertinent mathematical computations employing requisite formulas. This intricate process culminates in the extraction of distinct frequency bands, executed through meticulous filtering mechanisms [21].

Positioned as a swift and incredibly adaptive technique within the realm of signal analysis, the Discrete Wavelet Transform method stands analogous to EWT but boasts a robust and comprehensive mathematical foundation. This mathematical robustness ensures that both the scaling function and wavelets seamlessly conform to the information embedded within the analyzed signal, obviating the necessity for any a priori knowledge of the signal [22].

2.6. Feature Extraction

Feature extraction involves selecting specific characteristics from a dataset to effectively represent it. This process plays a crucial role in addressing classification and regression challenges. This study focuses on estimating battery capacity by utilizing voltage, current, temperature, and time data, along with attributes derived from various transformation methods. The study employs a total of 22 features, organized into four distinct sets. The first set (attributes 1-10) comprises parameters extracted directly from the dataset without any signal processing. It is shown in Table 2. The remaining sets encompass attributes derived from residual signals resulting from EMD, EEMD, and EWT methods applied sequentially to discharge voltage values.

When using the attributes in the above table, the following mathematical expressions were used.

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|x_n|^2} \quad \textbf{(2)}$$

$$x_{average} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}x_n} \quad \textbf{(3)}$$

2.7. Artificial neural networks

Artificial Neural Networks are one of the most widely used supervised learning methods for regression and classification problems. They are a machine learning approach inspired by biological nervous systems. ANNs are employed for various tasks, such as processing large datasets, recognizing complex patterns, and making predictions. The Multilayer Perceptron (MLP) is a widely preferred type of ANN and consists of three main layers: input, hidden, and output. The input layer contains the attributes necessary to achieve the targeted result. The hidden layer processes the information through neurons, and the output layer predicts continuous values [23]. The architecture used in the study is shown in Figure 2.

Table 2. Created set groups and features.

	Number	Feature Name	Description
	1	MaxVm	Maximum voltage at discharge
	2	MinVm	Minimum voltage
	3	AverageVm	Average voltage
	4	Vtime	Time of occurrence for minimum voltage
Set I	**5**	MinCurrent	Lowest current value
	6	CurrentTime	Time of occurrence for lowest current
	7	MaxTemperature	Highest temperature
	8	MinTemperature	Lowest temperature
	9	AverageTemperature	Average temperature

	10	TempTime	Time of occurrence for the highest temperature
Set II	**11**	Max_Vmres_Emd	Max, min, average, and RMS values of residue signal obtained by taking EMD components of discharge voltage values
	12	Min_Vmres_Emd	
	13	Average_Vmres_Emd	
	14	RMS_Vmres_Emd	
Set III	**15**	Max_Vmres_Eemd	Max, min, average, and RMS values of residue signal obtained by taking EEMD components of discharge voltage values
	16	Min_Vmres_Eemd	
	17	Avarage_Vmres_Eemd	
	18	RMS_Vmres_Eemd	
Set IV	**19**	Max_Vmres_Ewt	Max, min, average, and RMS values of residue signal obtained by taking EWT components of discharge voltage values
	20	Min_Vmres_Ewt	
	21	Avarage_Vmres_Ewt	
	22	RMS_Vmres_Ewt	

2.7. Artificial neural networks

Artificial Neural Networks are one of the most widely used supervised learning methods for regression and classification problems. They are a machine learning approach inspired by biological nervous systems. ANNs are employed for various tasks, such as processing large datasets, recognizing complex patterns, and making predictions. The Multilayer Perceptron (MLP) is a widely preferred type of ANN and consists of three main layers: input, hidden, and output. The input layer contains the attributes necessary to achieve the targeted result. The hidden layer processes the information through neurons, and the output layer predicts continuous values [23]. The architecture used in the study is shown in Figure 2.

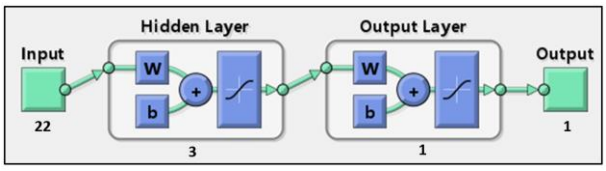

Figure 2. ANN architecture used in the study.

The basic unit of an artificial neural network is called an artificial neural cell or neuron. A neuron receives *input* signals, processes these signals, and produces an *output* signal. The basic mathematical function of a neuron is represented by Equation 4. The connections between neurons are represented by *weights*, and these weights are adjusted during the learning process [24].

$$Output = Activation\ function \times (weights \times inputs + bias) \quad \textbf{(4)}$$

The MLP conducts the learning process by assigning weights to the connections between the layers during training. This is achieved by utilizing the dataset and comparing the network's predictions with the actual results. A loss function quantifies the degree of agreement between the network's predictions and the actual results. The weights are subsequently updated using a backpropagation algorithm. This iterative process enables the network to improve its predictions [25].

The MLP conducts the learning process by assigning weights to the connections between the layers during training. This is achieved by utilizing the dataset and comparing the network's predictions with the actual results. A loss function quantifies the degree of agreement between the network's predictions and the actual results. The weights are subsequently updated using a backpropagation algorithm. This iterative process enables the network to improve its predictions [25].

2.8. Performance evaluation

In this study similar to other studies on RUL prediction, performance evaluations stemming from capacity value estimations were conducted using widely employed regression metrics: Mean Square Error (*MSE*), Mean Absolute Percentage Error (*MAPE*), and Root Mean Square Error (*RMSE*). The mathematical formulations for these performance metrics are provided below.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad \textbf{(5)}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \quad \textbf{(6)}$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(y_i - \hat{y}_i)^2}{N}} \quad \textbf{(7)}$$

Where y is the actual battery value, $\hat{y}$ is the prediction value and N is the number of predicted data.

2.9 Proposed methodology

In order to evaluate the efficacy of the proposed hybrid prediction algorithm, a series of methods were implemented. These methods detailed within this paper pertain to forecasting the future capacity of batteries. The precise procedural steps for these methodologies can be found in Figure 3, and they are elaborated upon in the following section.

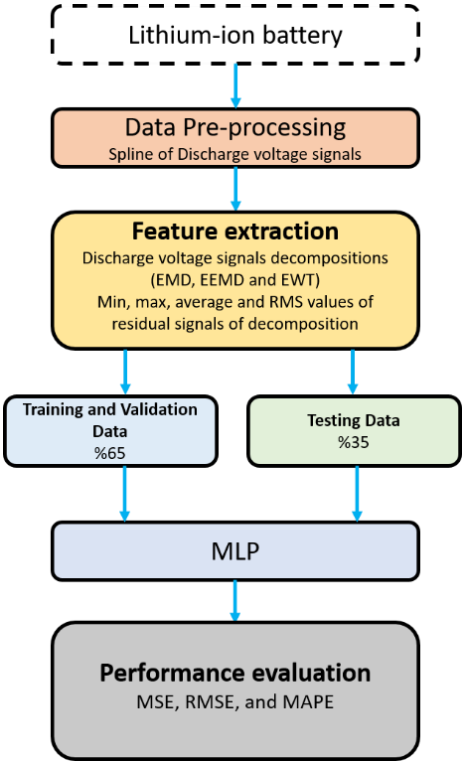

Figure 3. Proposed capacity estimation methodology.

Table 3. Feature sets used in the study.

Scenario number	Methods	Features sets
I	EMD	Set I and Set II
	EEMD	Set I and Set III
	EWT	Set I and Set IV
II	EMD-EEMD	Set I, Set II, and Set III
	EMD-EWT	Set I, Set II, and Set IV
	EEMD-EWT	Set I, Set III, and Set IV
	EMD-EEMD-EWT	Set I, Set II, Set III, and Set IV

In this study, data from NASA's battery database was utilized, specifically focusing on batteries labeled B0005, B0006, B0007, and B0018. The dataset encompasses various parameters in both charge and discharge states, including temperature, current, and voltage values during the discharge phase. To standardize the data before the feature extraction process, the spline command in MATLAB was employed to ensure an equal number of data points for each cycle.

As demonstrated in Table 3, four distinct feature sets were generated using signal processing techniques, including EMD, EEMD, and EWT. The study can be divided into two fundamental sections, each examining distinct strategies.

The common feature set used in both parts of the study is denoted as Set I. In Scenario I, each method is individually analyzed, facilitating a comparative discussion of their performance. Meanwhile, in Scenario II, the study delves into the impact of combining these methods, exploring the performance of a hybrid system encompassing all three approaches.

In the study focused on estimating capacity values, the data were partitioned into two distinct groups: training and test sets. The training dataset comprised roughly the initial 65%, while the test dataset encompassed approximately the remaining 35%. Specifically, for batteries B0005, B0006, and B0007, the initial 110 data points were designated as the training data, while the subsequent 58 data points were allocated to the test set. As for battery B0018, the first 86 data points were assigned to the training set, with the last 46 data points designated for the test set.

In this study, MLPs are utilized to predict battery capacity. The MLP architecture consists of a single hidden layer with 3 neurons. After conducting various tests, it was determined that this architecture produced the best results. The Levenberg-Marquardt algorithm was employed in the training process, as it is commonly used in the literature and is known for its speed compared to other algorithms. The training and test phases were conducted five times, and performance evaluation was based on average values.

Table 4. Options of Proposed MLP Model

Options	Abbreviation in MATLAB	Value
Maximum number of epochs	MaxEpochs	30
Size of mini-batch	MiniBatchSize	128
Option for data shuffling	Shuffle	once
Number of epochs for dropping the learning rate	LearnRateDropPeriod	10
Factor for dropping the learning rate	LearnRateDropFactor	0.1
Contribution of previous step	Momentum	0.9
Decay rate of gradient moving average	GradientDecayFactor	0.9
Denominator offset	Epsilon	1e-8
Maximum number of iterations	MaxIterations	1000
Frequency of neural network validation	ValidationFrequency	50
Factor for L2 regularization	L2Regularization	0.0001
Mode to evaluate statistics in batch normalization layers	BatchNormalizationStatistics	moving
Gradient threshold	GradientThreshold	Inf
Gradient threshold method	GradientThresholdMethod	l2norm
Option to pad or truncate sequences	SequenceLength	longest
Direction of padding or truncation	SequencePaddingDirection	right
Value to pad sequences	SequencePaddingValue	0

The options of the MLP consist of performance to normalization criteria are set as default and details are given in Table 4.

3. Results and discussion

In this section, we present the experimental results for battery test sets B0005, B0006, B0007, and B0018, respectively. The features in Table 2 extracted using EMD, EEMD, and EWT signal decomposition techniques were utilized to predict the capacity values over the battery's lifetime through MLP neural networks. The methodology proposed in Figure 3 was employed, and the performance was assessed using MSE, MAPE, and RMSE metrics. The analysis in this study was conducted using the Matlab programming language.

Analyses were conducted for four different batteries, and the results are depicted in Figures 4-7 and Tables 3-7. Figure 4 and Table 5 display the prediction results for battery B0005. The models were analyzed individually, as well as in their binary and ternary combinations. It can be observed from the figure and the values in the table that all methods yielded results close to the capacity value. When analyzing the methods individually, EEMD produced the best results, while the error values were the lowest in the EMD-EEMD-EWT combination. Figure 5 and Table 6 display the prediction values for Battery 6. It is evident that the predicted values from the models closely match the actual battery values. Notably, these predictions exhibit a strong alignment with the actual values during the training phase and the initial segments of the test phase; however, they start to diverge noticeably after the 150th data point. In contrast to Battery 5, for this particular battery, the model based on the EEMD method outperformed other single models and binary-ternary combinations. Figure 6 and Table 7 contain the predictions for battery 7. Among the individual models, EWT appears to perform the best, exhibiting the lowest MSE, MAPE, and RMSE values, which signify its accuracy in predicting battery capacity values. When considering binary combinations of models, EMD-EEMD-EWT emerges as the top-performing combination, boasting the lowest MSE, MAPE, and RMSE values among the binary pairs. This suggests that integrating EMD, EEMD, and EWT decomposition techniques into a single model yields the most precise predictions for battery capacity values. Therefore, in the context of this dataset and analysis, the EMD-EEMD-EWT combination model demonstrates superior predictive capabilities compared to the individual models and other binary combinations.

Table 5. Performance evaluations of B0005 battery.

Model	MSE	MAPE	RMSE
EMD	5.29×10^{-6}	0.001434	0.002301
EEMD	3.70×10^{-6}	0.001163	0.001922
EWT	8.48×10^{-6}	0.001810	0.002913
EMD-EEMD	4.83×10^{-6}	0.001346	0.002199
EMD-EWT	8.70×10^{-6}	0.001841	0.002949
EEMD-EWT	7.96×10^{-6}	0.001734	0.002822
EMD-EEMD-EWT	1.35×10^{-6}	0.000703	0.001160

Celal Bayar University Journal of Science
Volume 20, Issue 2, 2024, p 1-13
Doi: 10.18466/cbayarfbe.1429043

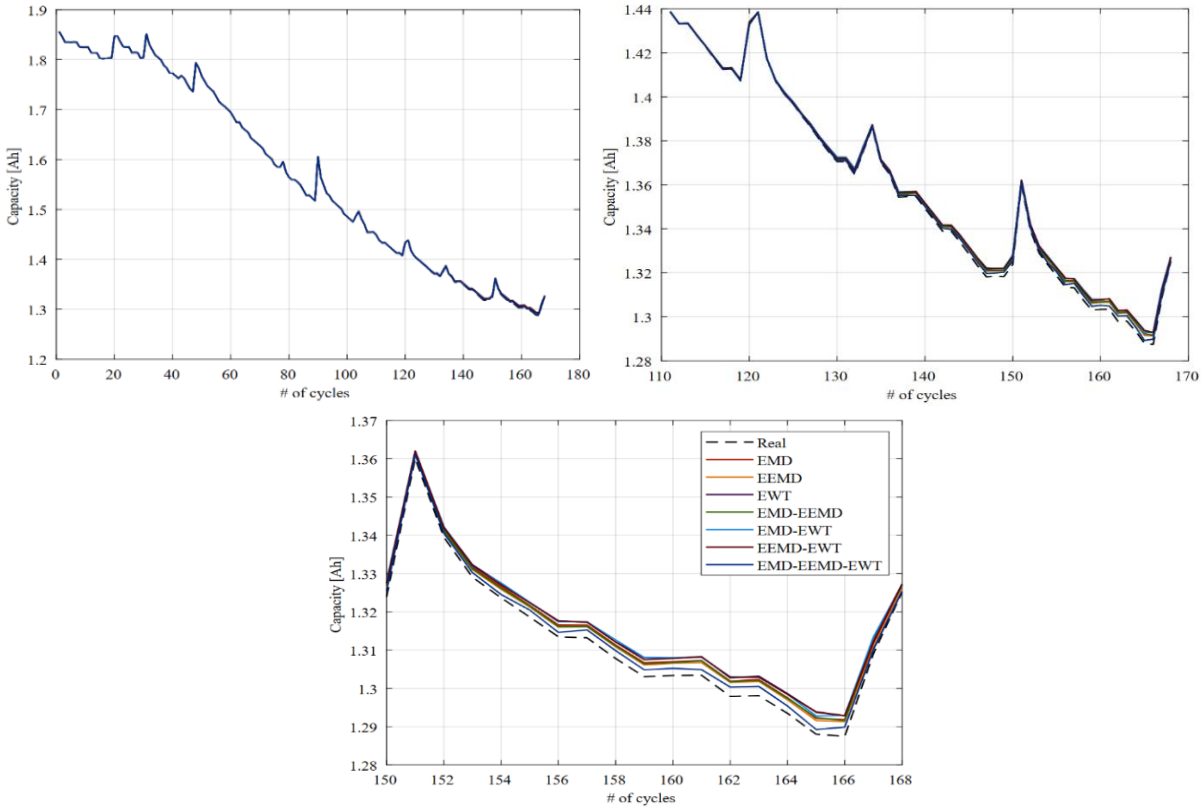

Figure 4. B0005 battery prediction results.

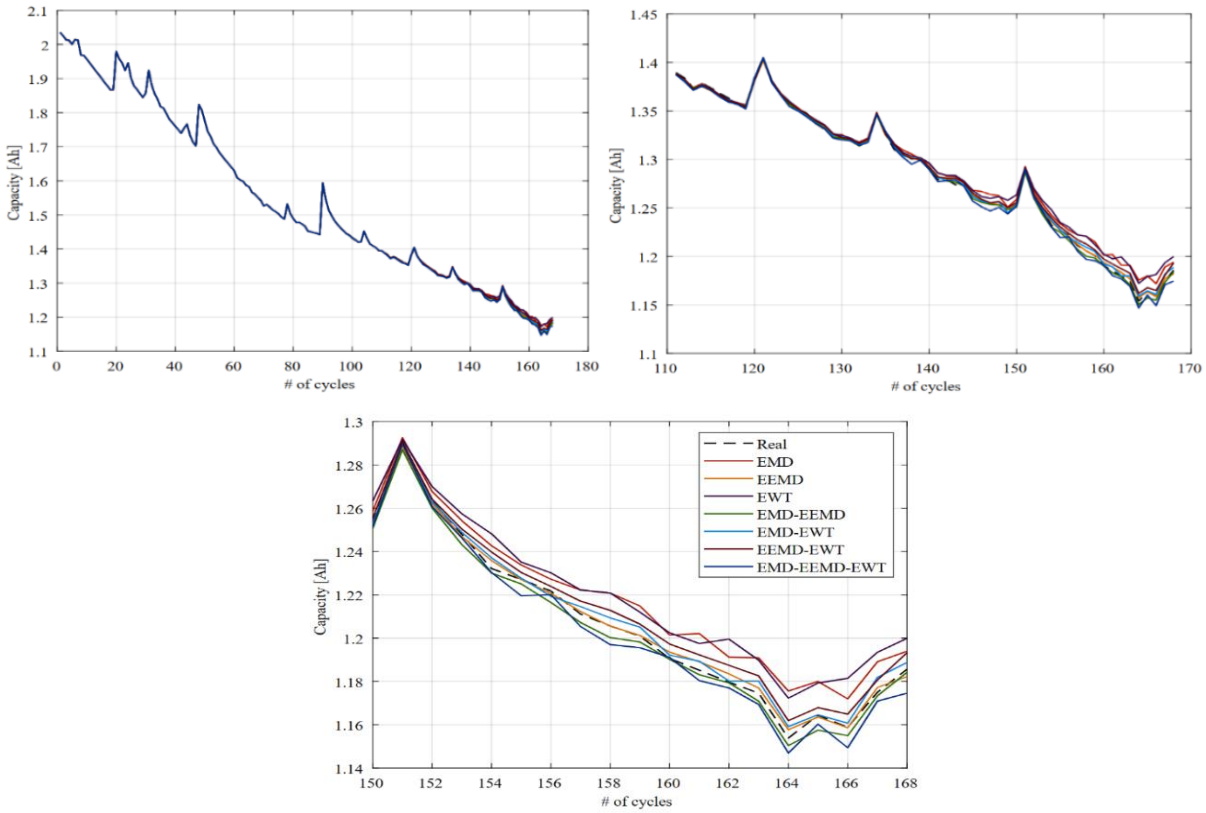

Figure 5. B0006 battery prediction results.

Table 6. Performance evaluations of B0006 battery.

Model	MSE	MAPE	RMSE
EMD	5.96×10^{-5}	0.004619	0.007721
EEMD	4.34×10^{-6}	0.001254	0.002083
EWT	7.31×10^{-5}	0.005011	0.008548
EMD-EEMD	6.41×10^{-6}	0.001646	0.002533
EMD-EWT	6.48×10^{-6}	0.001612	0.002546
EEMD-EWT	1.39×10^{-5}	0.002235	0.003729
EMD-EEMD-EWT	1.63×10^{-5}	0.002574	0.004041

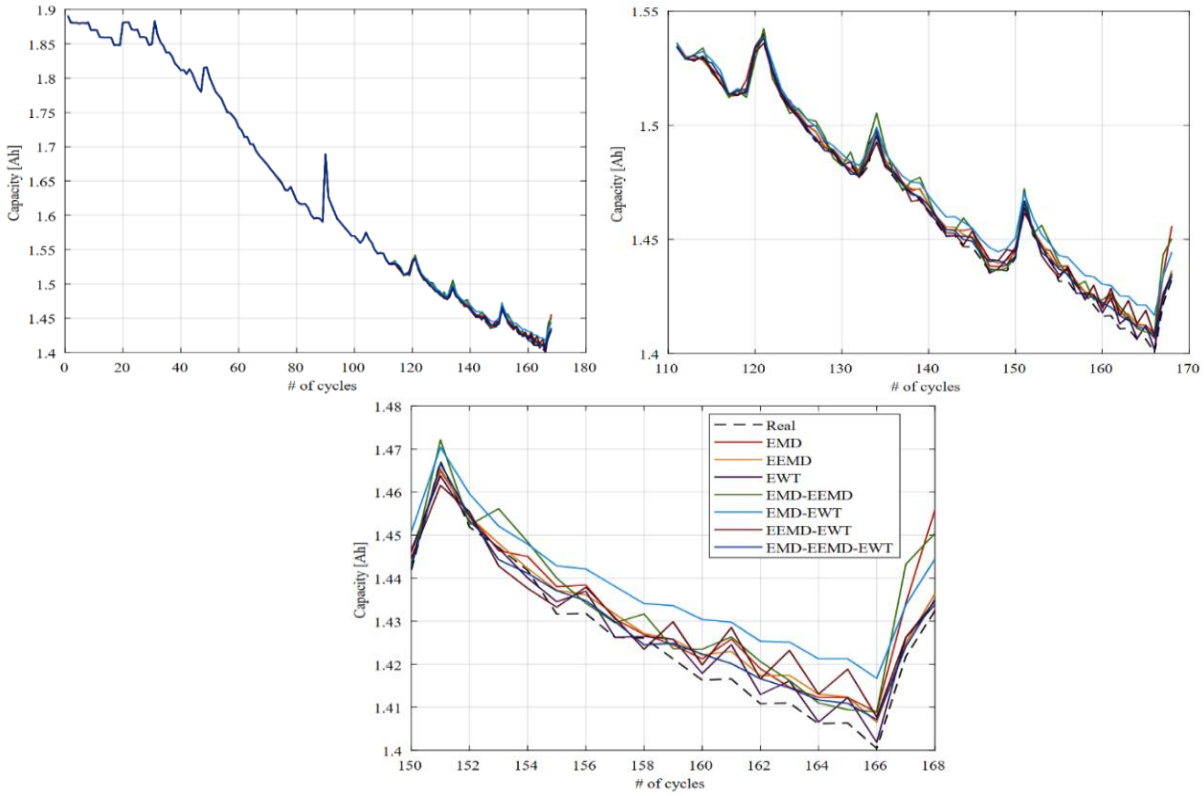

Figure 6. B0007 battery prediction results.

Table 7. Performance evaluations of B0007 battery.

Model	MSE	MAPE	RMSE
EMD	2.72×10^{-5}	0.002539	0.005218
EEMD	1.23×10^{-5}	0.002001	0.003509
EWT	8.47×10^{-6}	0.00133	0.00291
EMD-EEMD	3.94×10^{-5}	0.003155	0.006279
EMD-EWT	6.56×10^{-5}	0.00467	0.008101
EEMD-EWT	2.03×10^{-5}	0.002203	0.004507
EMD-EEMD-EWT	8.29×10^{-6}	0.001666	0.00288

Among the individual models, EMD shines as the standout performer, boasting the lowest MSE, MAPE, and RMSE values, underscoring its remarkable precision in forecasting battery capacity according to Figure 7 and Table 8. When exploring the realm of binary model combinations, EMD-EEMD takes center stage as the top-performing duo, displaying the most favorable MSE, MAPE, and RMSE values among the binary pairs. This underscores the efficacy of amalgamating EMD and EEMD decomposition techniques into a singular model, yielding the utmost accuracy in predicting battery capacity values. For this battery, the EMD model in isolation and the EMD-EEMD combination model emerge as the optimal choices for predicting battery capacity values, as substantiated by the provided metrics.

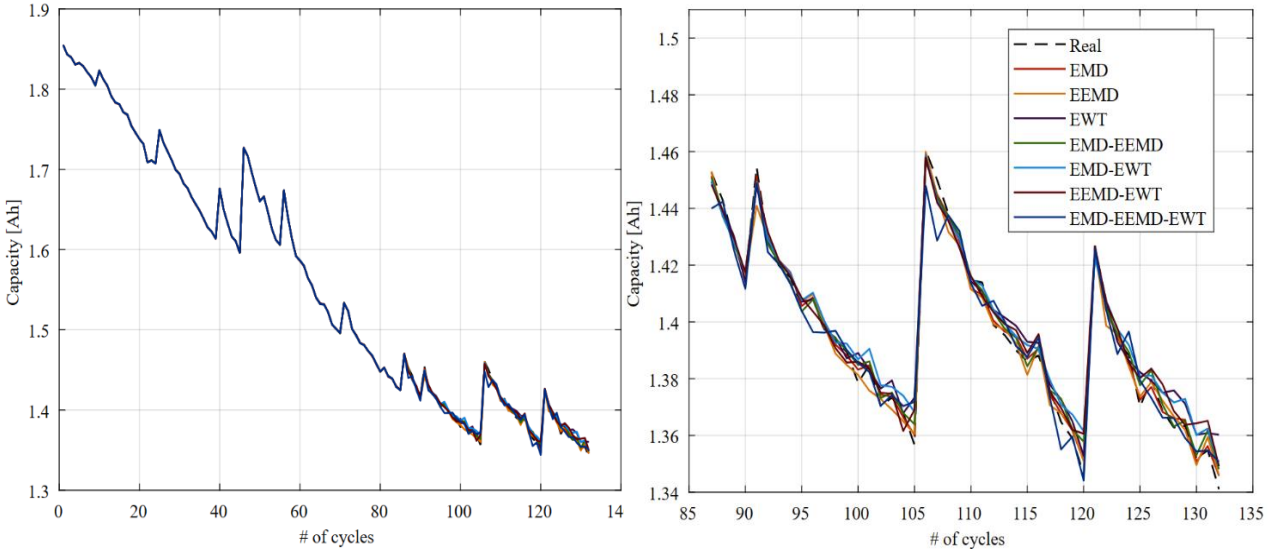

Table 8. Performance evaluations of B0018 battery.

Model	MSE	MAPE	RMSE
EMD	7.15×10^{-6}	1.58E-03	2.67E-03
EEMD	1.78×10^{-5}	0.002442	0.004223
EWT	4.06×10^{-5}	0.00339	0.006369
EMD-EEMD	1.49×10^{-5}	0.002107	0.003864
EMD-EWT	3.34×10^{-5}	0.003411	0.005783
EEMD-EWT	3.30×10^{-5}	0.003307	0.005741
EMD-EEMD-EWT	4.33×10^{-5}	0.003401	0.006581

Table 9 shows the performance metric values of the lithium-ion batteries according to the features obtained from different decomposition models. Lower values of MSE, MAPE, and RMSE indicate better model performance. Among the batteries analyzed, Battery B0005 demonstrates the most favorable results across all three metrics with the EMD-EEMD-EWT model, which boasts the lowest MSE, MAPE, and RMSE values, making it the optimal choice. Battery B0006, on the other hand, showcases distinct performance characteristics. The EEMD model stands out with the lowest MSE and RMSE values, while the EMD-EEMD
model achieves the lowest MAPE value. Consequently, both the EMD-EEMD model and the EEMD model exhibit commendable performance across all three evaluation criteria for Battery B0006. When examining Battery B0007, the EMD-EEMD model emerges as the top performer, displaying the lowest MSE and RMSE values. Meanwhile, the EWT model secures the lowest MAPE value. Notably, both the EMD-EEMD model and the EWT model consistently perform well across all three metrics for Battery B0007. Lastly, for Battery B0018, the EMD-EEMD-EWT model excels, recording the lowest MSE, MAPE, and RMSE values. This establishes it as the preeminent model in this context.

In addition to utilizing various established datasets for validation, we assessed the accuracy of our proposed technique by comparing it to other conventional, model-based, and intelligent approaches, as detailed in Table 10.

Table 9. Performance evaluations of models' mean values.

Model	MSE	MAPE	RMSE
EMD	3.96×10^{-5}	0.002288	0.004815
EEMD	2.42×10^{-5}	0.001965	0.003128
EWT	4.82×10^{-5}	0.003881	0.006438
EMD-EEMD	3.67×10^{-6}	0.001351	0.002311
EMD-EWT	5.53×10^{-6}	0.001741	0.002825
EEMD-EWT	9.37×10^{-6}	0.002620	0.004026
EMD-EEMD-EWT	1.64×10^{-5}	0.002211	0.003756

To ensure a fair and meaningful comparative analysis, we employed the same battery type and NASA datasets. Our comprehensive study considered key factors pertaining to the prediction of RUL for lithium-ion batteries, including input features, capacity, battery type, and error metrics. For example, in the work by Li et al. [26], a Particle Filtering (PF)-based technique for RUL prediction achieved an RMSE of 0.04408. Furthermore, Gao and Huang [12] introduced a hybrid approach using SVM and Particle Swarm Optimization (PSO) for RUL prediction, yielding RMSE values of 0.0213 and 0.0514, respectively. Additionally, the model developed by Zhou and Huang [24], based on EMD and Auto-Regressive Integrated Moving Average (ARIMA), achieved an RMSE below 1%. In the case of the Relevance Vector Machine (RVM) [27], the relative error remained below 1%. On the contrary, the MLP-based hybrid EMD-EEMD-EWT model, incorporating 22 input parameters, delivered favorable outcomes in terms of accuracy, adaptability, and robustness.

Table 10. Comparison with existing studies.

Reference	Feature parameter	Algorithms	Performance
[26]	Impedance, aging, number of charging cycle	*PF*	RMSE: 0.2902
[12]	Discharge cycle data	*PSO and SVM*	MSE: 0.0213
[28]	Voltage, current, temperature, capacity and time	*SVM*	RMSE: 0.2159, 0.3108
[27]	Capacity	*EMD and RVM*	MSE: 4.4972×10^{-5}, 1.6437×10^{-5}
[29]	Voltage	*FFNN*	MAE: 29.4218
[30]	Voltage, current, temperature, capacity	*Deep neural network*	RMSE: 3.427
[31]	Voltage, current, temperature, capacity	*BPNN*	RMSE: 0.0819 MSE: 6.7114×10^{-5}
This study	Voltage, current, temperature	*EMD*, *EEMD*, *EWT*, and *MLP*	MSE: 3.67×10^{-6} MAPE: 0.001351 RMSE:0.002311

PF: *Particle Filtering;* ***PSO:*** *Particle Swarm Optimization;* ***SVM:*** *Support Vector Machine;*
RVM: *Relevance Vector Regression;* ***FFNN:*** *Feed Forward Neural Network;*
BPNN: *Back Propagation Neural Network.*

Considering the B0005 battery analysis results, the best performance values of MSE (1.35×10^{-6}), MAPE (0.000703) and RMSE (0.001160) values belong to the EMD-EEMD-EWT model. Considering the experiments with this model B0005 battery makes it the most suitable choice. Considering the results of the B0006 battery, the EEMD model gave the most successful results with the lowest MSE (4.34×10^{-6}), MAPE (0.001254) and RMSE (0.002083) performance values. When the B0007 Battery was examined, the EMD-EEMD model emerged as the best performing model, showing the lowest MSE (8.29×10^{-6}) and RMSE (0.00288) values, and the EWT model achieved the lowest MAPE (0.00133) value. It does. In particular, both the EMD-EEMD model and the EWT model showed close performance for the B0007 Battery.

Finally, for Battery B0018, the EMD model highlights the lowest MSE (7.15×10^{-6}), MAPE (1.58×10^{-3})and RMSE (2.67×10^{-6}) values. The EMD model showed the best results for B0018.

It can be seen that different models come to the fore for different battery groups. In this regard, the fact that the data sets of battery groups and especially the voltage values have different values can change the effect on the estimation algorithm and for batteries

It can be thought that it gives different results. For this reason, the error values were compared by taking the average values of the models. The most successful model emerged as the EMD-EEMD model with MSE (3.67×10^{-6}), MAPE (0.001351) and RMSE (0.002311) values.

Moreover, in the performance comparison with the studies in the literature, the proposed method outperforms in terms of the three performance metrics.

Due to the fact that the other studies in the literature utilized a few input and basic regression methods, their performances of them much lesser than ours. On the other hand, the proposed method uses 22 features obtained from different signal processing methods and MLP models.

Moreover, it is considered that the preprocessing composing elimination of misunderstanding samples and interpolation with equal size is effective in this performance.

Contributions of the paper are as follows

- This study proposed a RUL prediction method based on an expanded feature set with 22 features in total obtained from residues of EMD,

EEMD, EWT, and different combinations of them.

- Then the hybrid models are composed of Multilayer Perceptron (MLP) and different combinations for different scenarios.
- Performances of the method are carried out for different mentioned hybrid models.
- Performances of the method are carried out for different batteries comparatively.
- The best results in terms of mean values are obtained for the EMD-EEMD model as given in Table 9 in the paper.
- The results show that the value of this model has a more effective performance than similar studies in the literature like in Table 8.

In addition, the proposed models can be used to estimate the remaining life after pre-processing the data received from different sensors of the batteries that are a part of electric vehicles. However, it is anticipated that the models to be created and their performance may vary depending on the battery characteristics.

4. Conclusion

Given the significant influence of capacity regeneration on the prediction of Remaining Useful Life (RUL) in lithium-ion batteries, we introduce a novel approach to enhance RUL prediction accuracy. Our method combines decomposition techniques with a focus on the capacity regeneration phenomenon. To begin, we employ the Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), and Empirical Wavelet Transform (EWT) methods to decompose the voltage curve's time series into multiple scales. Subsequently, we build time series prediction models based on Multilayer Perceptron (MLP) neural networks for each of these components. Finally, performance metrics for the models were calculated. While the findings of each method yielded good results, it can be stated that the metric values of the hybrid models used provide even better results.

In future studies, examining the model with the help of features obtained by different signal processing methods like CEEMDAN and machine learning methods like LightGBM, as well as methods based on deep learning like LSTM-based methods, can be explored.

Author's Contributions

Ozancan Bayri: Data Acquisition, Writing, Software, Signal Processing, and Editing.

Sıtkı Akkaya: Supervising, Writing, Software, Analysis, Evaluation and Editing.

Ethics

There are no ethical issues after the publication of this manuscript.

References

[1]. Adnan, M., The Future of Energy Storage: Advancements and Roadmaps for Lithium-Ion Batteries. 2023, MDPI. p. 7457.

[2]. Fang, H. Challenges with the ultimate energy density with Li-ion batteries. in IOP Conference Series: Earth and Environmental Science. 2021. IOP Publishing.

[3]. Hanifah, R.A., S.F. Toha, and S. Ahmad, Electric Vehicle Battery Modelling and Performance Comparison in Relation to Range Anxiety. Procedia Computer Science, 2015. 76: p. 250-256.

[4]. Ji, Y., et al., An RUL prediction approach for lithium-ion battery based on SADE-MESN. Applied Soft Computing, 2021. 104: p. 107195.

[5]. Pang, X., et al., A lithium-ion battery RUL prediction method considering the capacity regeneration phenomenon. Energies, 2019. 12(12): p. 2247.

[6]. Deng, Y., et al., Feature parameter extraction and intelligent estimation of the State-of-Health of lithium-ion batteries. Energy, 2019. 176: p. 91-102.

[7]. Dai, H., et al., A Novel Estimation Method for the State of Health of Lithium-Ion Battery Using Prior Knowledge-Based Neural Network and Markov Chain. IEEE Transactions on Industrial Electronics, 2019. 66(10): p. 7706-7716.

[8]. Li, P., et al., State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network. Journal of Power Sources, 2020. 459: p. 228069.

[9]. Chang, Y., H. Fang, and Y. Zhang, A new hybrid method for the prediction of the remaining useful life of a lithium-ion battery. Applied Energy, 2017. 206: p. 1564-1578.

[10]. Chen, Z., et al., State of Health Estimation for Lithium-ion Batteries Based on Fusion of Autoregressive Moving Average Model and Elman Neural Network. IEEE Access, 2019. 7: p. 102662-102678.

[11]. Wei, J., G. Dong, and Z. Chen, Remaining Useful Life Prediction and State of Health Diagnosis for Lithium-Ion Batteries Using Particle Filter and Support Vector Regression. IEEE Transactions on Industrial Electronics, 2018. 65(7): p. 5634-5643.

[12]. Gao, D. and M. Huang, Prediction of remaining useful life of lithium-ion battery based on multi-kernel support vector machine with particle swarm optimization. Journal of Power Electronics, 2017. 17(5): p. 1288-1297.

[13]. Li, X., C. Yuan, and Z. Wang, State of health estimation for Li-ion battery via partial incremental capacity analysis based on support vector regression. Energy, 2020. 203: p. 117852.

[14]. Fei, Z., et al., Early prediction of battery lifetime via a machine learning based framework. Energy, 2021. 225: p. 120205.

[15]. Sahaand, B. and K. Goebel, Battery Data Set, NASA ames prognostics data repository. NASA Ames Research Center, 2007.

[16]. Zhao, L., Y. Wang, and J. Cheng, A Hybrid Method for Remaining Useful Life Estimation of Lithium-Ion Battery with Regeneration Phenomena. Applied Sciences, 2019. 9(9): p. 1890.

[17]. Huang, N.E., et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 1998. 454(1971): p. 903-995.

[18]. Wu, Z. and N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method. Advances in adaptive data analysis, 2009. 1(01): p. 1-41.

[19]. Torres, M.E., et al. A complete ensemble empirical mode decomposition with adaptive noise. in 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). 2011. IEEE.

[20]. Gilles, J., Empirical wavelet transform. IEEE transactions on signal processing, 2013. 61(16): p. 3999-4010.

[21]. Liu, W. and W. Chen, Recent advancements in empirical wavelet transform and its applications. IEEE Access, 2019. 7: p. 103770-103780.

[22]. Hu, Y., et al., An enhanced empirical wavelet transform for noisy and non-stationary signal processing. Digital signal processing, 2017. 60: p. 220-229.

[23]. Yegnanarayana, B., Artificial neural networks. 2009: PHI Learning Pvt. Ltd.

[24]. Zou, J., Y. Han, and S.-S. So, Overview of artificial neural networks. Artificial neural networks: methods and applications, 2009: p. 14-22.

[25]. Abraham, A., Artificial neural networks. Handbook of measuring system design, 2005.

[26]. Li, L., et al., Battery Remaining Useful Life Prediction with Inheritance Particle Filtering. Energies, 2019. 12(14): p. 2784.

[27]. Zhang, C., et al., Capacity Prognostics of Lithium-Ion Batteries using EMD Denoising and Multiple Kernel RVM. IEEE Access, 2017. 5: p. 12061-12070.

[28]. Ali, M.U., et al., Online Remaining Useful Life Prediction for Lithium-Ion Batteries Using Partial Discharge Data Features. Energies, 2019. 12(22): p. 4366.

[29]. Wu, J., C. Zhang, and Z. Chen, An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks. Applied Energy, 2016. 173: p. 134-140.

[30]. Khumprom, P. and N. Yodo, A Data-Driven Predictive Prognostic Model for Lithium-ion Batteries based on a Deep Learning Algorithm. Energies, 2019. 12(4): p. 660.

[31]. Ansari, S., et al., Multi-Channel Profile Based Artificial Neural Network Approach for Remaining Useful Life Prediction of Electric Vehicle Lithium-Ion Batteries. Energies, 2021. 14(22): p. 7521