MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Remaining Useful Life Estimation for LFP Cells in Second Life Applications

Sanz-Gorrachategui, Ivan; Pastor-Flores, Pablo; Pajovic, Milutin; Wang, Ye; Orlik, Philip V.; Bernal-Ruiz, Carlos; Bono-Nuez, Antonio; Artal-Sevil, Jesús Sergio

TR2021-023 April 04, 2021

Abstract

The increasing deployment of battery storage applications in both grid storage and electric vehicle fields is generating a vast used battery market. These batteries are typically recycled but could be reused in Second Life applications. One of the challenges is to obtain an accurate Remaining Useful Life (RUL) estimation algorithm, which determines whether a battery is suitable for reuse and estimates the number of second life cycles the battery will last. In this paper, the RUL estimation problem is considered. We propose several Health Indicators (HI), some of which have not been explored before, along with simple yet effective estimation and classification algorithms. These algorithms include classification techniques such as Regularized Logistic Regression (RLR), and regression techniques such as Multivariable Linear Regression (MLR) and Multi-Layer Perceptron (MLP). As a more advanced solution, a multiple expert system combining said techniques is proposed. The performance of the algorithms and features is evaluated on a recent Lithium Iron Phosphate (LFP) dataset from Toyota Research Institute. We obtain satisfactory results in the estimation of RUL cycles with errors down to 49 RMSE cycles for cells that live up to 1200 cycles, and 0.24% MRE for the prediction of the evolution of capacity.

IEEE Transactions on Instrumentation and Measurement

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Remaining Useful Life Estimation for LFP Cells in Second Life Applications

Iván Sanz-Gorrachategui, Pablo Pastor-Flores, Milutin Pajovic, Ye Wang, Philip V. Orlik, Carlos Bernal-Ruiz, Antonio Bono-Nuez, Jesús Sergio Artal-Sevil

***Abstract*— The increasing deployment of battery storage applications in both grid storage and electric vehicle fields is generating a vast used battery market. These batteries are typically recycled but could be reused in Second Life applications. One of the challenges is to obtain an accurate Remaining Useful Life (RUL) estimation algorithm, which determines whether a battery is suitable for reuse and estimates the number of second life cycles the battery will last. In this paper, the RUL estimation problem is considered. We propose several Health Indicators (HI), some of which have not been explored before, along with simple yet effective estimation and classification algorithms. These algorithms include classification techniques such as Regularized Logistic Regression (RLR), and regression techniques such as Multivariable Linear Regression (MLR) and Multi-Layer Perceptron (MLP). As a more advanced solution, a multiple expert system combining said techniques is proposed. The performance of the algorithms and features is evaluated on a recent Lithium Iron Phosphate (LFP) dataset from Toyota Research Institute. We obtain satisfactory results in the estimation of RUL cycles with errors down to 49 RMSE cycles for cells that live up to 1200 cycles, and 0.24% MRE for the prediction of the evolution of capacity.**

***Index Terms*— Remaining Useful Life, Second Life Battery Applications, Lithium-ion batteries, LFP, Capacity Prediction**

I. Introduction

ENERGY Storage Systems (ESS) based on lithium-ion cell technologies are becoming the standard for many different storage applications due to the high energy density, high efficiency, and declining manufacturing costs [1], [2]. Additionally, they also provide longer service lives than other traditional alternatives, such as Lead-Acid chemistries, and are less contaminant than Nickel related chemistries such as NiCd [3], [4]. Lithium-ion technologies, such as LFP, NMC, LTO, or NCA, are some of the most popular lithium-based variants and are the main options used in Electric Vehicle (EV) applications [5]–[8]. The increasing number of these ESSs due to the expansion of EV is beginning to generate an enormous second-hand battery market, and reusing these batteries is a promising application [9].

Second Life of Lithium-Ion batteries has been studied from an economical, technical, and environmental approach [10]. A common conclusion is that a suitable second life application for these batteries are low-demanding applications in terms of power requirements and depth of discharge (DoD), such as energy storage in on-grid systems [10]–[12]. However, some key challenges need to be addressed before this becomes a reality, such as estimating the Remaining Useful Life (RUL) of these used batteries. The aging of lithium-ion batteries has been extensively studied and has proven to be a challenging problem [13]–[18]. RUL estimation studies have traditionally been focused on in-situ applications i.e. estimating the RUL of batteries that are already embedded in an application and monitored, where the main goal is to predict the failure threshold [19]–[24]. In the problem of RUL estimation for second life applications, knowledge of the past use of the battery is usually required [25].

Regarding the techniques used for the RUL estimation problem, there are two traditional approaches to the problem: model-based analysis and data-driven analysis. While model-based analysis is suitable for battery behavioral models, cell aging is a much more complex process, with many more agents involved. Thus, a data-driven approach is suitable for this task. However, until very recently, publicly available datasets were small [19], [26]. The dataset introduced in [27] was published recently (March ’19) by researchers of the Toyota Research Institute (TRI) and the MIT, and is the largest currently available. This dataset contains information on 124 cells, which are cycled until their End of Life (EoL), as defined as their capacity falling below 80% of their initial capacity. This provides more than 90000 full discharge cycles, with several measured features. In the original paper, the authors use this dataset for RUL prediction purposes, introducing some novel health indicators (HI). Their focus is the early RUL estimation on a monitored cell before it shows capacity-fade.

This paper takes the previous study as a starting point. We use the same dataset, but instead, focus on RUL and capacity estimation for second-life applications. We assume that a cell at an unknown moment of its life needs to be evaluated, and

This work has been supported in part by the Fulbright Commission, and by RIS3 Aragon Government and EU project SAHI (LMP16_18)

Authors I.S-G., P.P-F., C.B-R., A.B-N. and JS.A-S. are with the Electrical and Communication Engineering Department, University of Zaragoza, Zaragoza, Aragón 50018, Spain (email {isgorra, pablop, cbernal, antoniob, jsartal}@unizar.es)

Authors Y.W. and P.V.O. are with Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA (email {yewang, porlik}@merl.com)

Authors I.S-G. and M.P. were with MERL when this work was performed. M.P. is now with Analog Devices (email: micopajovic@gmail.com)

through some test cycles, we extract features from processing the measured voltage and charge waveforms. For this purpose, we use some HI that have already been introduced in the literature, together with some less explored indicators. With this set of measurements, we train some simple but effective machine learning algorithms to classify the cells in "long RUL" or "short RUL" classes and predict the exact number of RUL cycles and the future capacity after a set amount of cycles.

This paper is organized as follows. In section II, we discuss different approaches to the RUL estimation problem, and we describe the HI that we use. In section III, we propose a simple classification algorithm that can cluster short RUL or high RUL batteries with high accuracy. Section IV describes some regression algorithms to estimate the number of remaining cycles of the cells and predict the future capacity in a set amount of cycles. We summarize our conclusions in Section V.

II. Machine Learning Approach

In the literature, the RUL estimation problem has been addressed with different approaches. Data-driven or machine learning approaches are rising in popularity in recent years; however, the lack of convenient and large datasets has hampered progress in this area. The different approaches to address the estimation of RUL of a cell have been traditionally classified into these two categories [19], [28]:

- **Direct estimation of Remaining Useful Life.** When considering direct estimation, we assume that the capacity is available as a feature for our algorithms. This means that when the cell is not in an online application, we can perform test cycles to obtain the capacity, or when the cell is in an online application that cycles it completely, we can directly measure its capacity without disconnecting it.
- **Indirect estimation of Remaining Useful Life.** This type of estimation does not rely on capacity as a feature to determine aging and remaining useful life. It is aimed at industrial applications, where the battery is not cycled completely, and where it is not possible to disconnect cells to measure features.

When studying if a cell is suitable for a second life application, all the features can be measured through laboratory cycles, thus can be used in a direct estimation algorithm. On the other hand, some features (such as capacity) might not be directly measurable in an industrial application. The focus of this paper is the direct estimation of RUL and future capacity, although in the last section we will provide some insight for indirect estimation problems.

We use the dataset from [27] since it is the largest publicly available. It contains information from 124 commercial Lithium Iron Phosphate (LFP)/graphite cells, which were cycled until their EoL. The specific model of the cells is APR18650M1A from A123 Systems, and they have a 1.1Ah nominal capacity. The cells have variated high-current charge patterns, but they have all been discharged with the same discharge current (4C) along with all their life. During all cycles, the ambient temperature was controlled to be 30ºC. The cell lives range from 170 cycles to 2237 cycles in extreme cases.

In their data collection, waveforms for voltage, temperature, charge, and time were measured for each cycle of each cell. The dataset also provides the observed capacity and an estimate of the internal resistance of the cell once for each cycle.

The study in [27] uses this dataset and aims to estimate RUL after the 100th cycle of a cell that has been monitored from the beginning of its life, thus using information from its first 100 cycles. In said study, the authors conclude that the variance of the difference of the charge (Q) in the discharge waveform between cycles 100 and 10, $var(\Delta Q_{100-10})$ is a key feature for determining the RUL of the cell.

However, our study is motivated by a different concern, where we have a cell that has been used in a high current, deep cycle application (such as EV), and we would like to determine whether the cell is suitable for another application, by estimating its RUL and its future capacity. Thus, we do not have full information on its past usage, but we can perform some test cycles to the cell to collect data and calculate some useful features. Let us say we perform δ cycles to an already used cell, which is in its *i*th cycle of life. We can process the feature $var(\Delta Q_{(i+\delta)-i})$, however, this turns out to be very noisy unless δ is large enough. Making δ bigger helps to solve this issue, but it means wasting life cycles of the cell for testing purposes. For this reason, we consider some other features, presented in the following sub-section.

A. *Selected features*

Unlike the original study [27], we are going to consider that each cycle in the dataset is an example, instead of each cell. This approach is convenient considering our second life hypothesis: as we do not know what condition a second-hand battery will be in when we receive it, the algorithms need to be trained with cycles of batteries in different moments of their life. Furthermore, this way we increase the dataset from 124 cell examples to more than 90000, which is the total amount of cycles of all the cells. Each example will have its own set of features and will have as output the number of RUL cycles the cell had at that specific cycle. Some of these features, such as the measured capacity and internal resistance, are already provided by the dataset for each cycle of each cell. Other features are computed from processing the waveforms of each cycle.

As introduced before, the hypothesis is that we are conducting δ test cycles to an already used cell. This means that for most of the features, we will have available δ measurements, while for some others we will have available δ-1, as is described below.

- **Capacity, C [Ah]:** Amount of charge extracted in each full-discharge cycle. Fig. 1a shows an example of the evolution of this feature for one of the cells in the dataset. Degradation can be seen over the life of the cell, with an accelerated decay near the EoL.
- **Internal Resistance, IR [Ω]:** Internal resistance value. One measurement per cycle is provided in the original dataset, obtained by averaging ten current pulses at 80%

of State of Charge (SoC). An example of its evolution is shown in Fig. 1b, where it can be observed how the internal resistance increases abruptly when the cell is near its EoL. Some specific cells with an abnormally high IR value (30% higher than the average) have been taken out as outliers.

- **TIEDVD [min]:** Time-lapse between two voltage thresholds (V_{max}, V_{min}) while discharging a constant and known current. Seen in the literature in industrial applications for indirect estimation purposes ([19], [22]) when it is not possible to completely discharge the cell to measure its capacity (e.g. an Uninterruptible Power Supply). This can be used here as another feature obtained from the voltage-time waveforms of each cycle. We have tried different threshold voltages for estimation purposes, and have found that the couple (*3.3V*, *3.15V*) shows good estimation capabilities for this dataset. It can be appreciated in Fig. 1c how it decreases very linearly with the remaining useful life.
- **Capacity Fade ΔC_δ [C]:** Variation in capacity between two cycles separated δ cycles apart. Similarly, as what happened with $var(\Delta Q_{(i+\delta)-i})$, a greater value for δ improves the prediction capabilities, at expense of aging more the cell for test purposes. We have found that δ=10 is a good tradeoff in accuracy. Later on, this aspect will be further developed. An example of its evolution is shown in Fig. 1d. As happens with Capacity and Internal Resistance, in the last hundreds of cycles, its behavior changes drastically.

Besides these features, which have been used previously in the literature, we introduce two new health indicators we found for this dataset which provide useful information.

- **Capacitance peak, Cpk [AhV^{-1}]:** As the cell ages, a small disturbance appears in the low voltage part of the discharge waveform. It grows until a certain moment in its life and then becomes smoother when the cell is near its EoL. This generates a peak in the cell equivalent capacitance [27], [29]–[32], as seen in Fig. 2. This capacitance peak can be used as another feature. Fig. 1e shows how this peak evolves along with cell life, increasing in the first hundreds of cycles, and then decreasing in the last 150 cycles approximately. However, there are some issues with this feature: When the peak is near its maximum, it is a very noisy feature. On the other hand, when the peak is near its minimum (the first 80 cycles of life), the peak cannot be detected correctly.
- **Voltage at Capacitance Peak, Vpk [V]:** We also use the voltage at which the capacitance peak is observed, as an additional feature, as it varies with cell life. The evolution of this feature is shown in Fig. 1f. Similarly to the case of the Capacitance peak, it can be seen how in the first 80 cycles of life it cannot be captured correctly

Similar capacitance peaks obtained with lower discharge currents have been studied before for LFP batteries, and have been used as health indicators [27], [32]. However, these approaches aimed to determine degradation mechanisms, and have not been extensively applied to high current waveforms.

Given these processed features, our algorithm aims to estimate:

- **Remaining Useful Life, RUL [cycles]:** Number of remaining cycles from the current cycle until EoL. Cells with more than 1300 RUL cycles or with less than 300 RUL cycles have been taken out as outliers since the vast majority of the dataset seems to range between 500 and 1200 life cycles. We end up using 113 cells, which add up to more than 80000 test cycles in the selected dataset.

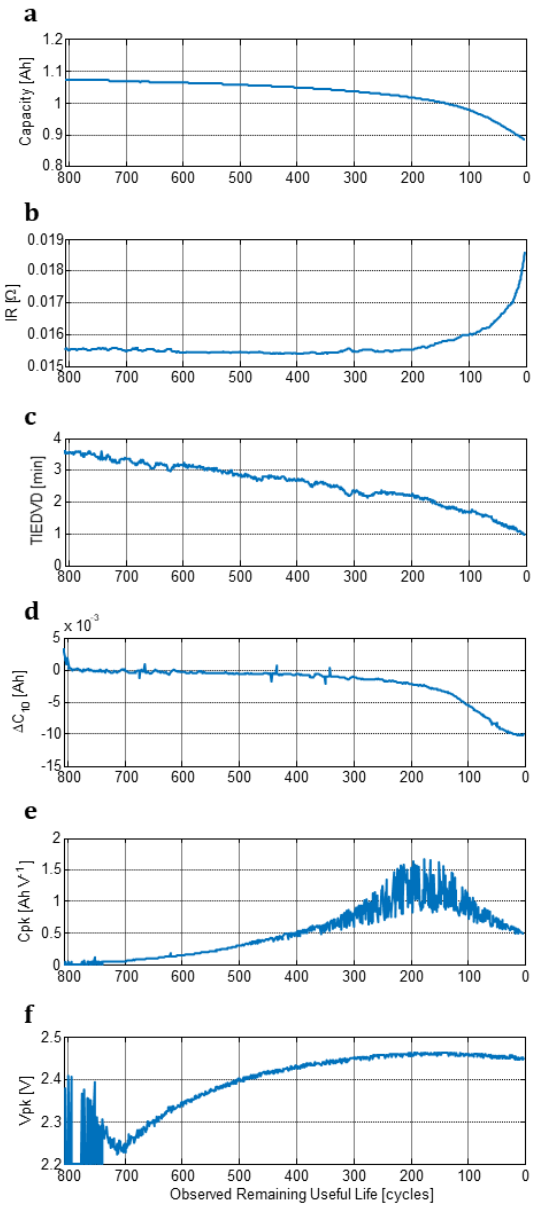

Fig. 1. Feature evolution for one cell vs its life cycles (#100). a, Capacity evolution. b, Internal Resistance evolution. c, TIEDVD evolution. d, ΔC10 Evolution. e, Capacitance Peak evolution. f, Voltage at Capacitance peak evolution

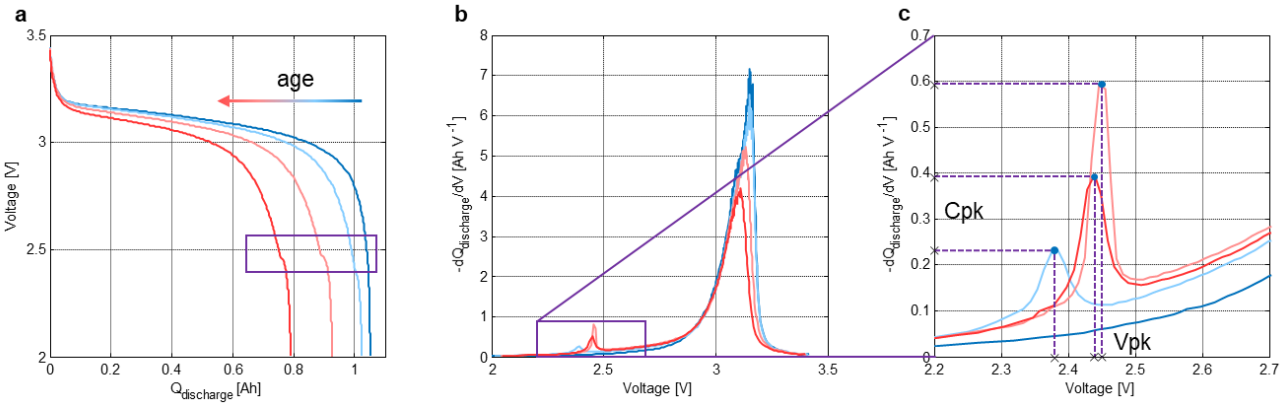

Fig. 2. Capacitance peak advent. a, Voltage vs. Charge curve and its variation as the cell ages. b, Charge derivative to Voltage (Capacitance), and its variation as the cell ages. c, Zoom in the lower voltage part, where the secondary Capacitance peak appears

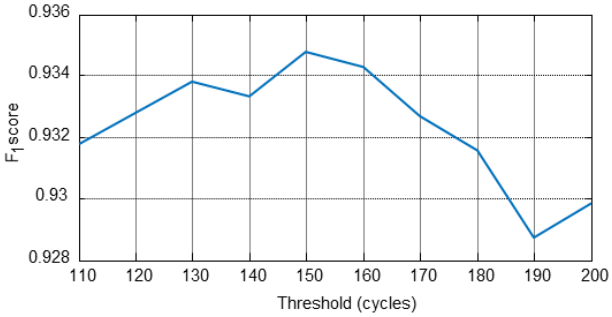

Fig. 3. F1 score for classification of cells into "short RUL" and "long RUL", with different threshold values

III. RUL Classification

As discussed in the previous section, there is a clear trend in some specific features when the cells are close to their EoL. The features Capacity, Internal Resistance, and Capacity Fade change drastically when the cell is in its last 100 to 200 cycles. In this interval, the 'aging knee' of the Capacity curve takes place [10], and according to [33], batteries below this life horizon should not be considered for a second life application. This change of behavior also suggests the possibility of training an algorithm to determine whether a cell is in this EoL region.

Thus, we have defined the classes "Short RUL cell" and "Long RUL cell" to classify those cells below or above this threshold (in remaining cycles). To establish the best threshold between the classes, an initial classification algorithm (logistic regression) has been tested. We have obtained the F1 score for different threshold values with this algorithm, and the best value has been established in 150 RUL cycles (Fig. 3).

A. *Classification algorithm*

The algorithm that has been proposed for the classification task is regularized logistic regression. The model that this algorithm tries to fit is given by the expression in (1).

$$\hat{y}(\boldsymbol{x}, \boldsymbol{w}) = \frac{1}{1 + e^{-\boldsymbol{w}^T \boldsymbol{x}}} \quad (1)$$

Where the predicted output $\hat{y}$ is the probability for the cell to belong to "Short RUL cell" and "Long RUL cell" classes respectively, $\boldsymbol{x}$ is an n-dimensional feature array, and $\boldsymbol{w}$ is an n-dimensional weight array. The output class is obtained by thresholding $\hat{y}$ with 0.5, with 0 meaning "short RUL" and 1 meaning "long RUL".

The algorithm fits $\boldsymbol{w}$ by minimizing the cost function J in (2).

$$J(\boldsymbol{w}) = \frac{1}{m}\sum_{i=1}^{m}\left(E(\boldsymbol{x}_i, \boldsymbol{w}, y_i)\right) + \lambda \sum_{j=1}^{n} \boldsymbol{w}_j{}^2 \quad (2)$$

In this equation, the first term represents the traditional cost function for conventional logistic regression, and the second term represents the Ridge regularization function that makes the algorithm less prone to overfitting. The parameter λ weights the regularization term and is set in the training. The parameter m represents the number of examples in the training dataset and the error function $E(\boldsymbol{x}_i, \boldsymbol{w}, y_i)$ represents the traditional cost function used for conventional logistic regression, given by (3):

$$E(\boldsymbol{x}_i, \boldsymbol{w}, y_i) = -y_i \log\left(\hat{y}(\boldsymbol{x}_i, \boldsymbol{w})\right) - (1 - y_i) \log\left(1 - \hat{y}(\boldsymbol{x}_i, \boldsymbol{w})\right) \quad (3)$$

B. *Determining the optimal number of test cycles δ*

The number of test cycles δ has been introduced before as a means of obtaining more measurements for the features of the cell under test. Specifically, a larger δ allows us to gather more information about the RUL of the cell according to (4). A specific number of cycles δ gives us δ measurements for each feature except ΔC, and δ-1 measurements for ΔC.

$$n_{measurements} = 6\delta - 1 \quad (4)$$

This comes at expense of reducing the RUL for test purposes. Thus, there is a tradeoff between δ and the success rate achieved. To determine the best δ, logistic regression has been used as the initial algorithm for the classification task, due to its simplicity. We have tested different sets of variables, including and excluding some of them and applying certain transformations. The best combination of features we found is gathered in Table I, which makes a total of 4δ measurements. Here, Cpk and Vpk have been combined into a single feature.

TABLE I
FEATURES USED IN THE ALGORITHMS

Feature	Samples used as inputs	Nº of samples
C	C_δ	1
ΔC	$[\Delta C_2, \ldots \Delta C_\delta]$	δ-1
Vpk, Cpk	$[Vpk_1/\log_{10}(Cpk_1), \ldots Vpk_\delta/\log_{10}(Cpk_\delta)]$	δ
IR	$[IR_1, \ldots IR_\delta]$	δ
TIEDVD	$[TIEDVD_1, \ldots TIEDVD_\delta]$	δ

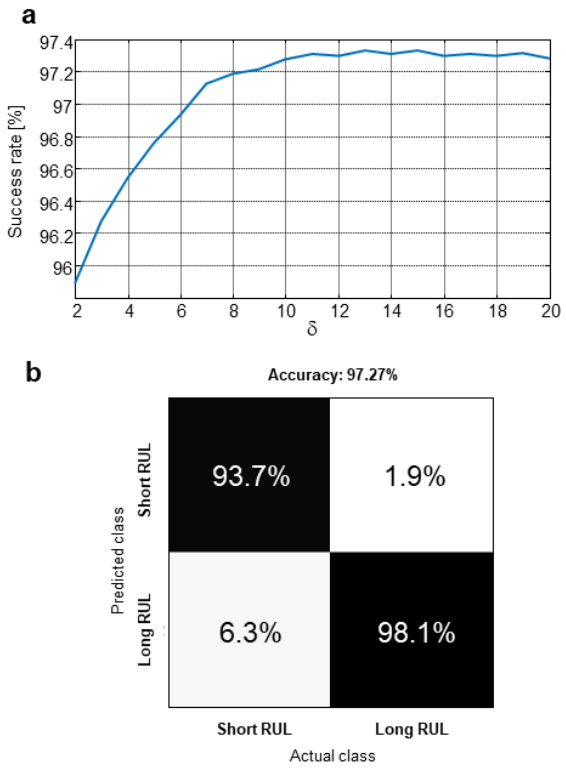

Fig. 4. Results for the regularized logistic regression algorithm. a, Success rate for each δ value. b, Confusion matrix for the case of δ=10

The logistic regression algorithm has been evaluated for different δ values, with 60% of the dataset for training, 20% for cross-validation, and 20% for testing. The sub-sets were created randomly from the examples in the dataset. Fig. 4a shows the success rate of the classification task for the test subset with the logistic regression algorithm, as a function of δ. It can be seen that values of δ above 10 do not improve the accuracy of the method, while additional measurements further reduce the RUL of the cell. Thus, we will only consider the case of δ = 10 in the rest of the analysis. Fig. 4b shows the confusion matrix for this case.

The overall accuracy in the classification obtained with this method is around 97.27%, with 93.7% accuracy for detecting "Short RUL cells", and 98.1% accuracy for "Long RUL cells". We emphasize that we achieve this high accuracy for cells in an unknown moment of their life, i.e., we do not know how many cycles they have been used before the measurement tests.

IV. RUL Regression

After observing the classification capabilities of the features, we would like to know how well they are capable of performing actual RUL estimation, i.e., predicting the specific number of remaining cycles for a given cell. For this purpose, as an initial approach, multivariable linear regression is considered.

A. *Direct RUL Estimation approach*

The first RUL regression method that we consider is the Multivariable Linear Regression (MLR) algorithm. We use 60% of the dataset for training, while 20% is used for cross-validation and 20% for testing. The predicted output of this algorithm for a given measurement input array is given by the expression (6).

$$\hat{y}(\boldsymbol{x}, \boldsymbol{w}) = \boldsymbol{w}^T \boldsymbol{x} \tag{6}$$

Where $\hat{y}$ is the predicted RUL, $\boldsymbol{x}$ is a 4δ +1 dimensional measurement array, and $\boldsymbol{w}$ is a 4δ +1 dimensional weight array (due to the bias term). The algorithm fits $\boldsymbol{w}$ by minimizing the cost function J in (7).

$$J(\boldsymbol{w}) = \frac{1}{2m}\sum_{i=1}^{m}(\boldsymbol{w}^T \boldsymbol{x}_i - y_i)^2 + \lambda \sum_{j=1}^{n} \boldsymbol{w}_j{}^2 \tag{7}$$

Where the first term is the traditional least-squares function and the second term is the Ridge regularization function, introduced to avoid overfitting as in the case of the classification problem. Again, the parameter m is the number of training examples and λ, the regularization weight.

To select from the 40 features those that are more relevant for regression purposes, the sequential forward selection (SFS) and sequential backward selection (SBS) algorithms have been used [34]–[36]. The SFS method starts with an empty set of measurements and evaluates each of them independently, selecting the one that performs the best. Then, saves it and adds each of the other measurements, selecting the one that performs the best. The algorithm keeps on selecting measurements until all of them have been added. Then selects the iteration with less error as the best performing mix of features. The SBS method does something similar, but starting with a full dataset and evaluating the error when each measurement is removed. Those that have less impact on the error are less important, and keep on being eliminated. Each simulation of the algorithms gave different results due to the random generation of the training, validation, and test subsets. However, two common aspects have been observed:

- The performance of both algorithms in their optimum points is always similar.
- The best performance versus the number of measurements tradeoff is obtained when including just one measurement for each type of feature. The performance can be slightly increased when adding more measurements. Specifically, IR seems to be the least important feature in that regard, and most of its measurements can be omitted.

The best performance in these algorithms is obtained when including the first 36 measurements for the SFS algorithm. The variables finally discarded are IR_1, IR_2, IR_4, and IR_8, all of them, among the 10 measurements of Internal Resistance. They seem to provide redundant information, which is already given by other IR measurements.

As an alternative to MLR, some more sophisticated algorithms have been considered. One of these algorithms is the Multilayer Perceptron (or Feed-Forward Neural Network). The inputs to this algorithm are the same as for the case of Multivariable Linear Regression. This algorithm has been trained using the Levenberg-Marquardt optimization algorithm with $\mu_0 = 10^{-3}$ and the Mean Squared Error (MSE) as the optimization goal, for a maximum of 1000 epochs with early stopping as a validation technique. A single hidden layer has

been used, and the number of neurons in the said layer has been swept. The optimal was found at 32 neurons.

The algorithms have been trained with the whole dataset (maintaining 60% for training, 20% for validation, and 20% for test). However, after observing the accuracy in the classification problem, we conclude that the "Short RUL" and "Long RUL" classes have a clear different behavior. Thus, we propose "expert" algorithms, which have been trained and tested with two different subsets: one for short RUL cells (those below 150 RUL cycles) and the other for long RUL cells (those above 150 RUL cycles). These expert algorithms have been combined in the Multiple Expert System (MES) (Fig. 5), where, the "RUL Classifier" block (the logistic regression algorithm described in the previous section) chooses which expert to use after classifying the input vector in one of the classes. Once the expert has been chosen, the input is fed to such expert (either "Short RUL expert" or "Long RUL expert", thus obtaining the final output.

The test results for the algorithms are collected in Table II. The error metrics are provided in terms of Root Mean Square Error (RMSE) (8) and Mean Relative Error (MRE) (9).

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2} \tag{8}$$

$$MRE = \frac{1}{m}\sum_{i=1}^{m}\frac{(\hat{y}_i - y_i)}{y_i}\cdot 100(\%) \tag{9}$$

Where m is the number of examples in the test dataset, $\hat{y}_i$ is the predicted RUL for the i-th example, and y_i is the actual RUL values for the i-th example.

In the table, the algorithms labeled as "global" are those trained with the whole dataset (maintaining 60% for training, 20% for cross-validation, and 20% for test). Although some of these error rates may seem unimpressive, especially in the case of MRE near 50%, cells with very low RUL (only a few cycles) have a large impact on this averaged error. A small error of only a few cycles for cells with low RUL (e.g., one of two cycles) leads to a large MRE. To clarify this, Fig. 6 provides information on the RMSE and MRE obtained for each value of Observed RUL. Note that in the case of the MRE (Fig. 6b), the error is within 10-20% for practically the whole range of observed RUL. It can be seen that MRE is not a good representation of the method on the first 100 cycles. The third column in the table gives MRE conditioned on larger values of RUL, which yields much better MRE results.

The results for the MES are given in the table, where it may be appreciated how it enhances the performance of the global algorithm, obtaining an average MRE of 15.2% in the estimation of RUL for cells in an unknown moment of their life, and below 10% conditioned to long RUL cells. Fig. 7 shows an example of RUL estimation for all of the test cycles for one specific cell in the dataset (#100). It can be appreciated how the RUL prediction curve follows accurately the RUL line. This is especially noticeable for low RUL values, where the Short Expert comes into play (below 150 RUL cycles), where despite the higher MRE the absolute prediction is much more accurate.

As a means of comparison, the last row of the table collects the RUL estimation error from the Elastic Net algorithm developed in [27] with the same dataset, which used the linear regression framework. The results in terms of RMSE are closer to the MLR algorithm. It must be noted that the approach followed in that reference is not the same that has been followed here since they used information from the first 100 cycles of each cell. This has an impact on the lower MRE since they do not try to predict very low RUL values.

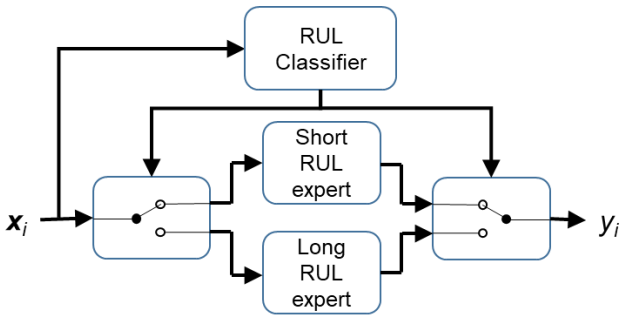

Fig. 5. MES scheme, where the RUL classifier acts as a selector of the expert

TABLE II
ERROR METRICS FOR DIFFERENT ALGORITHMS

Algorithm	RMSE (cycles)	MRE (%)	MRE (%) (>150 cycles)
Global MLR	90	53.81	18.51
Global MLP	52	23.03	10.51
MLP Short RUL expert (<150 cycles)	13	28.73	-
MLP Long RUL expert (>150 cycles)	61	10.32	10.32
MES	49	15.2	9.79
Elastic Net	86	10.1	-

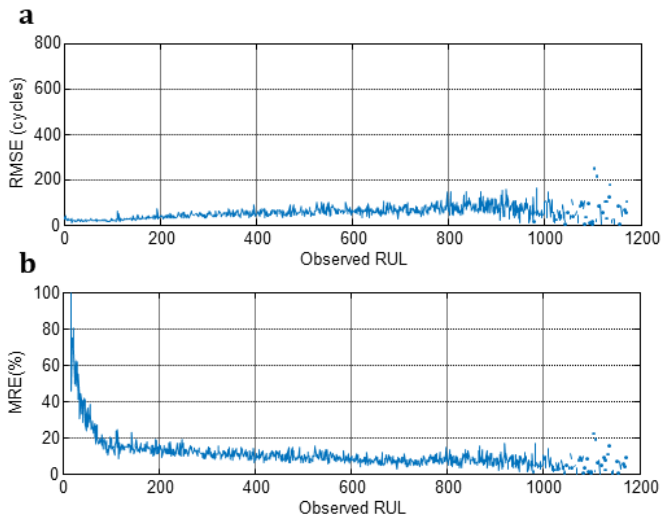

Fig. 6. Error rates for the MLP algorithm vs. observed RUL. a, RMSE. b, MRE

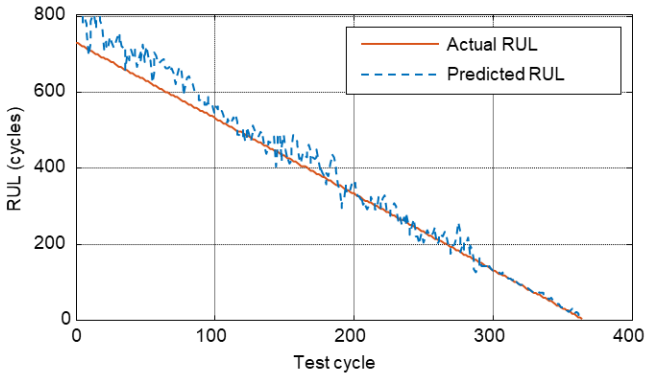

Fig. 7. RUL estimate with MES for all the cycles belonging to cell #100

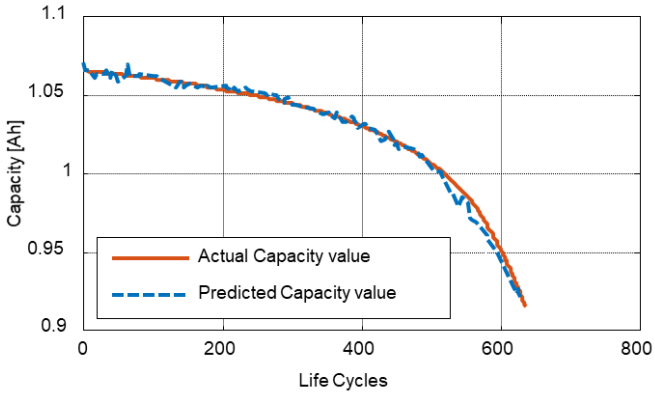

Fig. 8. Capacity prediction for cell #100 after 150 cycles

TABLE III
ERROR RESULTS FOR CAPACITY PREDICTION

Cycles ahead	δ	MRE (%)
100	1	0.45
	10	0.24
150	1	0.67
	10	0.45
200	1	0.82
	10	0.64

TABLE IV
ERROR RESULTS FOR RUL PREDICTION BASED ON CAPACITY

Cycles ahead	MRE (%)
100	6.9
150	8.6
200	9.99

B. *Capacity prediction approach*

In the problem of capacity prediction, the algorithm will predict the future capacity of the cell, after a set number of cycles ahead from the cycle where the features have been measured. As test examples, we have considered 100, 150, and 200 cycles ahead. The closer the prediction horizon is located, the more accurate will be the prediction. The same features as in the RUL estimation algorithm have been chosen. Besides, different δ values for the measurement cycles have been considered. Since MLP had obtained the best results among the global algorithms, it has been chosen for this task. The results are shown in Table III.

The accuracy in the prediction of capacity is very high according to the MRE metrics obtained. Among the δ and cycles ahead values that have been tested, the best MRE obtained is 0.24%. However, a higher window observation δ or a closer prediction horizon provides better error figures. As an example, Fig. 8 shows the output for cell #100 in the dataset, when predicting capacity 150 cycles ahead of the measurement moment.

Observing the accurate capacity prediction results, we wonder if there is a possibility of using this algorithm for RUL estimation purposes. In this scenario, we define an EoL condition based on capacity (e.g. 80% of nominal capacity) and we train algorithms to predict cell capacity *X* cycles ahead. We keep track of the battery along with its life, and when the prediction falls below the EoL condition, we predict RUL as *X* cycles. Table IV shows the accuracy in this RUL prediction method, which is higher than for the case of the MES algorithm but with a fixed number of cycles ahead.

C. *Indirect RUL Estimation approach*

So far, we have described the study of the RUL estimation problem from a Direct Estimation approach, meaning that capacity measurements are available as a feature for further analysis. This is meaningful from the second life application perspective since we are conducting some test cycles to a cell in a controlled environment and it can be charged and discharged at will.

However, it is interesting to see how well these features can perform in other scenarios e.g., a system where the battery cannot be completely discharged for test purposes, such as in an Uninterruptible Power Supply (UPS). In these systems, direct measurement of capacity and other capacity-related features is generally unavailable and thus, RUL estimation algorithms must do as best as they can with a subset of features. This has been introduced as “Indirect RUL Estimation” in the literature [19], [28].

Thus, from the features we have been considering until now, Capacity and ΔC are not suitable for this approach. Even more, capacitance peak related features (Vpk, Cpk) should not be considered as well (they are found in the last 10% of SoC). For this reason, only TIEDVD and IR features are considered for this approach. From the algorithms that have already been introduced, the MES performs the best. The RMSE and MRE metrics are contained in Table V. The error metrics are far from those obtained in the Direct Estimation approach, but still may be of interest in these applications.

TABLE V
ERROR RESULTS FOR MES IN INDIRECT ESTIMATION

Algorithm	RMSE (cycles)	MRE (%)	MRE (%) (>150 cycles)
MES	109	45.19	23.94

V. Conclusions

Remaining Useful Life estimation is becoming a popular topic in battery-related research. Due to the complexity of the problem, data-driven approaches are appropriate for the task, although until very recently, publicly available datasets were small and simple. We address the problem of RUL estimation from the second life application point of view. The main hypothesis assumes that past information about cell use is not available. By conducting some simple charge and discharge tests and analyzing the voltage waveforms, we can determine the number of RUL cycles with high accuracy. For this purpose, we have proposed several HI to be analyzed. Among them, there are well-known features such as Capacity or Internal Resistance, previously described features such as TIEDVD, and two novel HI (C_{peak}, V_{peak}), which have been observed in this specific dataset.

We propose some simple but effective algorithms to classify used cells into "Short Remaining Useful Life" and "High Remaining Useful Life" categories with an average accuracy higher than 97%. We also proposed methods to estimate the exact number of remaining cycles for cells in an unknown moment of their lives, with 49 cycles of RMS Error, providing an overall MRE of 15.2% and 9.79% for high RUL cells. In the capacity prediction problem, we obtain high accuracy when predicting the future value in a set amount of cycles, with relative errors below 1% in all the studied scenarios. As the last contribution, we study the viability of performing indirect estimation in embedded applications with a sub-set of features, obtaining approximate values for the RUL of these cells. The high accuracy in the prediction of short RUL cells (highly decayed and not useful) makes this approach a suitable method for determining whether a cell would be useful for a second life application such as grid-oriented ESS. On top of this, the prediction of the exact number of RUL cycles is made with high accuracy and is considered good for this purpose.

This work has been developed with the dataset by TRI since it is the most complete in terms of the number of cells and cycles among the datasets currently available. Even though this approach would be valid for determining the RUL of any chemistry and type of cell, a specific dataset would be needed for said kind of cell to be able to train the algorithms. Additionally, the applicability of the novel HI described in the paper would need to be researched for other chemistries, which is a subject for future work.

Additionally, future lines of work also include developing advanced algorithms, to enhance the prediction of RUL and the classification of the cells.

References

[1] B. Nykvist and M. Nilsson, "Rapidly falling costs of battery packs for electric vehicles," *Nat. Clim. Chang.*, vol. 5, no. 4, pp. 329–332, 2015, doi: 10.1038/nclimate2564.

[2] B. Dunn, H. Kamath, and J. M. Tarascon, "Electrical energy storage for the grid: A battery of choices," *Science (80-.).*, vol. 334, no. 6058, pp. 928–935, 2011, doi: 10.1126/science.1212741.

[3] B. Diouf and R. Pode, "Potential of lithium-ion batteries in renewable energy," *Renew. Energy*, vol. 76, pp. 375–380, 2015, doi: 10.1016/j.renene.2014.11.058.

[4] S. B. Peterson, J. Apt, and J. F. Whitacre, "Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization," *J. Power Sources*, vol. 195, no. 8, pp. 2385–2392, Apr. 2010, doi: 10.1016/J.JPOWSOUR.2009.10.010.

[5] K. Young, C. Wang, L. Y. Wang, and K. Strunz, "Electric Vehicle Battery Technologies," in *Electric Vehicle Integration into Modern Power Networks*, New York, NY: Springer New York, 2013, pp. 15–56.

[6] C. S. Ioakimidis, A. Murillo-Marrodán, A. Bagheri, D. Thomas, and K. N. Genikomsakis, "Life Cycle Assessment of a Lithium Iron Phosphate (LFP) Electric Vehicle Battery in Second Life Application Scenarios," *Sustainability*, vol. 11, no. 9, p. 2527, May 2019, doi: 10.3390/su11092527.

[7] M. Farhadi and O. Mohammed, "Energy Storage Technologies for High-Power Applications," *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 1953–1961, May 2016, doi: 10.1109/TIA.2015.2511096.

[8] D. Ansean, M. Gonzalez, J. C. Viera, J. C. Alvarez, C. Blanco, and V. M. Garcia, "Evaluation of LiFePO4 batteries for Electric Vehicle applications," in *2013 International Conference on New Concepts in Smart Cities: Fostering Public and Private Alliances (SmartMILE)*, Dec. 2013, pp. 1–8, doi: 10.1109/SmartMILE.2013.6708211.

[9] R. Reinhardt, I. Christodoulou, S. Gassó-Domingo, and B. Amante García, "Towards sustainable business models for electric vehicle battery second use: A critical review," *J. Environ. Manage.*, vol. 245, pp. 432–446, Sep. 2019, doi: 10.1016/J.JENVMAN.2019.05.095.

[10] E. Martinez-Laserna *et al.*, "Battery second life: Hype, hope or reality? A critical review of the state of the art," *Renew. Sustain. Energy Rev.*, vol. 93, pp. 701–718, Oct. 2018, doi: 10.1016/J.RSER.2018.04.035.

[11] Y. Jiang, J. Jiang, C. Zhang, W. Zhang, Y. Gao, and N. Li, "State of health estimation of second-life LiFePO4 batteries for energy storage applications," *J. Clean. Prod.*, vol. 205, pp. 754–762, Dec. 2018, doi: 10.1016/J.JCLEPRO.2018.09.149.

[12] E. Hossain, D. Murtaugh, J. Mody, H. M. R. Faruque, M. S. Haque Sunny, and N. Mohammad, "A Comprehensive Review on Second-Life Batteries: Current State, Manufacturing Considerations, Applications, Impacts, Barriers & Potential Solutions, Business Strategies, and Policies," *IEEE Access*, vol. 7, pp. 73215–73252, 2019, doi: 10.1109/ACCESS.2019.2917859.

[13] R. . Wright *et al.*, "Power fade and capacity fade resulting from cycle-life testing of Advanced Technology Development Program lithium-ion batteries," *J. Power Sources*, vol. 119–121, pp. 865–869, Jun. 2003, doi: 10.1016/S0378-7753(03)00190-3.

[14] M. Broussely, S. Herreyre, P. Biensan, P. Kasztejna, K. Nechev, and R. . Staniewicz, "Aging mechanism in Li ion cells and calendar life predictions," *J. Power Sources*, vol. 97–98, pp. 13–21, Jul. 2001, doi: 10.1016/S0378-7753(01)00722-4.

[15] I. Bloom *et al.*, "An accelerated calendar and cycle life study of Li-ion cells," *J. Power Sources*, vol. 101, no. 2, pp. 238–247, Oct. 2001, doi: 10.1016/S0378-7753(01)00783-2.

[16] S. F. Schuster *et al.*, "Nonlinear aging characteristics of lithium-ion cells under different operational conditions," *J. Energy Storage*, vol. 1, pp. 44–53, Jun. 2015, doi: 10.1016/J.EST.2015.05.003.

[17] H. Chaoui and C. C. Ibe-Ekeocha, "State of Charge and State of Health Estimation for Lithium Batteries Using Recurrent Neural Networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8773–8783, Oct. 2017, doi: 10.1109/TVT.2017.2715333.

[18] S. Sepasi, R. Ghorbani, and B. Y. Liaw, "Inline state of health estimation of lithium-ion batteries using state of charge calculation," *J. Power Sources*, vol. 299, pp. 246–254, 2015, doi: 10.1016/j.jpowsour.2015.08.091.

[19] D. Liu, W. Xie, H. Liao, and Y. Peng, "An integrated probabilistic approach to lithium-ion battery remaining useful life estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 660–670, 2015, doi: 10.1109/TIM.2014.2348613.

[20] L. Li, Y. Peng, Y. Song, and D. Liu, "Lithium-Ion Battery Remaining Useful Life Prognostics Using Data-Driven Deep Learning Algorithm," *Proc. - 2018 Progn. Syst. Heal. Manag. Conf. PHM-Chongqing 2018*, pp. 1094–1100, 2019, doi: 10.1109/PHM-Chongqing.2018.00193.

[21] R. Khelif, B. Chebel-Morello, S. Malinowski, E. Laajili, F. Fnaiech, and N. Zerhouni, "Direct Remaining Useful Life Estimation Based on Support Vector Regression," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2276–2285, 2017, doi: 10.1109/TIE.2016.2623260.

[22] D. Liu, H. Wang, Y. Peng, W. Xie, and H. Liao, "Satellite lithium-

ion battery remaining cycle life prediction with novel indirect health indicator extraction," *Energies*, vol. 6, no. 8, pp. 3654–3668, 2013, doi: 10.3390/en6083654.

[23] C. Chen and M. Pecht, "Prognostics of lithium-ion batteries using model-based and data-driven methods," *Proc. IEEE 2012 Progn. Syst. Heal. Manag. Conf. PHM-2012*, pp. 1–6, 2012, doi: 10.1109/PHM.2012.6228850.

[24] J. Liu and Z. Chen, "Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Health Indicator and Gaussian Process Regression Model," *IEEE Access*, vol. 7, pp. 39474–39484, 2019, doi: 10.1109/ACCESS.2019.2905740.

[25] E. Martinez-Laserna *et al.*, "Technical Viability of Battery Second Life: A Study From the Ageing Perspective," *IEEE Trans. Ind. Appl.*, vol. 54, no. 3, pp. 2703–2713, May 2018, doi: 10.1109/TIA.2018.2801262.

[26] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, and L. Zhang, "Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach," *IEEE Access*, vol. 6, pp. 50587–50598, 2018, doi: 10.1109/ACCESS.2018.2858856.

[27] K. A. Severson *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nat. Energy*, vol. 4, no. 5, pp. 383–391, 2019, doi: 10.1038/s41560-019-0356-8.

[28] L. Li, Y. Zhu, L. Wang, D. Yue, and D. Li, "Indirect remaining useful life prognostics for lithium-ion batteries," *2018 24th Int. Conf. Autom. Comput.*, no. September, pp. 1–5, 2019, doi: 10.23919/iconac.2018.8748973.

[29] M. Günther, U. Feldmann, and J. ter Maten, "Modelling and Discretization of Circuit Problems," *Handb. Numer. Anal.*, vol. 13, pp. 523–659, 2005, doi: 10.1016/S1570-8659(04)13006-8.

[30] J. L. Wyatt, *Foundations of nonlinear network theory*. 1978.

[31] J. R. Macdonald and M. K. Brachman, "The Charging and Discharging of Nonlinear Capacitors," *Proc. IRE*, vol. 43, no. 1, pp. 71–78, 1955, doi: 10.1109/JRPROC.1955.277920.

[32] D. Ansean *et al.*, "Lithium-Ion Battery Degradation Indicators Via Incremental Capacity Analysis," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2992–3002, May 2019, doi: 10.1109/TIA.2019.2891213.

[33] E. Martinez-Laserna *et al.*, "Evaluation of lithium-ion battery second life performance and degradation," 2016, doi: 10.1109/ECCE.2016.7855090.

[34] M. Dash and H. Liu, "Feature Selection for Classification, in Intelligent Data Analysis," *IDA ELSEVlER Intell. Data Anal.*, vol. 1, no. 97, pp. 131–156, 1997, [Online]. Available: www.elsevier.comAocate/ida.

[35] P. A. Devijver and J. Kittler, *Pattern recognition : a statistical approach*. Prentice/Hall International, 1982.

[36] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997, doi: 10.1016/S0004-3702(97)00043-X.