Useful energy prediction model of a Lithium-ion cell operating on various duty cycles

Damian Burzyński*

*Poznan University of Technology, Faculty of Control, Robotics and Electrical Engineering, ul. Piotrowo 3a, 60-965 Poznan, Poland

Abstract

The paper deals with the subject of the prediction of useful energy during the cycling of a lithium-ion cell (LIC), using machine learning-based techniques. It was demonstrated that depending on the combination of cycling parameters, the useful energy (\(RUE_c\)) that can be transferred during a full cycle is variable, and also three different types of evolution of changes in \(RUE_c\) were identified. The paper presents a new non-parametric \(RUE_c\) prediction model based on Gaussian process regression. It was proven that the proposed methodology enables the \(RUE_c\) prediction for LICs discharged, above the depth of discharge, at a level of 70% with an acceptable error, which is confirmed for new load profiles. Furthermore, techniques associated with explainable artificial intelligence were applied to determine the significance of model input parameters – the variable importance method – and to determine the quantitative effect of individual model parameters (their reciprocal interaction) on \(RUE_c\) – the accumulated local effects model of the first and second order.

Keywords

energy prediction, cycle life modelling, lithium-ion battery, machine learning, predictive models, useful energy prediction.

1. Introduction

In recent years, it is possible to observe a rapid increase in the demand for equipment designed for electric energy storage. At present, it is estimated that the highest increase refers to the industry branch related to electromobility, where solutions based on lithium-ion batteries are currently dominant [40]. It is predicted that by the end of 2030, the demand for batteries will increase ten times in this sector alone [13, 41]. The prevailing share of lithium-ion batteries on the market follows temperature (\(T_c\)) or ambient temperature, as well as the charge or energy throughput (expressed most frequently as the number of full equivalent cycles – FECs). At this point, it must be emphasised that some of these parameters are correlated with each other and their effect on the amount of energy throughput of LICs is non-linear [38]. The above-mentioned facts make analysing and predicting the LICs’ ability to transfer energy over the period of their lifetime highly complicated and require advanced methods [21]. In addition to this, testing electrochemical cells is very time-consuming due to their life. The testing of a single variant may last even a few years. For this reason, a more and more frequent practice is to conduct the testing under conditions which result in their accelerated degradation (e.g. under an increased load or at extreme temperatures), and then to predict their operational parameters under the assumed operating conditions [8, 12].

1.1. Related papers

Based on the current state of knowledge, it can be concluded that in the case of LICs, usually separate degradation models are developed for cycling (cycle life) and storage period (calendar aging). The calendar aging models usually include two parameters: the ambient temperature (\(T_a\)) at which a cell is stored and the state of charge (\(SoC\)) [2, 3, 23, 64]. On the other hand, degradation models during cycling are multi-parameter models – their parameters include: values of discharge/charge current (\(I_d/I_{ch}\)), depth of discharge (\(DoD\)), cell temperature (\(T_c\)) or ambient temperature, as well as the charge or energy throughput. Based on the current state of knowledge, it can be concluded that in the case of LICs, usually separate degradation models are developed for cycling (cycle life) and storage period (calendar aging). The calendar aging models usually include two parameters: the ambient temperature (\(T_a\)) at which a cell is stored and the state of charge (\(SoC\)) [2, 3, 23, 64]. On the other hand, degradation models during cycling are multi-parameter models – their parameters include: values of discharge/charge current (\(I_d/I_{ch}\)), depth of discharge (\(DoD\)), cell temperature (\(T_c\)) or ambient temperature, as well as the charge or energy throughput. This is confirmed for new load profiles. Furthermore, techniques associated with explainable artificial intelligence were applied to determine the significance of model input parameters – the variable importance method – and to determine the quantitative effect of individual model parameters (their reciprocal interaction) on useful energy prediction. The paper presents a new non-parametric prediction model based on Gaussian process regression. It was proven that the proposed methodology enables the \(RUE_c\) prediction for LICs discharged, above the depth of discharge, at a level of 70% with an acceptable error, which is confirmed for new load profiles. Furthermore, techniques associated with explainable artificial intelligence were applied to determine the significance of model input parameters – the variable importance method – and to determine the quantitative effect of individual model parameters (their reciprocal interaction) on \(RUE_c\) – the accumulated local effects model of the first and second order.

E-mail addresses: D. Burzyński (ORCID: 0000-0001-7874-1376): damian.burzynski@put.poznan.pl
It is possible to distinguish two main approaches in the modelling of LICs’ degradation: physico-chemical models and empirical models. The physico-chemical models usually consist of several partial differential equations, by means of which, the effect of the respective aging processes on the capacity fade of LICs is modelled. Examples of such models, for the Li2CoO2 cell can be found in papers [37, 45, 50]. The advantage of physico-chemical models is the possibility of mapping changes taking place in each part of a LIC as a consequence of cycling, however, the identification of their parameters and implementation are highly complicated. A frequent approach among researchers is the use of methods based on regression [16, 36, 48, 53, 58]. In such models, in the majority of cases, the effect of selected cycling parameters on LIC degradation is analysed. A paper by [15] studied the effect of certain parameters of the charging half-cycle, the charge currents (ICh), and final charge voltages (UfCh), on the process of the actual capacity fade of a LiCoO2 cell. The testing was conducted at one charge current level (1C) and at a temperature of 25°C. Xiong et al. [67] developed a model which enables prediction of the remaining useful life of a battery (RUL) for a LiNiCoAlO2 cell. The tests were conducted under conditions of accelerated aging – at discharge currents I0 = 1C and I0 = 2C and at temperatures T0 = 25°C and T0 = 40°C. Effect of fluctuating ambient temperature on the capacity loss process of LiFePO4 cell was investigated and modeled in the paper [33]. In paper [35] Dessaint implemented two multi-parametric models which allow for the determination of the maximum number of cycles which may be carried out by LiNiMnCoO2 and LiFePO4 cells, depending on the actual cycling parameters (ICh, I0, DoD, T0). On the other hand, in paper [57] Wang et al. took into account the charge throughput in addition to typical cycling parameters (ICh, I0, DoD, T0) to determine the capacity fade of a LiFePO4 cell. Among other approaches related to the modelling of degradation during LIC cycling, it is necessary to mention the destructive method based on the Palmgren-Miner damage accumulation theory [11,46] and the non-destructive method based on electrochemical impedance spectroscopy (EIS). Zhang et al. [69] proposed a forecasting methodology for the RUL of lithium-ion batteries using EIS and Gaussian process regression (GPR). In turn, Saha et al. in [47] used EIS and the Bayesian statistical approach to develop a new RUL method.

Moreover, as methods based on machine intelligence are being developed, it is possible to observe their more and more frequent use in problems related to the prediction of operational parameters of LICs such as, e.g. remaining useful life (RUL) [17, 32, 51] or capacity [43, 49]. These methods are strictly based on data which are usually acquired from many-months of experimental procedures [5, 10, 22, 28, 55]. In [19], Hannan et al. proposed the use of a deep neural network to determine the SoC for LiNiMnCoO2 cells. Li et al. [29] developed a method for estimating the capacity of LiFePO4 cells, based on a convolutional neural network. Moreover, the use of support vector machines was proposed in papers by [26, 63] to predict the state of health (SoH) of batteries. In [24], Li et al. proved that the SoC of a battery may be predicted with high accuracy by a structure which is a merger of the neural network and fuzzy logic. In the past, approaches based on sample entropy [20], fuzzy logic [6] and Rao-Blackwellization particle filter [14] have also been used to determine the RUL.

The GPR technique used in that paper had previously been used to predict the capacity fade or RUL. In [44], Howey et al. developed a model to predict the capacity fade of a LiCoO2 cell under variable load and temperature conditions. On the other hand in [52] Haritharan et al. used the deep GPR to estimate the end of life (EoL) – the chemical composition of the cell was not specified. GPR was also used to predict battery degradation during calendar aging, and this issue was touched upon in paper by [31].

Because of the fact that, for some designers of systems powered from cells, SoH is not always a sufficient assessment indicator, other methods which enable the determination and prediction of the available power and energy in LICs have been developed recently [39, 65]. In [66], an attempt was made to predict the energy available in the battery for two cells - LiNiMnCoO2 and LiFePO4, using particle filtering. In a paper by [25], the state of the available energy at different 2L and at different T0, taking into account degradation, was determined for a Li1/2Ti2O3 cell using a circuit model. Dong et al. [9] developed a state of energy estimation method for LiFePO4 cells using neural network achieving an error of less than 4%. The analyses were conducted at different ambient temperatures and discharge currents. In the papers [30, 68] methods based on predictive control theory were implemented to predict the energy of an LiNiMnCoO2 cell and a package of LiNiMnCoO2 cells. Another approach using an adaptive method based on an extended Kalman filter to estimate the remaining energy of LiFePO4 and LiNiMnCoO2 cells are presented in [60, 61]. Wang et al. [59] proposed a joint estimator based on particle filter to determine both the state of energy and SoC of the LiFePO4 cell. It should be noted here that the cell parameters studies were conducted with variation in discharge currents and at four ambient temperature levels. Fractional-order physics models can be also used to determine the state of energy with less than 5% error, as demonstrated in the paper [27].

1.2. Key contributions

Although many different approaches have been developed to date in the aspect of modelling the effects of aging phenomena occurring in LICs, in particular the prediction of available energy and power, this issue has not yet been sufficiently explored. In many applications, not only is knowledge of the predicted useful energy of the LIC over its lifetime required, but also of the predicted changes in degradation trends. In many research centres around the world, complex models are currently being developed that take into account many parameters of cyclic operation, such as ambient temperature, discharge current and, increasingly SoH. For this reason, the following contributions have been made in this paper:

1) A major contribution is the development of a machine learning based model that allows long-term prediction of the useful energy that the LIC is able to transfer during a single duty cycle taking into account its lifetime degradation which is a significant improvement of currently existing methods. The proposed model has a new structure of input parameters – given the negative effect of higher currents during charging proven in the literature [15], the model separately considers the effect of discharging and charging current.

2) The result of this research is a new dataset that can be used to test current methods and develop new predictive methods for the useful energy or SoH of LICs.

3) A demonstration of the variability of RUE, depending on the combination of cycling parameters, and also on the temperature and the state of charge, was determined for a LiNiMnCoO2 cell. Li et al. [29] developed a method for estimating the capacity of LiFePO4 cells, based on a convolutional neural network. Moreover, the use of support vector machines was proposed in papers by [26, 63] to predict the state of health (SoH) of batteries. In [24], Li et al. proved that the SoC of a battery may be predicted with high accuracy by a structure which is a merger of the neural network and fuzzy logic. In the past, approaches based on sample entropy [20], fuzzy logic [6] and Rao-Blackwellization particle filter [14] have also been used to determine the RUL.

1.3. Structure of the paper

The remaining part of the paper is organised as follows. Chapter II outlines the procedure for obtaining experimental data and determining the useful energy RUE, and also describes the results of preliminary analyses and calculations. Chapter III presents the applied methodology, GPR and XAI techniques. The results obtained and the discussion of the verification procedure are included in chapter IV. The final remarks and the conclusion are presented in chapter V.
2. Data preparation

In order to obtain data related to the ability of LICs to transfer energy during cycling, they were subjected to aging tests under various load and temperature conditions. Commercial Samsung 18650 LICs (cylindrical), with a nominal capacity of 2600 mAh ($C_{\text{nom}}$) and nominal voltage of 3.63 V ($U_{\text{nom}}$) were selected for the tests. The cathode material of the tested cells was composed of LiNi$_{0.33}$Mn$_{0.33}$Co$_{0.33}$O$_2$ compound, while the anode was made of graphite.

2.1. Experimental procedure

The procedure presented in Fig. 1 was adopted for the purpose of performing aging tests. Before starting the aging procedure of each cell, their reference charge ($Q_{\text{ref}}$) was determined in accordance with algorithm 1.

Then, each of the cells was subjected to cycling under unique temperature and load conditions. During the tests, a Panasonic MIR-254 temperature chamber was used to maintain a constant ambient temperature. The $T_a$ range in which the cells were cycled was between 10°C and 40°C. Momentary temperatures of cells were recorded by means of probes located in central points of the cells. The discharge half-cycles were completed using the constant-current (CC) method, while charge half-cycles were completed using the constant-current constant-voltage (CC-CV) method. Values $I_d$ and $I_{ch}$ were selected from the range between -2.6 A and -7.8 A (for discharge) and between 1.3 A and 7.8 A (during the CC charge phase). The average charge current from all half-cycles completed from $BoL$ ($SoH = 100\%$) to $EoL$ ($SoH = 80\%$) was adopted as the value of charge current during analyses. During the experimental procedures, LICs were kept in the voltage operating window specified by the manufacturer, i.e. from 2.75 V ($U_{d\text{sch}}$) to 4.2 V ($U_{c\text{h}}$). Several DoD levels were selected ranging between 16% and 100%, with an assumption that DoD is the charge which can be obtained in the given combination of cycling parameters. For example DoD= 100% corresponds to a discharge in the full voltage operating window. The interval between charge and discharge half-cycles was constant and was equal to 15 s. The cell cycling process was controlled by a Cadex C8000 tester dedicated to electrochemical cells (values of momentary currents, voltages and temperatures of LICs were measured using separate wires).

The current/voltage measurement accuracy amounted to ±0.001 A/V while the temperature measurement accuracy was ±0.1°C. To sum up, 29 cells in total were tested at four different $T_a$ (10°C, 15°C, 25°C and 40°C), at 3 different values of $I_d$ (-2.6 A, -5.2 A and -7.8 A) and at 5 different DoD levels (16%, 27%, 50%, 75% and 100%). The average value of charge current ($I_{\text{ch\_avg}}$) ranged between 1.2 A and 3.46 A. Table 1 contains detailed values of the cycling parameters adopted during the aging tests.

The condition for completion of each aging procedure was the degradation of the cell to $SoH = 80\%$. The actual $SoH$ of the cells was checked every 48 hours under reference conditions (described in algorithm 1) using the following relationship (1):

$$SoH = \frac{Q_{\text{ref}}}{Q_{\text{BoL}}} \times 100\%$$

Table 1 contains detailed values of the cycling parameters adopted during the aging tests.

![Fig. 1. Block diagram of the experimental procedure](image)

**Algorithm 1: Q_{ref} determination**

1. Interruption of the currently running aging variant after completion of the charge half-cycle (not applicable to BoL)
2. Set the temperature of the cell to 25°C
3. Charging of the cell to 4.2 V with 1.3 A current in accordance with procedure CC-CV until the current in the CV phase drops to 0.065 A
4. 15-second break
5. Discharge of the cell to 2.75 V with the -2.6 A current (CC method)
6. Calculation of the capacity of the cell using the Coulomb counting method - equation (2)

**TABLE 1:

| Cell ID | $T_a$ (°C) | $I_d$ (A) | $Q_{\text{BoL}}$ (Ah) | $Q_{\text{ref}}$ (Ah) | $Q_{\text{dsch}}$ (Ah) | $Q_{\text{ch}}$ (Ah) | $Q_{\text{dsch}}$ (Ah) | $Q_{\text{ch}}$ (Ah) | $Q_{\text{dsch}}$ (Ah) | $Q_{\text{ch}}$ (Ah) |
|---------|----------|-----------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| A01     | 10       | -2.6      | 2.89                 | 0.91                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 |
| A02     | 15       | -2.6      | 2.89                 | 0.91                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 |
| A03     | 25       | -2.6      | 2.89                 | 0.91                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 |
| A04     | 40       | -2.6      | 2.89                 | 0.91                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 | 1.63                 | 0.56                 |

2.2. Useful energy determination

In order to determine the useful energy throughput of a LIC during a single duty cycle over its entire life, the following methodology was adopted:

1. The assumption that one complete duty cycle consists of a discharge half-cycle and the immediately following charge half-cycle.
2. Calculation of the charge throughput during a single duty cycle (collected during discharge $Q_{\text{dsch}}$ and charge $Q_{\text{ch}}$) using the Coulomb Counting Method, on the basis of equation (2).
3) Determination of the relative useful energy \( RUE_c \) throughout during the duty cycle in accordance with relationship (3). \( RUE_c = 1 \) means that a cell transferred nominal energy (18.876 Wh) during a full duty cycle (a discharge followed by a charge).

4) After each control procedure, determination of the number of completed full equivalent duty cycles (FECs) for the needs of the prediction model.

5) Determination of \( FEC_{\text{cmax}} \) as the sum of \( RUE_c \) from all completed LIC duty cycles from BoL to EoL. An assumption was made that the transfer of nominal energy (18.876 Wh) during two consecutive charge and discharge half-cycles constitutes a full equivalent cycle (FEC).

6) Determination of the total energy throughput \( E_{\text{th_tot}} \) as a product of total charge throughput of LIC, multiplied by its voltage on terminals:

\[
Q_{\text{ch/dsch}} = Q_0 + \int_{0}^{t} I_{\text{ch/dsch}}(t)dt
\]

\[
RUE_c = \frac{Q_{\text{dch}} \cdot U_{\text{ave_dsch}} + Q_{\text{ch}} \cdot U_{\text{ave_ch}}}{2 \cdot C_{\text{nom}} \cdot U_{\text{nom}}}
\]

where \( Q_0 \) is the initial charge accumulated in a LIC, \( U_{\text{ave_dsch}} \) is an average voltage during a discharge half-cycle, \( U_{\text{ave_ch}} \) is an average voltage during a charge half-cycle, \( C_{\text{nom}} \) is the nominal capacity and \( U_{\text{nom}} \) is the nominal voltage.

2.3. Initial analysis and calculations

The tests carried out showed that for the LICs discharged to DoD > 70\% the value of the transferred \( RUE_c \) is variable during their considered life. In the case of variants with DoD < 70\%, LICs transferred the fixed energy throughout their life. Depending on the combination of the values of the cycling parameters, three types of evolution of \( RUE_c \) changes were observed during the analysed period of operation: a) approximately linear, b) a slow decrease in the first phase of life followed by a phase of accelerated loss of energy transfer, c) a characteristic inflection point occurred after the phase of releasing the assumed charge, followed by a phase of rapid loss of energy transfer. The phenomena described in points b) and c) were present in cells “A23”, “D14”, “D16”, “A30”, “H10”, “A20”, “B37”, “H5” and “C33”. Moreover, it was observed that the lower the average temperature during full duty cycle (\( T_c \)), the more rapidly the process of \( RUE_c \) loss proceeded. For LICs with \( T_c > 30^\circ C \) in most cases (except “B35”, “D16” and “A23”) this phenomenon proceeded approximately linearly. The selected characteristics of the \( RUE_c \) of cells, in which the effect described above occurred, are presented in Fig. 2.

For each completed aging variant, it was checked how the ability of the cell to transfer \( RUE_c \) decreased. For this purpose, the \( \Delta RUE_c \) difference (eq. 4) between \( RUE_c \) determined at BoL and EoL was calculated.

\[
\Delta RUE_c = \left( RUE_c_{\text{BoL}} - RUE_c_{\text{EoL}} \right) \times 100\%
\]

3. Proposed methodology

3.1. Data pre-processing

An inseparable part of prediction models based on machine learning is data pre-processing. The techniques used in data pre-processing enable assessment of the quality of the data and their usefulness, which translates directly into the ability of the created model to learn. Due to the numerical nature of the data used, it was not necessary to use methods related to the handling of categorical variables and null values. Before initiating the machine learning procedure, the data obtained from the experimental measurements were normalised. After the normalisation, the average value of each of the parameters of the learning dataset was equal to 0, and the standard deviation was equal to 1.
3.2. Proposed model

The amount of RUE, which LICs are capable of transferring during a full duty cycle over their lifetime depends on the combination of values of cyclic operating parameters and the number of completed FECs. As the tests demonstrated, LICs may have a similar ability to transfer RUE, even if they have completed a significantly different number of FECs. One additional issue is the fact that the cyclic operating parameters of a LIC are strongly correlated. For instance, higher values of DoD, I_d and I_d,avg result in an increase in the average temperature of the LIC, which, in turn, determines the amount of RUE. Furthermore, the impossibility of separating the impact of the above-mentioned parameters during a duty cycle makes it highly difficult to determine the precise effect of individual parameters on the amount of RUE. In this paper, in order to develop a prediction model for the relative useful energy RUE, the framework presented in Fig. 3 was adopted.

Taking the above-mentioned challenges into account, the author proposes the use of the non-parametric model, whose structure strictly matches the form of learning data. This model belongs to the group of machine learning models (with supervision); it is more frequently referred to as GPR and is a non-parametric regression technique based on Gaussian processes. This approach has many advantages, the most important being the lack of requirements to know the distributions of the model parameters and the correlations between them, and also a high prediction accuracy even when using a small learning dataset. Additionally, it forces out the use of techniques which enable the interpretation of the manner in which the implemented non-parametric model operations. These techniques belong to the so called explainable artificial intelligence XAI. In this paper, the author has applied a method that makes it possible to determine the significance of explanatory variables in the model. For the purpose of quantitative determination of the effect which the respective model variable has on the predicted value (taking into account the correlation of other parameters), the concept of accumulated local effects (ALE) has been used.

3.3. Gaussian process regression

Gaussian process regression (GPR) belongs to the group of kernel-based non-parametric models [42]. In this model, the Gaussian probability distribution is defined for each finite set of input variables x:

\[ f(x) \sim GP(\mu(x), \text{cov}(x, x')) \]

(5)

where \( \mu(x) \) is mean function, and \( \text{cov}(x, x') \) is the covariance function (kernel function).

The mean and covariance functions are expressed as:

\[ \mu(x) = E(f(x)) \]

(6)

\[ \text{cov}(x, x') = E((f(x) - \mu(x))(f(x') - \mu(x')) = k(x, x') \]

(7)

where E is an expected value.

The mean function may be equal to zero and may be a constant or mean value of the learning dataset. On the other hand, the kernel function may have various specific forms ranging from the standard ones, such as constant, linear, radial, squared exponential or matern, as well as being a composite of multiple kernel functions. The kernel functions contain parameters related to the scaling of responses of model y and input vector x and are referred to as hyperparameters \( \theta \). For the majority of standard kernels, the hyperparameters include: standard deviation \( \sigma_f \) and characteristic length \( \sigma_x \). For all the parameters of the input vector of the model, the characteristic length may be identical or different.

The matching of the structure of the model to the learning set takes place by defining, for each sample of the learning dataset \( x_i \), a function \( f(x_i) \) having the Gaussian distribution – hereafter called a hidden variable – and a set of basis functions \( h(x) \), which transfer the input vector \( x_i \) from the original feature space \( R^d \) to the extended feature space \( R^p \). The graphic representation of the GPR model is presented in Fig. 4. The known values (learning dataset) are marked in circles with continuous lines, while the hidden variables of the GPR model are mapped using rectangles. The hidden variables are connected by a thin horizontal line and, taking into account the marginalisation property of the model, each data sample \( (x_i, y_i) \) is independent of others. Points \( (x^*, y^*) \) in circles with dashed lines mean new data points.

By using the kernel functions, it is possible to determine how model prediction \( y \) for vector \( x \) is dependent on the response in other points \( x' \). It determines the similarity between the two points \( (x, x') \) in the scalar form \( k(x, x') \). More details related to kernel functions can be found in [18].

For the needs of the RUE prediction, the following structure of the model has been adopted:

\[ y(x) = h(x)^T w + f(x) + \varepsilon \]

(8)

where \( h(x) \) is the set of basis functions, \( w \) is the vector of basis function coefficients, \( \varepsilon \) is noise with normal distribution.

Model prediction \( y \) takes place on the basis of input vector \( x' \) and the learning dataset:

\[ P(y|f(x_i), x_i) \sim N \left( y_{i} | h(x_i)^T w + f(x_i), \sigma^2 \right) \]

(9)

The above-mentioned model may be expressed in the form of a vector using the following relationship:

\[ P(y|f, X) \sim N \left( y | Hw + f, \sigma^2 \right) \]

(10)

where \( \sigma \) is the standard deviation, \( I \) is the identity matrix and:

\[ X = \begin{bmatrix} x_1^T & x_2^T & \cdots & x_n^T \end{bmatrix}^T \]

(11)

\[ f = \begin{bmatrix} f(x_1) \cdot f(x_2) \cdots f(x_n) \end{bmatrix}^T \]

(12)

\[ H = \begin{bmatrix} h(x_1^T) \cdot h(x_2^T) \cdots h(x_n^T) \end{bmatrix}^T \]

(13)

\[ y = \begin{bmatrix} y_1 \cdot y_2 \cdots y_n \end{bmatrix}^T \]

(14)

Joint probability distribution of latent variables \( f \) with an assumption of a zero mean function \( \mu(x)=0 \) is expressed as follows:
where:

$$P(f|X) = N(f|0, K(X, X'))$$  \hspace{1cm} (15)$$

and:

$$K(X, X') = \begin{pmatrix} k(x_1, x_1') & k(x_1, x_2') & \cdots & k(x_1, x_n') \\ k(x_2, x_1') & k(x_2, x_2') & \cdots & k(x_2, x_n') \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1') & k(x_n, x_2') & \cdots & k(x_n, x_n') \end{pmatrix}$$  \hspace{1cm} (16)$$

The values of the hyperparameters of the kernel function can be determined by maximisation of the logarithmic membership function $P(y|X)$ with respect to parameters $w, \theta$ and $\sigma^2$ using the following relationship:

$$w, \theta, \sigma^2 = \arg \max_{w, \theta, \sigma^2} P(y|X, w, \theta, \sigma^2)$$  \hspace{1cm} (17)$$

Taking into account that:

$$P(y|X, w, \sigma^2) = N(y|Hw, K(X, X') + \sigma^2 I)$$  \hspace{1cm} (18)$$

the logarithmic membership function may be represented as:

$$\log P(y|X, w, \sigma^2) = -\frac{1}{2} \log \left| K(X, X') + \sigma^2 I \right| + \frac{1}{2} \log |y - Hw|^2$$  \hspace{1cm} (19)$$

After determining the parameters of the model, the probabilistic prediction of new values $y_{\text{new}}$ for a new input vector $x_{\text{new}}$ can take place as follows:

$$P(y_{\text{new}}|y, x, x_{\text{new}}) = \frac{P(y_{\text{new}}|y, x_{\text{new}})P(y|x, x_{\text{new}})}{P(y|x_{\text{new}})}$$  \hspace{1cm} (20)$$

where:

$$P(y_{\text{new}}|y, x_{\text{new}}) = \left[ P(y_{\text{new}}|f_{\text{new}}, x_{\text{new}}) f^T(f_{\text{new}}|x_{\text{new}}) df_{\text{new}} \right]$$  \hspace{1cm} (21)$$

The individual components of the above integral can be represented as follows:

$$P(y_{\text{new}}|f_{\text{new}}, x_{\text{new}}) = N(y_{\text{new}}|h(x_{\text{new}})^T w + f_{\text{new}}, \sigma^2_{\text{new}})$$  \hspace{1cm} (22)$$

and:

$$f^T(f_{\text{new}}|x) = \frac{1}{\sigma^2} \left( K(X, X') + \sigma^2 I \right)^{-1} (y - Hw)$$ \hspace{1cm} (23)$$

$$P(f_{\text{new}}|f, x, x_{\text{new}}) = N(f_{\text{new}}|K(s_{\text{new}}, x)K(X, X)^{-1} f, \Delta)$$  \hspace{1cm} (24)$$

where:

$$\Delta = k(x_{\text{new}}, x_{\text{new}}') - K(s_{\text{new}}, x)K(X, X)^{-1} K(X, x_{\text{new}}')$$  \hspace{1cm} (25)$$

Finally the probability density of prediction $y_{\text{new}}$ at the new point $x_{\text{new}}$, given with known $y$ and $X$ is given by the following equation:

$$P(y_{\text{new}}|y, x_{\text{new}}) = N(y_{\text{new}}|h(x_{\text{new}})^T w + \rho, \sigma^2_{\text{new}} + \Sigma)$$  \hspace{1cm} (26)$$

where:

$$\rho = K(x_{\text{new}}, x) \left( K(X, X') + \sigma^2 I \right)^{-1} (y - Hw)$$  \hspace{1cm} (27)$$

$$\Sigma = k(x_{\text{new}}, x_{\text{new}}') - K(s_{\text{new}}, x)K(X, X)^{-1} K(X, x_{\text{new}}')$$  \hspace{1cm} (28)$$

3.4. Variable importance

Variable importance is the technique which allows for the determination of the significance of each explanatory variable in reference to its effect on the implemented machine learning model [34,62]. This concept consists of the determination of the model error by permutations of the values of each variable. The given variable is significant if changes in its value cause changes in the model error, as the model relied on this variable for prediction. If the permutations of the variable do not cause changes in the model error, then it is considered insignificant. In other words, the more the model relies on a variable to make predictions, the more important it is for the model. The values of significance of the variables are determined on the basis of algorithm 2.

3.5. Accumulated local effects

Accumulated local effects (ALE) is the method which allows for the assessment of the relationship between the explanatory variables of the model and the predicted value [1,34]. This concept is based on conditional variable distributions. This means that mean differences in prediction and not their mean values are taken into account (as is the case with the partial dependence (PD) concept [34]). ALE reflects the way in which model forecasts change within the narrow range of values of variables. This method is particularly useful in the case of sets whose parameters are correlated with each other. The ALE function is defined as follows:

$$\text{ALE}(x_j) = \frac{1}{n} \sum_{i=1}^{n} \left[ f(x_j, x_i) - f(x_{\text{new}}, x_i) \right]$$  \hspace{1cm} (29)$$

where partial derivative $\frac{\partial f(x_j, x_i)}{\partial x_j}$ represents the local effect of $x_j$ on prediction function $f$ at $(x_{\text{new}}, x_i)$, $x_j$ is lower value of considered variable range, $\epsilon_j$ is a constant selected in order to vertically center the graph.

Averaging the local effect over the conditional distribution $p(x_{\text{new}}, x_j)$ makes it possible to separate the effect of other correlated variables located outside the considered range of values of variable $x_j$.

In order to determine local effects, the whole range of values of the given explanatory variable must be divided into multiple intervals. The "effect" is understood as the difference in prediction calculated separately for each occurrence of a variable in a given interval. All differences in the given interval are summed up and divided by the number of occurrences – this way, the obtained mean difference in predictions is the "local" effect. Ultimately the value of the accumulated local effect of a given variable is determined by summing up all the local effects and represents the average change in the prediction when the value of the variables changes within its range.
Algorithm 2: Variable importance determination

Input:
- \( X \) – variables matrix (explanatory variables)
- \( \gamma \) – vector of predicted values of original data
- \( f \) – trained model
- \( L(y,f) \) – loss function (any model performance method)

Steps:
1. Compute the original model error \( L^0 = L(y,f(X)) \)
2. For each explanatory variable \( X_j \):
   a. In order to break the relationship between variable \( X_j \) and true prediction \( y \) – generate feature matrix \( X_j \) by permuting variable \( X_j \) in the data \( X \).
   b. Compute model predictions \( Y \) based on the permuted data
   c. Calculate error based on the predictions of the permuted data \( L_j = L(Y,f(X_j)) \)
   d. Calculate variable importance \( V_{j} = L_j - L_0 \)
3. Sort explanatory variables by descending \( V_j \)

Output:
- \( V_j \) – variable importance of each explanatory variable

Numerically, ALE for a single variable can be calculated according to the following relationship:

\[
ALE(x_i) = \sum_{k=1}^{k(x_i)} \left[ f(z_{k,i},x_i^{(k)}) - f(z_{k-1,i},x_i^{(k)}) \right] - c_1 \tag{30}
\]

where \( z_i \) represents the boundary of the range of variable \( x_i \), \( k(x_i) \) is the index of the interval to which a given point of variable \( x_i \) belongs, \( n(k) \) means the number of points within the interval under consideration, and \( f \) is the prediction function.

The ALE method also allows for analysis of changes in the predicted value in the case of interaction of two (ALE second-order) or more variables (ALE higher-order). In this case, the main impact of the variables concerned is ignored – only the effect of their interaction is calculated. For two explanatory variables, the ALE function may be represented as:

\[
ALE(x_1, x_2) = \frac{1}{n_1 \cdot n_2} \sum_{i=1}^{n_1 \cdot n_2} \left[ f(z_{1,i},z_{2,i},x_1, x_2) - f(z_{1,i-1},z_{2,i-1},x_1, x_2) \right] \tag{31}
\]

and in the numerical form, the above relationship may be presented as follows:

\[
ALE(x_1, x_2) = \sum_{k=1}^{k(x_1,x_2)} \left[ f(z_{k,1},z_{k,2},x_1, x_2) - f(z_{k-1,1},z_{k-1,2},x_1, x_2) \right] \tag{32}
\]

where:

\[
\Delta f_{1,2}(K,k,m,x_1^{(k)}, x_2^{(k)}) = \left[ f(z_{k,1},z_{m,2},x_1, x_2) - f(z_{k-1,1},z_{m-1,2},x_1, x_2) \right] \tag{33}
\]

The rules for the determination of ALE for two variables are identical to those for one variable, whereby, the intervals are replaced by rectangular areas due to the necessity of accumulation of local effects in two dimensions. ALE can also be calculated for the interaction of three or more variables – the relevant relationship can be found in paper [1].

4. Results and discussion

The input parameters of the model were the cyclic operating parameters: the average LIC temperature during the full duty cycle (\( T_c \)), the value of current during the discharge half-cycle (\( I_d \)), depth of discharge (DoD), the value of mean current during the charge half-cycle (\( I_{d, avg} \)). The total aging processes leading to LIC degradation were included in the model by the number of full equivalent cycles (FEC). The predicted value was the relative useful energy (RUEc) throughout a cell during a full duty cycle.

4.1. Model training

The results of measurements of cycling of LICs, which were cycled with DoD>70% (20 LICs in total), were used as the learning dataset. Cells “A38”, “B32”, “B40” and “H9” were used to verify the model exclusively. Finally, the learning dataset contained 311 samples, and the verification set consisted of 99 samples – a fragment of the learning dataset after initial calculations is presented in Table III. The value of FEC=0 is equivalent to SOH~100% of the cell, while the last value of FEC=533 indicates SOH~80% and the accomplishment of the EoL status.

As the kernel function of the GPR model, the matern 3/2 function was used, which is expressed by the following equation:

\[
k_{matern3/2}(x,x') = \sigma_f^2 \left[ 1 + \frac{\sqrt{3r}}{\sigma_f} \right] \exp \left( -\frac{\sqrt{3r}}{\sigma_f} \right) \tag{34}\]

where \( r \) is Euclidean distance between \( x \) and \( x' \).

| \( T_c \) | \( I_d \) | DoD | \( I_{d, avg} \) | FEC | RUEc |
|---|---|---|---|---|---|
| 36.6 | -7.8 | 100 | 1.74 | 0 | 0.905 |
| 36.6 | -7.8 | 100 | 1.74 | 41 | 0.871 |
| 36.6 | -7.8 | 100 | 1.74 | 504 | 0.371 |
| 36.6 | -7.8 | 100 | 1.74 | 533 | 0.221 |

The values of characteristic parameters \( \sigma_l \) and \( \sigma_f \) were determined using the method of optimisation without constraints (quasi-Newton) and were 0.1687 and 0.7659 respectively, while the value of the logarithmic membership function was \( L=713.7 \). The linear function type was adopted for the function set \( h(x) \). The coefficient of determination \( R^2 \), mean average percentage error \( MAPE \) and root mean square error \( RMSE \) were used to assess the model quality. The above indicators are determined according to the following relationships:

\[
R^2 = \frac{\sum_{i=1}^{n} (f_i - \overline{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y}_i)^2} \tag{35}\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - f_i}{y_i} \right| \cdot 100\% \tag{36}\]
\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - f_i)^2} \]  

where \( n \) is the number of samples, \( f \) is the predicted value, \( y \) is the measured value.

The results of the learning process were compared with three supervised machine learning models: a regression tree, a support vector machine (SVM), and an ensemble of regression trees and are summarized in Table IV. The same input parameter structure was used for each model, and the predictor parameter was \( \text{RUE}_c \). For the regression tree, the lowest RMSE was achieved for a leaf size of 4. It was observed that increasing the leaf size beyond this value resulted in an increase in RMSE. For the SVM-based model, the lowest RMSE was obtained for a Gaussian-type kernel with a characteristic length of 0.56. The model with second lowest RMSE was the ensemble of the regression trees consisting of a combination of 100 trees with a leaf size of 5 each and a learning rate of 0.25.

The example characteristics of the three observed types of \( \text{RUE}_c \) evolution for two models with lowest RMSE (GPR model and Ensemble of trees model) are presented in Fig. 5 (a-c). The learning results of other LICs for all comparative models are included in Appendix A.

As opposed to the model based on Ensemble of regression trees, the response of the GPR model for each LIC included in the learning dataset was within the assumed confidence interval (95%). The MAPE of the GPR model developed for all samples of the learning dataset was 0.05%. This proves that the GPR model almost perfectly matches the learning dataset.

### 4.2. Prediction

The developed original model was verified using new load profiles which were not used to train the model. LICs with different mean cycling temperatures \( T_c \) at mean load current of \( I_L = -5.2 \) A were selected for verification. For instance, for \( T_c > 40 \)°C, the learning dataset included LICs loaded with \( I_L = -2.6 \) A and \( I_L = -7.8 \) A (cells "D6" and "H12"), therefore, the “A38” cell was selected for verification at higher temperatures. The verification process consisted of the \( \text{RUE}_c \) prediction by a model of 4 cells in the period between BoL and EoL, and of comparison of their values obtained during the experiment. The verification results are presented in Table V, and the characteristics for the 4 tested cases are presented in Fig. 6. From the results obtained it can be observed that in two cases (“H9” and “B40”) APE exceeded 5%. The reason for obtaining higher APE for the indicated LICs is primarily due to the low number of samples in the learning dataset containing conditions similar to those under which they were cycled. For the “H9” LIC the problem is in the near EoL phase. During the charging half-cycle, the average charging current for this LIC was 2.85 A. The learning dataset contains 29 samples (9.3% of all samples) with similar values of cycling parameters (especially with high average charging currents) to “H9” cell. Cell “B40” performed a high number of FECs (over 1000) while in the learning dataset only 7% of all samples are with similar values (mainly coming from “C26” LIC, which performed 909 FEC). This compares to 14.1% of samples with similar values in the learning dataset for cell “A38” and 14.5% for “B32”.

In the author’s opinion, the obtained results of the verification process confirm the effectiveness of the selected GPR method for the purpose of \( \text{RUE}_c \) prediction which a LIC can transfer during a full duty cycle throughout its life. Only in one test case did the MAPE slightly exceed 5% (the “B40” cell). Also in favor of the method used, is the fact that despite an almost ideal matching of the model to the learning data, the prediction of \( \text{RUE}_c \) under new operating conditions for the LIC is possible with little error.

---

**Table IV**

| Model                  | MAPE (%) | \( R^2 \) (-) | RMSE (-) |
|------------------------|----------|---------------|----------|
| Ensemble of trees      | 0.73     | 0.95          | 0.0311   |
| **GPR**                | 0.05     | 0.99          | 0.0141   |
| Regression tree        | 3.54     | 0.85          | 0.0551   |
| **SVM**                | 2.72     | 0.89          | 0.0465   |

---

*Fig. 5. Learning results obtained for three observed types of \( \text{RUE}_c \) evolution of GPR model and Ensemble of regression trees: a) after about 300 FECs, the cell lost the ability to transfer the fixed energy – the phase of significant decrease in \( \text{RUE}_c \) takes place, b) under these conditions, throughout the lifetime of the LIC, the decrease in \( \text{RUE}_c \) was approximately linear, c) the phase of slow degradation occurs until approximately 400 FECs are completed, then it is possible to observe the phase of accelerated ability to transfer \( \text{RUE}_c \)***
The GPR model also allows for prediction within ranges exceeding the ranges of variables of the learning dataset, which is not possible for some techniques based on artificial intelligence – for instance, neural networks or fuzzy logic.

### 4.3. Model explanation

In order to determine the significance of the explanatory variables of the model, the variable importance (VI) technique was used. RMSE was selected as the loss function of the model (eq. 37). The calculations were repeated 8 times in order to negate the effect of uncertainty caused by permutations of values. The obtained results of VI are characterized by a low standard deviation; they were averaged and listed in Table VI. Based on the results, the respective model input parameters were ordered in terms of their significance - the effect of FEC on the predicted value of RUE, is the highest in the used dataset (the highest VI), then the effect of I_F and T_c is on a comparable level. From among the adopted model parameters, DoD and I_h_avg have the lowest effect on RUE, (in both cases the calculated significance is at a similar level).

Given the fact that interpretation of the structure of the developed GPR non-parametric model is very difficult, in order to determine the effect which the respective model parameters have on the value of the RUE, prediction, the author has applied the ALE concept. Fig. 7 (a–e) presents the plots of the first-order accumulated local effects which determine the impact of individual input parameters of the GPR model on RUEc. The value of ALE at the given point determines the difference in prediction in relation to the mean prediction of RUEc. ALE=0 represents the mean prediction of RUEc. For instance, for T_c ≈ 25°C, the RUEc prediction will be lower by 0.05 in relation to the mean prediction which is an exclusive consequence of temperature (the effect of other cyclic operating parameters is separated). Based on the results obtained, it is possible to determine, unambiguously, the effect of T_c on RUEc – the lowest values of RUEc are obtained at average cycling temperatures below 30°C and they increase gradually up to the range of values between 30°C and 44°C, where RUEc reaches its peak value and then falls for T_c higher than 44°C (Fig. 7 – a). In the case of values of currents I_F and I_h_avg the effect on RUEc is linear. The lower the values of I_F and I_h_avg are, the higher the RUEc is. Furthermore, cycling of the LICs at I_F values lower than -5.2 A and higher than 2.25 A for I_h_avg will result in lower RUEc values over their lifetime – the visible breaking point of characteristics (Fig. 7 – b, d). The plot of ALE for DoD showed the linear dependence of this factor on RUEc (Fig. 7 – c). The amount of transferred RUEc drops sharply when the cells complete more than 500 full equivalent duty cycles (Fig. 7 – e).

Additionally, the results obtained using the ALE method confirm the results obtained using the VI technique – the highest dispersion of ALE values (from -0.286 to 0.225) is characteristic of FEC, which is compliant with the highest value of VI for this parameter. For I_F and T_c the differences in relation to the mean prediction vary from -0.114 to 0.08 for I_F and from -0.108 to 0.027 for T_c respectively, which also correlates with the VI results. The lowest dispersion of values was observed for I_h_avg – from -0.063 to 0.031, which means the lowest effect on RUEc in the used dataset.

In order to determine the effect of interaction of the model parameters on RUEc, the values of the second-order accumu-
lated local effects were determined in the paper for all the possible combinations of pairs of the GPR model input parameters. The results obtained in the form of value maps are presented in Fig. 8 (a–j). Analysis of the interaction effect of model parameter pairs showed that the highest influence on $RUE_c$ is exerted by combinations of the following pairs $T_c$–$FEC$, $I_d$–$FEC$ and $I_{ch_{avg}}$–$FEC$. For instance, for LICs, which completed a large number of FECs (over 1500) at $T_c$ below 25°C, the differences in relation to the mean $RUE_c$ prediction even amount to -0.4, and at $T_c$ above 50°C +0.2 respectively (Fig. 8 a). A high level of influence was also shown by the $I_{ch_{avg}}$–$FEC$ pair at around $FEC=1350$ and $I_{ch_{avg}}= 3A$ – the ALE value in this area reaches up to +0.3 (Fig. 8 d). The level of influence of the $DoD$–$FEC$ pair over the entire value range does not exceed the range between -0.04 and +0.06 (Fig. 8 d). To sum up – the pairs containing $FEC$ are characterised by the highest influence due to the highest effect of $FEC$ on $RUE_c$ which is shown in Fig. 7 – e. The interaction effect of parameter pairs, i.e. $T_c$–$DoD$ and $T_c$–$I_d$ is locally at the level of hundredths (+0.02 and -0.015) – Fig. 8 – h,j. On the other hand, the effect of other pairs is at a negligible level (ALE at a level of thousandths or lower) – Fig. 8 – e-g,i.

5. Conclusion

The paper presents the author’s prediction model for the useful energy that can be transferred during a single duty cycle under various temperature and load conditions, taking into account the energy throughput to date by LiNi$_{0.33}$Mn$_{0.33}$Co$_{0.33}$O$_2$ cells.

APPENDIX

A. Results of the learning procedure – MAPE of comparative models for individual LICs.

The methodology presented in the paper allows for the obtaining of a $RUE_c$ prediction over the entire LIC life with high accuracy, even when learning datasets with a small number of samples are used. The developed model is resistant to the overfitting phenomenon – despite an almost ideal matching of the model to the learning dataset used, the prediction of trends in the evolution of $RUE_c$ is possible under new load and temperature conditions, which has been confirmed by the author with 4 test cases. The appropriate selection of types for the kernel function $k(x, x')$ and basis functions $h(x)$ was critical to achieving the high accuracy of model prediction.

In comparison to classic modelling techniques, the advantage is that the model structure does not need to be specified. In fact, this is the disadvantage of methods based on machine learning - the structures of such models and their parameters have no physical meaning. The author proposes a solution for the above-mentioned problem by using the XAI techniques, which enable determination of both the significance of model input parameters, and the quantitative determination of their effect on the predicted value of $RUE_c$.

The information obtained in this way may be used, among other things, to improve the algorithms used specifically in battery or battery pack management systems, helping to optimise their operation and extend their life. What is more, the methodology proposed by the author can be applied to electrochemical cells with different chemical compositions.
66. Xiong R, Zhang Y, He H et al. A double-scale, particle-filtering, energy state prediction algorithm for lithium-ion batteries. IEEE Transactions on Industrial Electronics 2017; 65(2): 1526-1538, https://doi.org/10.1109/TIE.2017.2733475.

67. Xiong R, Zhang Y, Wang J et al. Lithium-Ion Battery Health Prognosis Based on a Real Battery Management System Used in Electric Vehicles. IEEE Transactions on Vehicular Technology 2019; 68(5): 4110-4121, https://doi.org/10.1109/TVT.2018.2864688.

68. Zhang X, Wang Y, Liu C, Chen Z. A novel approach of remaining discharge energy prediction for large format lithium-ion battery pack. Journal of Power Sources 2017; 343: 216-225, https://doi.org/10.1016/j.jpowsour.2017.01.054.

69. Zhang Y, Tang Q, Zhang Y et al. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. Nature Communications 2020; 11(1): 6-11, https://doi.org/10.1038/s41467-017-02251-3.