Time Series Feature extraction for Lithium-Ion batteries State-Of-Health prediction

Inès Jorge [*], Tedjani Mesbahi, Ahmed Samet, Romuald Boné

Université de Strasbourg, 67000, France
Institut National des Sciences Appliquées (INSA Strasbourg), 67000, France
CNRS, ICube Laboratory UMR 7357, Strasbourg, 67000, France

ARTICLE INFO

ABSTRACT

Lithium-Ion batteries are a core component of many devices recently designed. Despite their very high performances, their use in electric vehicles involves sometimes harsh conditions, which results in a limited driving range and a variable cycle life from one vehicle to another. The goal of prognostics and health management applied to Lithium-Ion batteries in electric vehicles is to better understand the ageing mechanisms that take place during the whole cycle life of a battery through the observation of operating data. Predicting with accuracy the State-Of-Health of a battery, in relationship with usage data, is a key step in the development of electric vehicles and of the improvement of any battery powered devices. This paper proposes an approach based on the extraction of features from current, voltage and temperature curves during charge and discharge to predict the evolution of the State-Of-Health of a battery. A sliding window of cycles is used as input to a Long Short Term Memory neural network that outputs a multi-step ahead prediction of the State-Of-Health. Our study is based on several datasets, namely two major benchmarks published by the MIT and the NASA Prognostics Centre of Excellence.

1. Introduction

With a market size valued at 41.1 billion USD in 2021, Lithium-Ion (Li-Ion) batteries are receiving a lot of attention [1]. Due to their very high specific energy density (up to 200 Wh/kg), high operating temperature range (from 0° to 60 °C), low self-discharge and absence of memory effect, Li-Ion batteries are used in most portable applications and especially in Electric Vehicles (EV) [2]. Research in the transportation field has brought EV in line with the performances of thermal vehicles, and the demand for Li-Ion batteries is very likely to increase in the coming years. However, the spread of electric vehicles is slowed down by their purchase cost, limited driving range and cycle life. In the field of smartphones, the battery is identified as the major cause of failure, leading to the replacement of the product [3]. Throughout their life, Li-Ion batteries degrade with usage (driving periods) and time (calendar ageing). In automotive applications, the most common criterion is to consider a battery out of use after reaching 80% of its initial storage capacity [4]. This is referred to as the *End-Of-Life* (EOL) criterion. The *State-Of-Health* (SOH) of a battery is defined as the ratio between the storage capacity of the battery at any time ($Q(t)$) and its initial capacity (Q_{nom}).

$$SOH(t) = \frac{Q(t)}{Q_{nom}} \quad and \quad SOH(t)\% = \frac{Q(t)}{Q_{nom}} * 100 \tag{1}$$

Considering this definition, the *Remaining Useful Life* (RUL) is the operating time left to a battery (often expressed as a number of cycles) before its SOH reaches 80%. Prognostics and Health Management (PHM) in the specific field of Li-Ion batteries aims at determining the future evolution of SOH, and by extension, the RUL. Being able to predict such values with accuracy can improve many aspects in battery management. A better understanding of the ageing mechanisms means a better use of the battery, and a better confidence in the vehicle for the users. The economic aspect can also be mentioned, with lower costs linked to maintenance if the failure is predicted in advance. By increasing the cycle life of the batteries, the tension on primary resources is lowered, which induces in the end a lower environmental impact. The ultimate goal of PHM for Li-Ion batteries is to be able to link the ageing of a battery with its global operating environment. By taking operating data such as current, voltage and temperature into account in the predictive models, a feedback could be made to the user in order to provide them with an accurate estimation of the SOH of the battery and advice how to maintain it. A study of the literature about PHM of Li-Ion batteries shows that several approaches have already been developed, especially for SOH and RUL prediction [5–8]. There are two main types of approaches for PHM of batteries: *model-based*

and *data-driven models*. As this paper focuses on a data-driven model, a strong emphasis is put on the description of this type of approaches. More details on model-based approaches can be found in Refs. [5,9,10].

Data-driven models take as input operating data from a given system to predict its future behaviour. Machine learning algorithms and more specifically Neural Networks have proved very efficient in all fields concerned by predictive prognostics due to their ability to model complex and non linear phenomena. PHM models can be classified into three categories: Non-linear Auto-Regressive models (NAR), Non-linear Auto-Regressive models with Exogenous Variables (NARX), and Non-linear Regressive models with Exogenous Variables (NRX).

NAR models make use of one single vector that is given as input to the model and that is the target output. That means future values of one unique time series are determined by looking only at its past values. NAR is by far the most employed category in PHM of Li-Ion batteries. Indeed, in a great number of approaches, the prediction of future SOH values is made by studying the historical evolution of SOH. Li et al. [11] and Qu et al. [7] have both used Empirical Mode Decomposition (EMD) on SOH curves combined with Long Short Term Memory neural networks (LSTM) to predict future SOH values. More recently, Liu et al. [8] have used a combination of EMD, LSTM and Gaussian process to predict iteratively the future trend of SOH, with a measure of uncertainty, and Li et al. [12] have designed an embedded model based on Sequence to Sequence LSTM models for SOH prediction.

NAR models are very efficient in the field of PHM for batteries due to the very smooth, slow and almost linear degradation trend of SOH in experimental datasets. Indeed, the SOH of a battery describes a decreasing curve, with a steady plateau at the beginning and a knee point from which it starts dropping. From one cycle to another, the variation of SOH is very limited, and the shape of the curve facilitates short term predictions. Therefore, very accurate predictions of future SOH values from past ones can be reached with models that are dedicated to time series such as LSTM. Some NAR models are specifically designed to identify and predict capacity regeneration phenomena during Li-Ion cycling [13–15], with good results. Nevertheless, they are based on the hypothesis that the use conditions of the battery will remain the same throughout all its life, which is very unlikely to happen in the case of an electric vehicle.

Contrary to NAR models, NARX models combine the observed time series with other exogenous variables as input. In the field of PHM of Li-Ion batteries, very few approaches are based on NARX models. In [16], a multi-layer ANN for SOH and RUL prediction is described, and takes as input a combination of voltage, current, temperature and capacity measurements, which represent SOH.

NRX models are different from NAR and NARX models since they only make use of exogenous variables to predict future values of a given time series. That means that concerning Li-Ion batteries, past values of SOH are not taken into account to predict future ones. NRX models for battery PHM consist in taking only data from current, voltage, temperature or internal resistance to predict SOH, or to use current, voltage, temperature, internal resistance and SOH to predict RUL exclusively, without predicting SOH. You et al. [6] have extracted sequences of current and voltage to predict the SOH with bi-directional LSTMs. More recently, Ren et al. [17] have extracted empirical features from charge and discharge current and voltage curves. These empirical features are used as input to an Auto-CNN-LSTM network for RUL prediction. Audin et al. [18] have recently developed an approach where time series of current, voltage and temperature are used to predict future values of SOH. They make use of auto encoders to extract features from time series, and employ LSTM neural networks to make the SOH prediction. Feature extraction combined with Support Vector Regression (SVR) and Ant-Lion Optimiser (ALO) for SOH estimation of Li-Ion batteries was developed by Li et al. [19] and tested on the NASA dataset.

Concerning RUL prediction, the approach described in [20] consists in gathering several features for each cycle such as the SOH of a cell,

its internal resistance, average temperature, charging time..., to cite a few, and to use them as input vector to an ANN that makes a regression on the RUL of the cell at any given cycle.

In some approaches, the impedance of the battery is used as a feature to predict the evolution of the SOH. Rather than using operating data such as current, voltage and temperature, electrochemical impedance spectroscopy (EIS) tests are conducted in order to determine the impedance of a battery over a wide range of frequencies. Zhang et al. have developed a model based on gaussian process regression to estimate the RUL and capacity of a battery using the EIS spectrum [21]. Recently, Li et al. [22] have reviewed the existing approaches to predict the temperature of Li-Ion batteries according to the EIS, concluding that it was a promising approach, but which faces some challenges. The EIS brings valuable information about the ageing stage of a battery. However, our aim is to develop a model that could be used on board of any vehicle with non intrusive sensors to collect operating data. Most vehicles are equipped with smart Battery Management Systems (BMS) that continuously record and store operating data, which is very practical for on line estimation of the SOH of a battery.

In this paper, we describe an innovative approach to predict SOH, based on two major datasets, published by the MIT, and by the NASA Prognostics Centre of Excellence (PCoE). Current, voltage and temperature time series during charge and discharge are studied, and features are extracted from them in order to be used as input to a SOH NRX predicting model.

Our key contributions are the following: *(i)* we introduce a new purely exogenous approach based only on ageing features that are extracted from operating data exclusively, without using past SOH values; *(ii)* these ageing features are fed to an LSTM that uses all available data throughout the life of a battery, without using iterative predictions; *(iii)* we develop several models that are able to predict SOH at different horizons, from 25 cycles ahead up to 400 cycles ahead.

The remainder of this article is organised as follows: Section 2 describes the methodology proposed for SOH prediction. The experimental process is detailed in Section 3, and results are presented and discussed in Section 4. A conclusion about the work described in this article is given in Section 5.

2. Window exogenous LSTM for SOH prediction

In this paper, we introduce a deep learning model for SOH prediction, based on the extraction of features from temporal curves of current, voltage and temperature, which are used as input to a window-based exogenous LSTM. The architecture is referred to as SOH-window-XLSTM. The global framework of our model is described in Fig. 1, and the following subsections detail the principles of feature extraction, feature selection and window LSTM for SOH prediction.

2.1. Feature extraction

There are two different data types in ageing datasets of Li-Ion batteries: "historical" data and "local" time series. The first one concerns the global evolution of features that are computed at each cycle. In Fig. 2(a), the SOH of several batteries from the MIT dataset is plotted as a function of the number of cycles. The longer the lifetime of the battery, the darker the curve. In Fig. 2(b), the internal resistance (IR) of the same batteries is plotted as a function of the number of cycles, and the colour of the curves varies in the same way. For both SOH and IR, there is one single value per cycle. SOH and IR are considered as Historical Features (HF).

The second type of data concerns local time series of current, voltage and temperature that evolve with the use mode of the battery and are represented as a function of time within a given cycle. Fig. 3 shows the temporal evolution of the charging current for several given cycles of the same battery from the MIT dataset. This time there is a whole vector for one cycle, unlike HF.
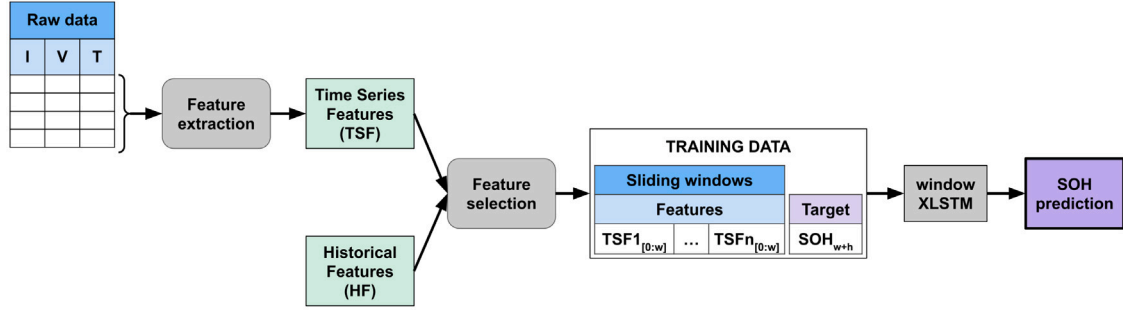
Fig. 1. SOH prediction based on the SOH-window-XLSTM model.
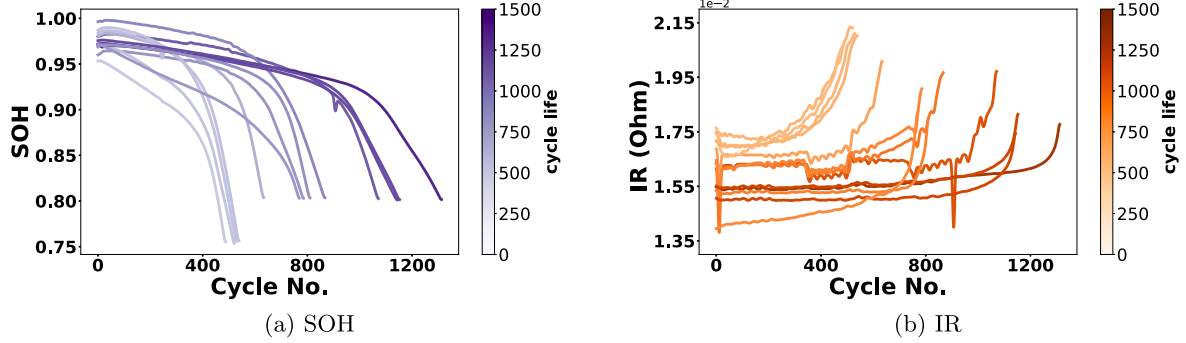

(a) SOH

(b) IR

Fig. 2. Historical features. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
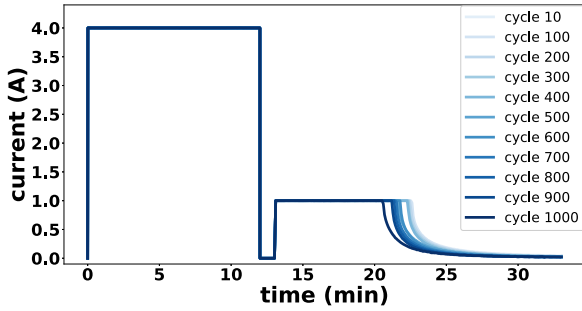

Fig. 3. Raw time series : Charge current.

Historical data and local time series do not vary on the same scale but both carry information on battery ageing. According to the literature, NAR models are very efficient and lead to very accurate predictions of SOH. However, building a PHM strategy for Li-Ion batteries should take into account all available information about the use of the battery. Studying only past values of SOH can give very accurate predictions of SOH for the coming cycles only if the use of the battery stays the same. If the use conditions vary, the prediction of SOH will not be impacted consequently when using NAR models. In this approach, features are computed and extracted from six different temporal curves: charge current, voltage, temperature, and discharge current, voltage and temperature, namely I_C, V_C, T_C, I_D, V_D and T_D. For each of these curves, and at each cycle of the life of each cell, several features are computed in the temporal, statistical and spectral domains. The list of all computed features for each of the aforementioned curves is the following, separated into domains [23]. In all equations, s represents the time series signal vector, t is the corresponding time vector and N is the length of s.

2.1.1. Temporal domain
- **Total Energy:** $\frac{\sum_{i=1}^{N} s^2}{t_N - t_0}$
- **Area Under the Curve (AUC):** $\sum_{i=1}^{N} (t_i - t_{i_1}) * \frac{s_i + s_{i-1}}{2}$

2.1.2. Statistical domain
- **Mean** : $\frac{1}{N} \sum_{i=1}^{N-1} s_i$
- **Root Mean Square (RMS):** $\sqrt{\frac{1}{N} \sum_{i=1}^{N} s_i^2}$
- **Maximum** : the largest value in s
- **Minimum** : the smallest value in s

2.1.3. Spectral domain
The Fast Fourier Transform of the signal is computed ($freq, fmag = fft(t, s)$), from which several features are extracted:

- **Fundamental frequency** : the lowest frequency of the Fourier transform
- **F_{max}** : the maximum frequency of the Fourier transform
- **Power Bandwidth (PBW)** : the width of the frequency interval in which 95% of the power of the signal is located

By reducing I, V and T° time series to a combination of several scalar features, local time series can then be represented at the same scale as HF, in the form of Time Series Features (TSF). For example, the global evolution of the RMS value of charge current is shown in Fig. 4 over the whole cycle life of different cells, just as HF can be represented. Every curve corresponds to a different battery, and the darker the curve, the longer the cycle life.

2.2. Feature selection

After computing features from the original time series, each cycle is represented by a vector of features that includes TSF and HF. This vector has a very high dimensionality, because the same number of
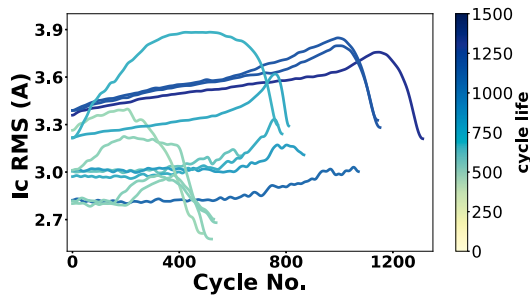
Fig. 4. Example of a Time Series Feature : I_C RMS.

features is computed from each of the six time series described earlier. A common practice is to reduce the dimensionality of the input vector of a learning algorithm by selecting a subset of features that best represent the predicting problem [24].

The approach described in this paper uses wrappers for selecting features [25]. Wrappers have the advantage of evaluating the relevance of features according to the performance of a predictive model. The wrapper technique for feature selection uses the training process on a given machine learning model to select the best combination of features. The aim is to obtain the best possible performances with a given algorithm by testing iteratively different subsets, numbers and combinations of features. The strategy is to start with an initial set of features and to add or remove several features after each training process in order to study the impact on the performances. This process is called the Sequential Feature Selection (SFS). After the feature selection process, 10 features are kept, mainly coming from charge curves. More precision on the selected features can be found in the Appendix.

2.3. Long Short Term Memory

LSTMs belong to the family of RNNs. RNNs have the ability to process data sequences by keeping track of their past elements. This means the output of an RNN layer is computed as a function of the sequence at time t, but also as a function of the output of the RNN cell several steps before (referred as the *state, h*).

$$output(t) = f_{RNN}(h(t-1), i(t)) \tag{2}$$

Function f_{RNN} represents the global operation taking place inside the RNN cell. It involves several weight matrices linked to the transformation of the previous state and input state, a bias vector, and the *tanh* activation function.

However, simple RNN layers cannot keep track of very long sequences and lack the ability to learn long-term dependencies. This is called the vanishing gradient problem, theorised by Bengio et al. in [26], and studied in [27]. LSTM addresses this problem by adding a data flow that carries information across time steps. The output of an LSTM cell at time *t* is then computed according to three values: the input $i(t)$, its previous state $h(t-1)$ and the memory value that is carried along the time sequence $C(t)$.

$$output(t) = f_{LSTM}(h(t-1), i(t), C(t)) \tag{3}$$

Function f_{LSTM} represents the global operation taking place inside the LSTM cell. It involves several weight matrices linked to the transformation of the previous state, input state and memory, bias vectors, *tanh* and *sigmoid* activation functions.

The next value of carried memory is computed as a function of the input and output of the cell, and past value of memory.

2.4. Sliding window approach for SOH prediction

The architecture described in the article, based on a window exogenous LSTM, makes use of past values of TSF to predict future values of SOH. Although LSTMs are built to learn long term dependencies in time series, small input windows are used in this model. Rather than giving as input the whole operating data of current, voltage and temperature coming from a battery since its beginning of life, a smaller window of several consecutive cycles is considered. This allows more flexibility in the model. There is no need to store large data sequences, and data preprocessing is made simpler because the size of the window is fixed. The window dataset that is built consists in several overlapping windows of TSF, and the output corresponds to a future value of SOH. Predictions can be made on a very short scale, or for longer time horizons. The training dataset for SOH prediction in this approach is built in a way that several predicting horizons are made possible. Future values of SOH are predicted from 25 up to 400 cycles ahead. Another advantage of this model is that an online prediction of SOH is made. Most approaches [7,28] implement iterative predictions, which consists in updating the input signal given to the predictive model with the last prediction. Our SOH-window-XLSTM is trained with "offline data", and prediction can be made "online", on unseen data corresponding to a window of 25 consecutive cycles at any moment of the cycle life of a battery cell.

3. Experiments

3.1. Public Li-Ion battery ageing datasets

The approach proposed in this paper is built on the use of two different datasets that were published by two different research institutes, at different times and with different battery chemistries tested.

The first one was published by the MIT in collaboration with Toyota engineering and with the Department of Materials Science and Engineering of Stanford university[1] in 2019 [29]. It is the largest available dataset regarding Li-Ion battery ageing. The cells used for testing are Li-Ion Iron Phosphate (LFP) / graphite cells from A123 manufacturer, model APR18650M1A. These cells have a 3.3 V nominal voltage and a 1.1 Ah nominal capacity. They can provide discharge currents up to 30 A. The cells were tested in a 30 °C climate chamber. The batteries are always discharged at a constant current of 4.4 A (4 C). The most important factor in the tests is the charging policy. Batteries are charged following a multi-step constant-current/constant-voltage (CC–CV) policy which makes it possible to reduce the charging time.

The second dataset was published earlier, in 2010 by the NASA PCoE.[2] It is known as a reference benchmark since it was the first public dataset on battery ageing. This dataset contains 6 groups of cells for a total of 34. The cycling experiment was segmented into three parts: charging, discharging and impedance spectroscopy. The charging protocol is the same for all tests, a CC–CV charging protocol with 1.5 A and a 4.2 V threshold. Various different discharging currents were used. Moreover, cells were cycled at different ambient temperatures, which influences the ageing speed.

3.2. Data structure

The window-based approach with LSTM described in Section 2 requires a specific structure of data. Any cycle in the life of a battery can be represented by a vector that combines HF and TSF as follows:

$$x_{cycle_i} = \{x_i^1, x_i^2, \dots, x_i^n\} \tag{4}$$

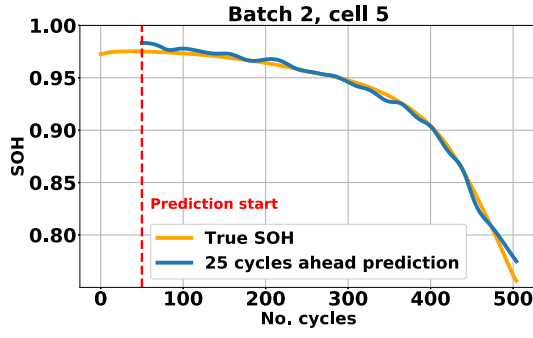where *n* corresponds to the number of HF and TSF combined.

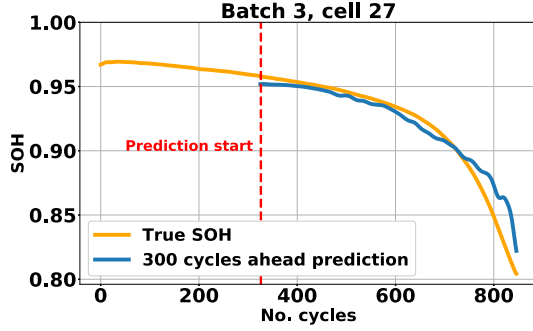**Fig. 5.** Real VS 25-cycle-ahead predicted SOH for battery b2c5.

Fig. 6. Real VS 300-cycle-ahead predicted SOH for battery b3c27.

Table 1
XLSTM performances for SOH prediction on MIT batteries.

Metric	Type	Prediction horizon (No. of cycles ahead)								
		25	50	100	150	200	250	300	350	400
MAE ($*10^{-2}$)	Average	1.05	1.2	1.4	1.3	1.9	1.6	2.0	2.1	2.4
	Best	**0.77**	1.07	1.19	0.9	1.5	1.1	1.5	1.6	2
STD ($*10^{-2}$)	Average	1.13	1.5	1.6	1.9	2.1	1.9	2.3	2.4	2.8
	Best	**0.7**	1.2	1.1	1.04	1.7	1.5	1.4	1.9	2
RMSE ($*10^{-2}$)	Average	1.5	1.9	2.1	2.3	3	2.5	3	3.2	3.7
	Best	**1.04**	1.6	1.6	1.4	2.2	1.9	2.1	2.5	2.8
RMSPE (%)	Average	1.7	2.16	2.44	2.67	3.3	2.82	3.49	3.7	4.22
	Best	**1.14**	1.8	1.8	1.6	2.6	2.14	2.5	2.9	3.14
NMSE ($*10^{-1}$)	Average	1.18	1.8	2.1	2.5	3.6	2.7	3.9	4.3	5.7
	Best	**0.48**	1.24	1.2	0.8	2.2	1.4	1.7	2.7	3.1

As the architecture involves an LSTM that is made to deal with data sequences, the input of the network is made of several consecutive cycles, as shown in (5):

$$w_j = \{x_{cycle_i}, x_{cycle_{i+1}}, \ldots, x_{cycle_{i+m}}\} \qquad (5)$$

where m corresponds to the size of the window.

Each input vector has a corresponding output y_j, which is the predicted value of SOH. The final dataset is composed of a series of overlapping windows from several batteries. After feeding all the dataset composed of a series of Z windows $\{w_1, w_2, \ldots, w_Z\}$, the LSTM produces an output vector of SOH values $\{y_1, y_2, \ldots, y_Z\}$.

3.3. Data pre-processing

In order to get rid of all noise in the data coming from measurement or experimental errors, sensor defaults or defective battery cells, every feature is cleaned. HF and TSF are submitted to various treatments such as outlier excluding and denoising. A selection of the best batteries is also done for the NASA batteries, and only 6 batteries out of the 34 initial ones are kept. The remaining 6 batteries were chosen for their long cycle life and data quality.

Table 2
BiXLSTM performances for SOH prediction on MIT batteries.

Metric	Type	Prediction horizon (No. of cycles ahead)								
		25	50	100	150	200	250	300	350	400
MAE ($*10^{-2}$)	Average	0.98	1.3	1.4	1.2	1.8	1.5	2.0	2	2.1
	Best	**0.82**	1.1	1.13	0.91	1.4	1.1	1.5	1.5	1.7
STD ($*10^{-2}$)	Average	1.26	1.6	2	2.1	2.4	2.1	2.3	2.5	2.5
	Best	**0.72**	1.2	1	1.3	1.8	1.4	1.4	1.9	2.1
RMSE ($*10^{-2}$)	Average	1.6	2.1	2.4	2.4	3.1	2.6	3	3.2	3.3
	Best	**1.1**	1.7	1.5	1.6	2.3	1.77	2.1	2.4	2.7
RMSPE (%)	Average	1.81	2.3	2.8	2.86	3.6	3	3.49	3.7	3.8
	Best	**1.18**	1.8	1.67	1.94	2.71	2.06	2.5	2.8	3.15
NMSE ($*10^{-1}$)	Average	1.3	2	3.1	2.9	4.5	3.1	3.9	4.5	4.4
	Best	**0.52**	1.3	1.05	1.2	2.21	1.3	1.7	2.3	3

Each vector representing a cycle gathers features that have different ranges and units. In order to improve the efficiency of the learning process and to give the same importance to all features, a scaling step is necessary. For this work, a feature-wise maximum-absolute normalisation is applied. Each feature is scaled by its maximum absolute value and therefore is shifted to a $[-1,1]$ range or $[0,1]$ if all the values are positive.

3.4. Error metrics

In order to compare the performances of our models between them and with other approaches in the literature, several scoring measures were used. In a vast majority of works, the evaluation of models is based on the Root Mean Square Error (RMSE), Root Mean Square Percentage Error (RMSPE) and Mean Absolute Error (MAE). MAE and RMSE are very common error metrics in regression problems. Individual errors are not taken into account in the same way in RMSE and MAE. Large errors and outliers have a greater influence in RMSE due to the squaring information, whereas MAE is a linear score, which means that all the individual errors are weighted equally in the average. RMSPE has the advantage of being scale independent, which is convenient to compare predicting performances across different datasets [30].

We also add the Normalised Mean Square Error (NMSE) in order to compare the performances of our models with future work, and the Standard Deviation of the MAE (σ_{MAE}) in order to evaluate the reliability of the models. These quality measures are expressed as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_{pred,i} - y_i| \qquad (6)$$

$$\sigma_{MAE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (a_i - MAE)^2} \qquad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_{pred,i} - y_i)^2} \qquad (8)$$

$$RMSPE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{(y_{pred,i} - y_i)}{y_i} \right)^2} \qquad (9)$$

$$NMSE = \frac{\sum_{i=1}^{N} (y_{pred,i} - y_i)^2}{N * V} \qquad (10)$$

In all these formulas, $y_{pred,i}$ is the predicted SOH, y_i is the real SOH and N is the number of samples on which the error is calculated. V is the variance of y. For example, the use of the mean of y as the predicted values would give an NMSE of 1.
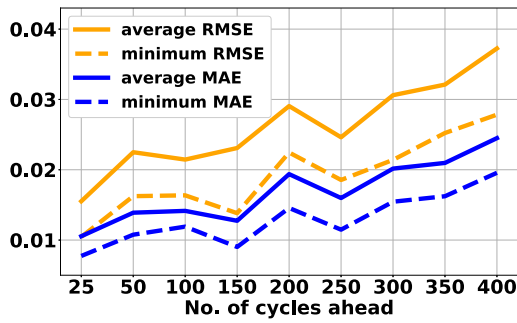
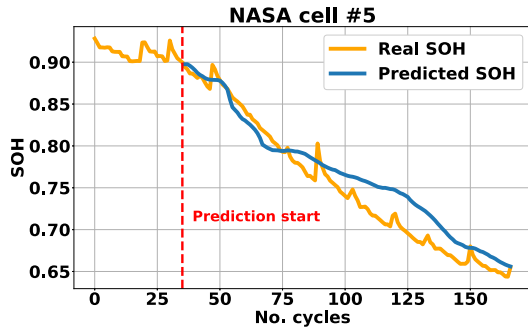**Fig. 7.** Evolution of the average and minimum MAE and RMSE according to the number of cycles ahead.

Fig. 8. Real VS 12-cycle-ahead predicted SOH for NASA battery # 5.

Table 3
XLSTM performances for SOH prediction on NASA batteries.

Metric	Type	B0005	B0006	B0007	B0018
MAE ($*10^{-2}$)	Average	1.6	2.2	2.9	4.5
	Best	1.2	1.9	1.5	1.9
STD ($*10^{-2}$)	Average	1.1	1.6	1.3	2.5
	Best	0.7	1.3	0.8	3.4
RMSE ($*10^{-2}$)	Average	2	2.7	3.2	5.2
	Best	1.4	2.3	1.6	2.05
RMSPE (%)	Average	2.6	4	4.1	7
	Best	1.9	3.4	2.05	5.2
NMSE ($*10^{-1}$)	Average	0.67	1	3	9
	Best	0.32	0.7	0.7	5.2

Table 4
Comparison of SOH prediction performances, 50 cycles ahead, on MIT batteries.

	MAE $*10^{-2}$	STD $*10^{-2}$	RMSE $*10^{-2}$	NMSE $*10^{-1}$
XLSTM	1.07	1.2	1.6	0.124
biXLSTM	1.1	1.2	1.7	0.13
AE-LSTM [18]	2.43	1.22	2.78	8.07

3.5. Training process

The SOH-window-XLSTM is composed of four layers, including two LSTM layers, one Dropout layer and one Dense layer, organised as follows : [LSTM, Dropout, LSTM, Dense]. The Dropout layer aims at preventing the network from overfitting [31]. The two LSTM layers are respectively composed of 64 and 256 neurons. The Dense layer contains only one neuron and no activation function as the network is trained to make a regression on one value of SOH. The dropout rate is of 0.1. The training process is completed through back propagation, with the Adam optimiser and a learning rate of 0.001.

The number of training epochs is determined by observing training and validation loss curves. In order to make sure the network is not overfitting, training is stopped if the validation loss has not decreased

during the last 500 epochs, and the weights that correspond to the best validation loss are restored.

3.5.1. Training models with the MIT dataset

For each training process, the whole dataset is divided into three parts: a training, a validation and a test set. The training set contains 64% of all available data, the validation set contains another 20% and the test set is composed of the remaining 16%. The three ensembles are distinct and every single battery is used only once. The sharing of battery between train, validation and test sets is made randomly.

Each measurement showed in the following section is computed as follows: for each model, 10 different random train/validation/test splits are made, and the model is trained from scratch 10 times. For every training process, performance measurements are done on the test set as the mean of all errors on all predictions. After completing the 10 consecutive training processes, the final error measurement is computed as the mean of all previously computed measures.

3.5.2. Training models with the NASA dataset

There are a lot less exploitable batteries in the NASA dataset than in the MIT dataset. Indeed, only 6 exploitable batteries were kept from the 34 initial ones, compared to 124 exploitable batteries in the MIT dataset. Therefore, it is much more complicated to build a performing training dataset, with large training, validation and test splits.

For each training process, only one battery was kept for validation, and the rest of the batteries were used to train the model. The validation set is also used as the test set, therefore, the prediction errors that are represented in the next section correspond to the validation battery. In order to have a reliable evaluation of the error, 10 successive trainings are done with the same train/validation split. For each train/validation split, we show the mean of the performance scores of the 10 successive trainings, and also the best performance.

The 6 used batteries from the NASA PCoE dataset are batteries B0005, B0006, B0007, B0018, B0054 and B0055. Batteries B0054 and B0055 are always part of the training set, which means no error measures will be shown for them in Section 3. When the predicting performances are shown for battery B0005 for example, it means the model has been trained on 5 batteries: B0054, B005, and B0006, B0007, B0018. The process is the same for batteries B0006, B0006, and B0018.

4. Results and discussion

4.1. Predicting performances

This section investigates the predicting performances of the SOH-window-XLSTM according to the dataset used. Observations on the quality of the models are made as a function of the predicting horizon. The size of the input window stays the same for all tests and was fixed at 25 consecutive cycles. Due to the maximum-absolute normalisation, MAE and RMSE measurements refer to signals that have an amplitude of 1, as explained in Section 3.3. Different models are trained according to the input data. As there are important differences between batteries that were tested in the MIT and in the NASA dataset, distinct models need to be trained. Therefore, the results regarding MIT batteries and NASA batteries will be described in two different sections. For both cases, all error metrics will be shown.

4.1.1. SOH prediction on the MIT dataset

Table 1 shows the predicting performances of our SOH-window-XLSTM model on MIT batteries. SOH predictions can be made from 25 up to 400 cycles ahead. For all cases, two metrics are given: the average prediction error and the minimum prediction error. As explained in Section 3, different successive trainings are made on different train/validation/test splits, and the average error is computed as the mean of all trainings. The prediction error of the best performing model is also shown for each predicting horizon.

Table 5
Computed and selected features.

Domain	Feature	Current		Voltage		Temperature	
		Charge	Discharge	Charge	Discharge	Charge	Discharge
Temporal	Total energy	√	√				
	Area under the curve			√	√		
Statistical	Mean				√		
	Root mean square						
	Max					√	√
	Min						
Spectral	Fundamental freq.						
	F$_{max}$					√	
	Power bandwidth				√		√

As expected, the best performances are obtained for short term predictions, 25 cycles ahead. As the predicting horizon grows longer, the predicting performances degrade. The best of our model reaches a 1.14% RMSPE for a 25-cycle-ahead prediction. Nevertheless, with very long term predictions, up to 400 cycles ahead, the RMSPE stays as low as 3.14%. In Fig. 7, the average and minimum MAE and RMSE of each model, corresponding to each predicting horizon, are plotted. The curves show that the prediction error increases almost linearly according to the number of cycles ahead. The structure of the model does not vary from one predicting horizon to the other, which means that the architecture is quite robust.

In Figs. 5 and 6, a comparison is made between the real SOH curve and the predicted one, for 25 and 300-cycle-ahead predictions with the best models. The batteries chosen to compare real SOH and predicted SOH were picked in the test set of the corresponding model. In both cases, the predicted curves are very close to the real SOH curve.

Given the results presented in Table 1, our models stay between $1.05.10^{-2}$ (25 cycles ahead) and $2.4.10^{-2}$ (400 cycles ahead) for the MAE, and the standard deviation of the MAE stays below 2.10^{-2} (for 400 cycles ahead). Considering that those measures refer to signals that have a [0,1] range, we can conclude that our SOH-window-XLSTM models give accurate predictions, at both short term and long term, with a high reliability.

4.1.2. SOH prediction on the NASA PCoE dataset

Table 1 shows the predicting performances of our SOH-window-XLSTM model, on NASA batteries. SOH predictions are made only 12 cycles ahead. Contrary to MIT batteries, the ones that were tested by the NASA PCoE have a short cycle life (168 cycles maximum, compared to more than 2000 cycles for some batteries in the MIT dataset). In order to have a reasonable number of training samples, the choice was made to limit the predicting horizon to 12 cycles.

Even though there is not much available training data, the predicting results are still satisfactory and vary in the same range as the ones obtained with MIT batteries. The three models corresponding to batteries B0005, B0006 and B0007 have comparable performances. For these three batteries, the best predicting models have a RMSPE of 1.9%, 3.4% and 2.05% respectively. However, the predicting performances of the model on battery B0018 is slightly lower, with a RMSPE of 5.2% . This might be due to the fact that battery B0018 has a lower cycle life than batteries B0005 to B0007. Indeed, batteries B0005, B0006 and B0007 are tested during 168 cycles, batteries B0054 and B0055 are tested during 90 cycles, and battery B0018 is the only one to be tested for 132 cycles.

Fig. 8 shows a comparison between the real SOH and the predicted SOH for battery B0005. The two curves are very close. Although local regeneration phenomena are not taken into account, our model catches the global trend of SOH degradation. With more training data, local variations could be predicted with more accuracy.

Although the predicting performances are quite good on those four batteries, the fact that the worst performance is obtained with the most "isolated" battery in the dataset shows that a low number of training data impacts negatively the performances of our predictive model. In all cases, the SOH-window-XLSTM trained on MIT data is more performing. There are 20 times more exploitable batteries in the MIT dataset than in the NASA dataset, and their lifetime is 5 times longer on average. Therefore, the MIT SOH-window-XLSTM model is trained with a lot more data, leading to better and more reliable results.

4.2. Comparison with other approaches

4.2.1. MIT prediction results

There have not been many approaches in the literature dealing with the SOH prediction from the MIT dataset. We compare our predicting performances with those of Audin et al. [18] that deal with the prediction of SOH, multi cycles ahead.

In [18], the prediction performances of their AE-LSTM model is given for a 25-cycle input window, and the output is the value of SOH, 50 cycles ahead. We compare their predicting performances to our best performing SOH-window-XLSTM mode, for a 50-cycle-ahead prediction.

We can see from Table 4 that our SOH-window-XLSTM model outperforms the AE-LSTM model. For a 50-cycle-ahead SOH prediction, our models shows a RMSE of $1.6 * 10^{-2}$ compared to $2.78 * 10^{-2}$ for the AE-LSTM model.

Our model is less complex as it only uses LSTM in the model itself. The data preprocessing requires low computational abilities as very simple features are extracted from time series.

For comparison, another type of RNN was trained with the same structure as our SOH-window-XLSTM, except that bi-directional LSTM (biLSTM) are used in stead of unidirectional LSTM. BiLSTM imply the same data transformations as detailed in Section 2.3, but two LSTM are trained on the input sequence : the first one on the input sequence as-is and the second one on a reversed copy of the input sequence. The comparative architecture involving biLSTM is referred to as SOH-window-biXLSTM.

The results obtained with the SOH-window-biXLSTM are summarised in Table 2. By comparing Tables 1 and 2, several observations can be made. The performances of the SOH-window-XLSTM and SOH-window-biXLSTM are very close. Regarding the MAE of the best models, the SOH-window-XLSTM has slightly better performances for short term predicting horizons (25 and 50 cycles ahead), and slightly lower performances for long term predicting horizons (350 and 400 cycles ahead). On average however, the MAE SOH-window-XLSTM has a lower standard deviation.

4.2.2. NASA prediction results

We compare our predictive performances with two articles published in 2019 by Qu et al. [7] and 2021 by Liu et al. [8]. Although those two approaches describe very well their architecture and training process, it is hard to compare our model with theirs. Considering that the exogenous architecture described in this work is based on the use of TSF only and not past SOH values, it is hard to compare our SOH-window-XLSTM model to the PA-LSTM [7] or the LSTM + GPR + IMF [28], that are both NAR models. Moreover, our approach aims at building a generic model from a maximum of available data. With

MIT batteries, it is easier to test the model on several unseen batteries as there are so many. With the NASA dataset however, we split the batteries in a way that the validation and test ensemble are merged together, and contain only one battery.

The major difference in the approach described in this article is that all available data from all batteries are used to train the model. Training data can be considered as "offline data", used to learn degradation patterns. Once the model is trained, predictions can be made "online" from any input window of the test battery. This means that it is not necessary to isolate the first part of the data from a given battery and use it in model learning to predict future SOH values. In [7], at least 30% of the data is used for incremental learning, and it can go up to 70% of the SOH curve. That means only the last 50 cycles can be predicted by the model. In [8], 50% of the SOH data is also used for training. Moreover, in both [7,8], predictions are made iteratively. We use input windows of 25 cycles, which represents less than 20% of the lifetime of battery B0018 and less than 15% of the lifetime of batteries B0005, B0006, and B0007.

As shown in Table 3, our model has a RMSE of $1.4 * 10^{-2}$, $2.3 * 10^{-2}$ and $3.9 * 10^{-2}$ for batteries B0005, B0006, and B0007 respectively. Our method is outperformed by [7,8], as the first one reaches an RMSE of $1.9 * 10^{-2}$, $1.97 * 10^{-2}$ and $2.08 * 10^{-2}$ for batteries B0005, B0006, and B0007, and the second one reaches an RMSE of $3.6 * 10^{-3}$ and $4.9 * 10^{-3}$ for batteries B0005 and B0006 (B0007 is not indicated).

However, those results should be analysed regarding the context. Our model is less complex and takes operating data such as current, voltage and temperature as input. Our approach is motivated by the observations made in Section 1. The very good results obtained by NAR models depend on the nature of the input data. Experimental ageing tests of Li-Ion batteries lead to a steady degradation trend, but one that does not represent the real use of an electric vehicle. By using TSF in a model, the SOH prediction is slightly less accurate, but more flexible and adaptable to different use cases.

5. Conclusion

This article focuses on the prediction of future SOH values of Li-Ion batteries, based on the study of two major datasets published by the MIT and the NASA PCoE. Our study of the existing state of the art shows that the use of NAR models could lead to very accurate prediction results, but that this approach was not the most adapted for the PHM of Li-Ion batteries.

We propose here a feature extraction strategy coupled with feature selection in order to use time series of current, voltage and temperature as input to a SOH predicting model. These features, coupled with an LSTM neural network can predict with accuracy and on the long term future values of SOH from an input window of 25 cycles. Although we do not reach the same accuracy as NAR models, our SOH-window-XLSTM model has promising performances and could adapt to more realistic situations. It is still complicated to provide quantitative arguments to support this hypothesis, but the principle of NAR models for SOH prediction implicitly require that the use conditions of a battery remain the same throughout its entire life. By observing operating data, our model could take into account any change in the use environment of the vehicle (through external temperature), or any change is the driving profile of a user (through current and voltage).

The next step in our work will be to use data that would be more representative of the real use of a vehicle to train our SOH-window-XLSTM and test its reliability. A custom ageing dataset is being developed, with the use of driving cycles to discharge the batteries.

Moreover, the SOH-window-XLSTM model could be completed in two different ways. The studied datasets only make use of isolated battery cells. The behaviour of such cells in a battery pack is different due to their very high number and the interactions between them. In order to adapt the model described in this article to a full battery pack used in an electric vehicle, it is necessary to model the interactions between cells that occur in the pack. In this way, a SOH prediction for a full battery pack could be considered. Moreover, the principal goal of PHM of Li-Ion batteries is to identify dangerous behaviours from the user. Neural Networks are black-box models which results are accurate but hardly interpretable. In order to provide a feedback to the user on how to optimise the performances of the battery, a layer of explainability could be added to the SOH-window-XLSTM.

Declaration of competing interest

Data availability

Data will be made available on request.

Acknowledgements

Appendix

Table 5 shows which features are kept as input to the SOH-window-XLSTM after the feature selection process.

References

[1] MarketsandMarket, Lithium-Ion Battery Market with COVID-19 Impact Analysis, by Type, Capacity, Voltage, Industry (Consumer Electronics, Automotive, Power, Industrial), & Region (North America, Europe, APAC & RoW) - Global Forecast to 2030, Technical report, 2021.

[2] Hideyuki Noguchi Masaki Yoshio, Lithium-ion batteries, 2009, pp. 9–48, Chapter 2.

[3] Mauro Cordella, Felice Alfieri, Christian Clemm, Anton Berwald, Durability of smartphones: A technical analysis of reliability and repairability aspects, J. Clean. Prod. 286 (2021) 125388.

[4] Akash Basia, Zineb Simeu-Abazi, Eric Gascard, Peggy Zwolinski, Review on state of health estimation methodologies for lithium-ion batteries in the context of circular economy, CIRP J. Manuf. Sci. Technol. 32 (2021) 517–528.

[5] Huixing Meng, Yan Fu Li, A review on prognostics and health management (PHM) methods of lithium-ion batteries, Renew. Sustain. Energy Rev. 116 (September) (2019) 109405.

[6] Gae Won You, Sangdo Park, Dukjin Oh, Diagnosis of electric vehicle batteries using recurrent neural networks, IEEE Trans. Ind. Electron. 64 (6) (2017) 4885–4893.

[7] Jiantao Qu, Feng Liu, Yuxiang Ma, Jiaming Fan, A neural-network-based method for RUL prediction and SOH monitoring of lithium-ion battery, IEEE Access 7 (2019) 87178–87191.

[8] Kailong Liu, Yunlong Shang, Quan Ouyang, Widanalage Dhammika Widanage, A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery, IEEE Trans. Ind. Electron. 68 (4) (2021) 3170–3180.

[9] Tedjani Mesbahi, Nassim Rizoug, Patrick Bartholomeus, Redha Sadoun, Fouad Khenfri, Philippe Le Moigne, Dynamic model of li-ion batteries incorporating electrothermal and ageing aspects for electric vehicle applications, IEEE Trans. Ind. Electron. 65 (2) (2018) 1298–1305.

[10] Xiaosong Hu, Le Xu, Xianke Lin, Michael Pecht, Battery lifetime prognostics, Joule 4 (2) (2020) 310–346.

[11] Xiaoyu Li, Lei Zhang, Zhenpo Wang, Peng Dong, Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and elman neural networks, J. Energy Storage 21 (December 2018) (2019) 510–518.

[12] Weihan Li, Neil Sengupta, Philipp Dechent, David Howey, Anuradha Annaswamy, Dirk Uwe Sauer, One-shot battery degradation trajectory prediction with deep learning, J. Power Sources 506 (2021) 230024.

[13] Qiuhui Ma, Ying Zheng, Weidong Yang, Yong Zhang, Hong Zhang, Remaining useful life prediction of lithium battery based on capacity regeneration point detection, Energy 234 (2021) 121233.

[14] Jiusi Zhang, Student Member, Yuchen Jiang, Remaining useful life prediction of lithium-ion battery with adaptive noise estimation and capacity regeneration detection, 2022.

[15] Xiaoqiong Pang, Rui Huang, Jie Wen, Yuanhao Shi, Jianfang Jia, Jianchao Zeng, A lithium-ion battery rul prediction method considering the capacity regeneration phenomenon, Energies 12 (11) (2019).

[16] Phattara Khumprom, Nita Yodo, A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm, Energies 12 (4) (2019).

[17] Ren Lei, A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life, IEEE Trans. Ind. Inform. 17 (5) (2021) 3478–3487.

[18] Paul Audin, Ines Jorge, Tedjani Mesbahi, Ahmed Samet, Francois De Bertrand De Beuvron, Romuald Bone, Auto-encoder LSTM for Li-ion SOH prediction: A comparative study on various benchmark datasets, in: Proceedings - 20th IEEE International Conference on Machine Learning and Applications, ICMLA 2021, IEEE, 2021, pp. 1529–1536.

[19] Qianglong Li, Dezhi Li, Kun Zhao, Licheng Wang, Kai Wang, State of health estimation of lithium-ion battery based on improved ant lion optimization and support vector regression, J. Energy Storage 50 (July 2021) (2022) 104215.

[20] Ines Jorge, Ahmed Samet, Tedjani Mesbahi, Romuald Bone, New ANN results on a major benchmark for the prediction of RUL of lithium ion batteries in electric vehicles, in: Proceedings - 19th IEEE International Conference on Machine Learning and Applications, ICMLA 2020, 2020, pp. 1246–1253.

[21] Yunwei Zhang, Qiaochu Tang, Yao Zhang, Jiabin Wang, Ulrich Stimming, Alpha A. Lee, Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning, 2020.

[22] Dezhi Li, Licheng Wang, Chongxiong Duan, Qiang Li, Kai Wang, Temperature prediction of lithium-ion batteries based on electrochemical impedance spectrum: A review, Int. J. Energy Res. 46 (8) (2022) 10372–10388.

[23] Marília Barandas, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, Hugo Gamboa, TSFEL: Time series feature extraction library, SoftwareX 11 (2020) 100456.

[24] Isabelle Guyon, André Elisseeff, An introduction of variable and feature selection, J. Mach. Learn. Res. Spec. Issue Var. Feature Sel. 3 (2003) 1157–1182.

[25] Ron Kohavi, George H.John, Wrappers for feature subset selection, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), in: LNCS, vol. 7920, 2013, pp. 654–678, (97).

[26] Yoshua Bengio, Patrice Simard, Paolo Frasconi, Learning long-term dependencies with gradient descent is difficult, IEEE Trans. Neural Netw. 5 (2) (1994) 157–166.

[27] Romuald Boné, Michel Crucianu, Jean Pierre Asselin de Beauville, Learning long-term dependencies by the selective addition of time-delayed connections to recurrent neural networks, Neurocomputing 48 (1–4) (2002) 251–266.

[28] Datong Liu, Jingyue Pang, Jianbao Zhou, Yu Peng, Data-driven prognostics for lithium-ion battery based on Gaussian process regression, in: Proceedings of IEEE 2012 Prognostics and System Health Management Conference, PHM-2012, IEEE, 2012, pp. 1–5.

[29] Kristen A. Severson, Peter M. Attia, Data-driven prediction of battery cycle life before capacity degradation, Nature Energy 4 (5) (2019) 383–391.

[30] Rob J. Hyndman, Anne B. Koehler, Another look at measures of forecast accuracy, Int. J. Forecast. 22 (4) (2006) 679–688.

[31] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958.