State of Health and Lifetime Prediction of Lithium-ion Batteries Using Self-learning Incremental Models

Murilo Camargos[1,2,3] and Plamen Angelov[1,2,3]

[1] *School of Computing and Communications, Lancaster University, Lancaster, LA1 4WA, UK*
m.camargos@lancaster.ac.uk
p.angelov@lancaster.ac.uk

[2] *Lancaster Intelligent, Robotic and Autonomous Systems (LIRA) Research Centre, Lancaster, UK*

[3] *The Faraday Institution, Quad One, Becquerel Avenue, Harwell Campus, Didcot, OX11 0RA, UK*

ABSTRACT

Lithium-ion batteries are key energy storage elements in the context of environmental-aware energy systems representing a crucial technology to achieve the goal of zero carbon emission. Therefore, its conditions must be monitored to guarantee the safe and reliable operation of the systems that use these components. Furthermore, lithium-ion batteries' prognostics and health management policies must cope with the nonlinear and time-varying nature of the complex electrochemical dynamics of battery degradation. This paper proposes an incremental-learning-based algorithm to estimate the State of Health (SoH) and the Remaining Useful Life (RUL) of lithium-ion batteries based on measurement data streams. For this purpose, a two-layer framework is proposed based on incremental modeling of the SoH. In the first layer, a set of representative features are extracted from voltage and current data of partial charging and discharging cycles; these features are then used to train the proposed model in a recursive procedure to estimate the battery's SoH. The second layer uses the capacity data for incremental learning of an Autoregressive (AR) model for the SoH, which will be used to propagate the battery's degradation through time to make the RUL prediction. The proposed method was applied to two datasets for experimental evaluation, one from CALCE and another from NASA. The proposed framework was able to estimate the SoH of 8 different lithium-ion cells with an average percentage error below 1.5% for all scenarios, while the lifetime model predicted the cell's RUL with a maximum average error of 25%.

Murilo Camargos et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ACRONYMS

SoH	State of Health
RUL	Remaining Useful Life
TS	Takagi-Sugeno
MF	Membership Function
RLS	Recursive Least Squares
MAPE	Mean Absolute Percentage Error
RMSPE	Root Mean Squared Percentage Error
FT	Fault Threshold
AR	Autoregressive
IC	Incremental Capacity
DV	Differential Voltage

1. INTRODUCTION

The lithium-ion batteries are key energy storage elements in the context of environmental-aware energy systems due to their recognized performance and energy density. For this reason they have been widely applied in microgrids, consumer electronics, and electric vehicles. However, the safety and operation costs of those systems become dependent on the battery's storage capacity and lifetime. Indeed, the battery's charge capacity is progressively reduced due to the aging and repeated charging (discharging) cycles (Birkl, Roberts, McTurk, Bruce, & Howey, 2017). Therefore, the PHM of lithium-ion batteries is essential to improve the reliability of those systems and extend their lifetime. In this sense, the PHM methodologies for lithium-ion batteries (Y. Zhang & Li, 2022; Omariba, Zhang, & Sun, 2018; Ge, Liu, Jiang, & Liu, 2021) are responsible for estimating their State of Health (SoH) and Remaining Useful Life (RUL).

In particular, the SoH of a battery denotes the ratio of its current parameters (e.g., charge capacity, impedance, and power)

and the same parameters at the beginning of its life (Cai, Lin, & Liao, 2022). Otherwise, the RUL of a battery is defined as time between the current observation instant and the battery collapse instant, denominated the battery's end-of-life (Dong, Han, & Wang, 2021). For a battery approach, it is necessary to estimate the SoH and RUL, since these variables cannot be directly measured in an everyday battery use. Some techniques as Coulomb counting and peak tracking in Incremental Capacity (IC)/Differential Voltage (DV) curves often require long measurement duration, are prone to noise amplification and might not be well generalized for different cells (Richardson, Birkl, Osborne, & Howey, 2019).

The physics of the battery's degradation can be used to derive mathematical models, e.g., battery equivalent circuit models or electrochemical models, to enable the SoH estimation and RUL prediction. However, those model-based methodologies (Downey, Lui, Hu, Laflamme, & Hu, 2019; Lui et al., 2021) required high fidelity models for the degradation process that is usually nonlinear and time-varying for batteries. Since some of those models' parameters are not known *a priori*, the application of model-based methods might also require parameter estimation techniques which becomes more challenging for complex mathematical models.

Hence, the use of data-driven methodologies (Wu, Fu, & Guan, 2016) is specially attractive in applications to SoH and RUL estimation of lithium-ion batteries. In general, the data-driven methodologies for lithium-ion batteries can be classified into three groups:

1. **Empirical methods** are built based on some historical data used to estimate parameters for a chosen structure, e.g., exponential models (Cai et al., 2022), autoregressive models (M. Camargos et al., 2021), and neural networks (Q. Zhang et al., 2022).
2. **Stochastic methods** describe degradation phenomenons as stochastic processes and the SoH as a random variable. The parameters of those stochastic models can be estimated through different methodologies, such as Gaussian process regression (Richardson et al., 2019; X. Li, Wang, & Yan, 2019) and Bayesian filtering (Si, 2015).
3. **Signal-based methods** aims at obtaining the relation between the capacity loss and the measured signals properties which can be extracted in the time or frequency domain (Khaleghi, Firouz, Mierlo, & den Bossche, 2019; Wang, Pan, Liu, Cheng, & Zhao, 2016).

Most of the aforementioned data-driven methodologies establish stiff relations between the process data and the degradation processes related to the SoH and RUL estimation. However, the battery's degradation and its relation with the measurable data is complex, time-varying, and tends to be particular for each cell under test. This context motivates the development of methodologies which are able to adapt themselves to the cell's behavior without compromising the generality ability. In this sense, some adaptive or incremental learning methodologies have recently been proposed for battery's SoH and RUL estimation (J. Zhang et al., 2022; Qin, Zhao, & Liu, 2022; Si, 2015). Although those methods are able to update their parameters to provide a better representation of the degradation processes based on the data streams, they are unable to modify their complexity to capture novel dynamic behaviors.

In this regard, evolving systems are effective tools for obtaining incremental models which update their structure and adapt their parameters through autonomous learning from data streams (Angelov, 2012). For this reason, evolving systems have been effectively applied for dealing with complex and time-varying dynamics aiding to solve different problems, such as fault diagnosis (Shah & Wang, 2021), classification (Soares, Angelov, & Gu, 2020), time-series prediction and forecasting (Severiano, de Lima e Silva, Cohen, & Guimarães, 2021), system identification (Škrjanc, 2021), and learning-based control (Cordovil, Coutinho, Bessa, Peixoto, & Palhares, 2022). Recently, the use of evolving fuzzy models has been proposed for solving the RUL prediction problems (M. O. Camargos, Bessa, D'Angelo, Cosme, & Palhares, 2020), including with applications to lithium-ion batteries (Ahwiadi & Wang, 2022; M. Camargos et al., 2021). In addition to flexibility and adaptability of those incremental models, the evolving fuzzy models provides interpretability for the data-driven approaches. For example, the prognostics based on evolving fuzzy models allows to relate the increment of the model structure to the degradation stage.

This paper addresses the problem of data-stream based SoH estimation and RUL prediction for lithium-ion batteries. To solve this problem, it is proposed a two-layer framework based on incremental modeling of the SoH. The first layer extracts features related to the voltage from charging cycles, then, uses the extracted features for incremental learning of the SoH behavior and SoH estimation. The second layer uses the capacity data for incremental learning of an Autoregressive (AR) model for the SoH, which is then applied for RUL prediction.

The remainder of this paper is organized as follows: Section 2 provides an overview on the class of self-learning incremental models used in this paper; Section 3 describes the proposed methodology for SoH estimation and RUL prediction; Section 4 presents the experimental procedures and setup; Section 5 presents the results for SoH estimation and RUL prediction applied to two lithium-ion batteries' datasets; and Section 6 draws the conclusions and indicates further research directions.

2. Self-learning Incremental Models

The self-learning incremental models used in this paper are represented using the Takagi-Sugeno (TS) representation. This

type of representation can be seen as a mixture of linear models whose mixing probabilities are given by fuzzy relations. Moreover, it is a powerful modeling technique that is capable of approximating nonlinear dynamics, multiple operating modes and significant parameter and structure variations (Angelov & Filev, 2004).

The TS fuzzy models are rule-based models composed by C IF-THEN rules that are used as an inference system. The antecedents are represented by fuzzy relations between the input data and a knowledge base of fuzzy sets while the consequents are usually linear functions (Nguyen et al., 2019). Each rule is represented as:

$$\begin{aligned}\text{Rule } i\text{: } &\textbf{IF } (x_1 \textbf{ is } \Phi_{i1}) \textbf{ AND } \ldots \textbf{ AND } (x_{n_x} \textbf{ is } \Phi_{in_x}) \\ &\textbf{THEN } \hat{y}_i = \mathbf{a}_i^\top \tilde{\mathbf{x}}\end{aligned} \tag{1}$$

where $\mathbf{x} = [x_1, \ldots, x_{n_x}]^\top \in \mathbb{R}^{n_x}$ is the vector of premise variables, $\mathbf{a}_i \in \mathbb{R}^{n_x+1}$ is the vector of estimated consequent parameters, and $\tilde{\mathbf{x}} = \begin{bmatrix}1 & \mathbf{x}^\top\end{bmatrix}^\top$. Moreover, $(x_j \textbf{ is } \Phi_{ij})$ denotes the fuzzy relation between x_j and the fuzzy set Φ_{ij} for $i \in \mathbb{N}_{\leq C}$ and $j \in \mathbb{N}_{\leq n_x}$. Throughout the text, $\mathbb{N}_{\leq k}$ will be used to denote the set of natural numbers up to k, such that $\mathbb{N}_{\leq k} = \{1, 2, \ldots, k\}$.

The fuzzy relation in the antecedents will define the activation degree of each rule, i.e., the level of contribution of each local linear model to the overall output. The activation degree is given by a Membership Function (MF) $\varphi_{ij} \colon \mathbb{R} \to [0, 1]$ that maps a given input's component x_j to the unit partition, for $j \in \mathbb{N}_{\leq n_x}$. For Gaussian-like antecedent fuzzy sets, the MF is of the form:

$$\varphi_{ij}(x_j) = \exp\left(-\alpha \|x_j - x_{ij}^\star\|^2\right), \tag{2}$$

where $x_{ij}^\star$ is the j-th component of the focal point of the i-th rule, $\alpha = 4/r^2$ and r defines the radius of the neighborhood of a data point, also known as the model's zone of influence. According to (Angelov & Filev, 2004), too large a value of r leads to averaging while too small values leads to over-fitting. In general, values of $r \in [0.3, 0.5]$ can be recommended (Angelov & Filev, 2004; Chiu, 1994). The final activation degree of the i-th rule is defined as the Cartesian product or conjunction of respective fuzzy sets:

$$w_i(\mathbf{x}) = \bigcap_{j=1}^{n_x} (x_j \textbf{ is } \Phi_{ij}) = \prod_{j=1}^{n_x} \varphi_{ij}(x_j). \tag{3}$$

The output of the TS fuzzy model is a convex combination among C consequent linear models weighted by the rules' activation degrees. The activation degrees must comply with the convex sum property, i.e., they need to be non-negative and sum one. From the center average defuzzification, the overall model output is given as

$$\hat{y} = \sum_{i=1}^{C} h_i(\mathbf{x})\, \mathbf{a}_i^\top \tilde{\mathbf{x}} \tag{4}$$

in which

$$h_i(\mathbf{x}) = \frac{w_i(\mathbf{x})}{\sum_{m=1}^{C} w_m(\mathbf{x})}. \tag{5}$$

Given a set of input and output data, the problem of identifying the TS model, i.e., finding the number of rules, the focal points in Eq. (2) and the parameters of the linear subsystems in Eq. (4), is divided into two parts:

1. Finding the antecendents' focal points: $\{\mathbf{x}_1^\star, \ldots, \mathbf{x}_C^\star\}$
2. Finding the parameters of each linear subsystem: $\{\mathbf{a}_1 \ldots, \mathbf{a}_C\}$

This first task can be solved by clustering the input-output data space while the second task can be solved by computing each linear model's parameters in the least-squares sense. In (Angelov & Filev, 2004), online learning strategies for these tasks are given. In such cases, the number of clusters changes as new data samples becomes available, therefore, C becomes C_k and both the antecedents and consequents parameters also change in time, becoming $\{\mathbf{x}_{k1}^\star, \ldots, \mathbf{x}_{kC_k}^\star\}$ and $\{\mathbf{a}_{k1} \ldots, \mathbf{a}_{kC_k}\}$.

In the clustering problem, a recursive variation of the so-called subtractive clustering (Chiu, 1994) algorithm is given. The proposed algorithm uses a Cauchy type function of first order to represent the potential of each data point to become a focal point. This function enables recursive calculation and is both monotonic and inversely proportional to the distance between two data points. The computed potential is then used to decide whether the new data point will be used to replace an old focal point or will represent a new focal point, or cluster center. The consequent parameters are updated using the Recursive Least Squares (RLS) algorithm.

The rule-base model will dynamically upgrade the number of clusters in the input-output data space or modify existing ones, while preserving rules that represents old knowledge. The details of this procedure can be found in (Angelov & Filev, 2004).

3. State of Health and Lifetime Prediction

In order to estimate the SoH and to predict the lifetime, i.e., estimate its RUL, of the batteries, we propose the a parallel architecture using two self-learning incremental models as described in Section 2. As shown in Figure 1, one model will be used to estimate the SoH while another one will be used to predict the RUL. They use the same learning procedure, as

described by (Angelov & Filev, 2004); however, their inputs are different: the SoH predictor uses extracted features from partial charge procedures while the RUL predictor uses past values of the charge capacity time-series to predict the next SoH, i.e., it is an AR model.

The parameter set for each model is given as:

$$\boldsymbol{\theta}_k^p = \left\{ C_k^p,\, \mathbf{x}_{k1}^{p\star}, \ldots, \mathbf{x}_{kC_k^p}^{p\star},\, \mathbf{a}_{k1}^p, \ldots, \mathbf{a}_{kC_k^p}^p \right\} \tag{6}$$

where the superscript $p \in [1, 2]$ indicates the SoH model and the lifetime model respectively. The estimation on both models is done as:

$$\hat{y}_k^p = f_p(\mathbf{x}_k^p | \boldsymbol{\theta}_{k-1}^p, \boldsymbol{\Omega}_p) \tag{7}$$

where $f_p \colon \mathbb{R}^{n_x^p} \to \mathbb{R}$ is a TS fuzzy model in the form of Eq. (4), n_x^p is the dimension of input $\mathbf{x}_k^p$, and $\boldsymbol{\Omega}_p$ is a set of time-invariant parameters. The learning procedure to update the parameter set is given as

$$\boldsymbol{\theta}_{k|k-1}^p = h\left(\boldsymbol{\theta}_{k-1}^p, \mathbf{x}_k^p, y_k, \hat{y}_k^p, \boldsymbol{\Omega}_p\right). \tag{8}$$

Both models use the same learning procedure h. The time-invariant parameter set of the SoH model contains only the model's zone of influence $\boldsymbol{\Omega}_1 = \{r_1\}$ while the lifetime model also contains the number of past capacity values to predict the next one $\boldsymbol{\Omega}_2 = \{r_2, L\}$.

Their inputs are different from each other; the lifetime model takes a lagged vector as input, i.e.,

$$\mathbf{x}_k^2 = [y_{k-1}, \ldots, y_{k-L}]^\top \in \mathbb{R}^{n_x^2}, \tag{9}$$

where $n_x^2 = L$ is the number of past values the lifetime model will take as inputs to predict the next one. When the prognostics task starts, i.e., when $k = t_P$, the SoH values estimated by the lifetime model will replace true SoH values in the input vector shown in Eq. (9). Then, at cycles $k \geq t_P + L$, all components in Eq. (9) will be previous estimates.

The SoH model uses features extracted from partial charge data as described in (Richardson et al., 2019). To overcome the necessity of having to identify the parameters of highly accurate battery models or the requirement of having long measurements that ensures the coverage of IC/DV curve's peaks and dealing with the noise that comes out of this process, the proposed feature extraction uses direct voltage data from partial charging procedures. After defining a specific voltage window $[V_{\text{lb}}, V_{\text{ub}}]$, M equispaced voltages are taken and the time it takes to go from one voltage to another is defined as the feature for the SoH estimation. The extracted features for the SoH estimator are given as:

$$\mathbf{x}_k^1 = [\tau_k(v_0, v_1), \ldots, \tau_k(v_M - 1, v_M)]^\top \in \mathbb{R}^{n_x^1}, \tag{10}$$

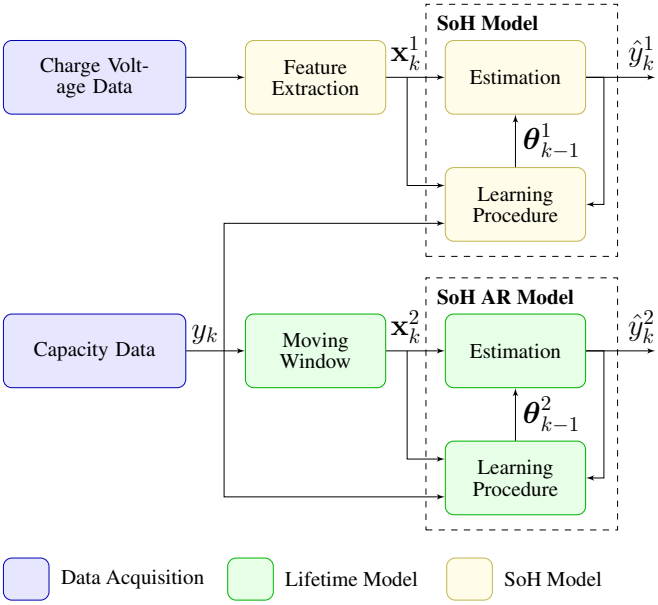

Figure 1. Parallel architecture of self incremental models for State of Health and lifetime prediction.

where $n_x^1 = M$ and

$$v_i = V_{\text{lb}} + i \cdot \frac{V_{\text{ub}} - V_{\text{lb}}}{M - 1} \tag{11}$$

in which $\tau_k(v_a, v_b)$ computes the time it takes to go from voltage v_a to voltage v_b in the k-th charging cycle.

4. Experimental Setup

In order to test the proposed architecture shown in Figure 1, datasets that represent the degradation of lithium-ion batteries from two sources were used. This type of battery is commonly found in industry and commercially, e.g., in electric vehicles, microgrids, and electronic devices (X. Li et al., 2019; Saha & Goebel, 2009).

In the first dataset, the cycle aging experiments of four lithium-ion batteries (B0005, B0006, B0007, B0018) are provided by NASA Ames Prognostics Center of Excellence (PCoE)[1] (Saha & Goebel, 2007). The testbed comprises commercial lithium-ion 18650-sized rechargeable batteries from the Idaho National Laboratory; a programmable 4-channel DC electronic load and power supply; voltmeters, ammeters, and a thermocouple sensor suite; custom electrochemical impedance spectrometry equipment; and environmental chamber to impose different operational conditions. The batteries run at room temperature (23° C). Charging is done in constant mode at 1.5 A, until the voltage reaches 4.2 V. Discharging is performed at a constant current level of 2 A, until the battery voltage reaches 2.7 V (Saha & Goebel, 2009).

[1] http://ti.arc.nasa.gov/project/prognostic-data-repository

The second dataset also contains four cycle aging experiments (CS2 35, CS2 36, CS2 37, CS2 38) of prismatic cells with graphite anode and a lithium cobalt oxide cathode. The data is provided by the Center for Advanced Life Cycle Engineering (CALCE)[2] from the University of Maryland. The cycling of the batteries was accomplished by multiple full charge-discharge tests using an Arbin BT2000 battery testing system under room temperature. The batteries were cycled at constant current of 1 C (1.1 A) with charging and discharging being cut off at the manufacturer's specified cutoff voltage (from 2.7 V to 4.2 V). The capacity of the tested batteries was estimated using the Coulomb counting method (He, Williard, Osterman, & Pecht, 2011; Xing, Ma, Tsui, & Pecht, 2013).

As the cells ages, its maximum available capacity will decrease. In this paper, the SoH is defined as the relative capacity of each cell, i.e., the computed capacity at cycle k divided by the cell's nominal capacity (2 A for NASA cells and 1.1 A for CALCE cells).

The results are obtained after defining the hyperparameters for both parallel models, i.e., $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$. For the SoH model, we use the standard value of $r_1 = 0.3$. For the lifetime model we perform a cross validation task to find the best values for (r_2, L). We choose a training cell and a validation cell to perform a grid search over the parameters. The optimal parameters are given as:

$$\begin{aligned} \underset{\rho,\ell}{\arg\min} \quad & \frac{1}{H - t_P} \sum_{k=t_P}^{H-1} \frac{|r_k(\rho,\ell) - \hat{r}_k(\rho,\ell)|}{r_k(\rho,\ell)} \\ \text{subject to} \quad & \rho \in [0.3, 0.35, 0.4, 0.45, 0.5], \\ & \ell \in [1, 2, 3, 4, 5, 6, 7, 8] \end{aligned} \tag{12}$$

where $\hat{r}_k(\rho,\ell)$ is the estimated RUL at the k-th cycle of the validation cell using a model trained with the training cell, $r_k(\rho,\ell)$ is the true RUL under the same conditions, t_P is the cycle in which prognostics task starts and H is the validation cell lifespan.

Here, we define the RUL as the time elapsed between the prognostics task starting time (t_P) and the time in which the system's degradation state reaches a given Fault Threshold (FT) (N. Li, Lei, Lin, & Ding, 2015). Formally, we can express the RUL as:

$$\hat{r}_{t_P} = \inf\left\{ n \in \mathbb{N} \,\middle|\, \hat{y}^2_{t_P+n} \geq \eta \right\}, \tag{13}$$

where $\hat{r}_k$ denotes the RUL computed at instant t_P, given that the true values of the state of health are known up until t_P, $\mathbb{N}$ is the natural numbers set, and η is the predefined FT. Moreover, Eq. (13) is a simplified version of the canonical RUL definition by (Chiachío, Chiachío, Sankararaman, Saxena, & Goebel, 2015), where the concept of failure domain is defined.

[2] https://web.calce.umd.edu/batteries/data.htm

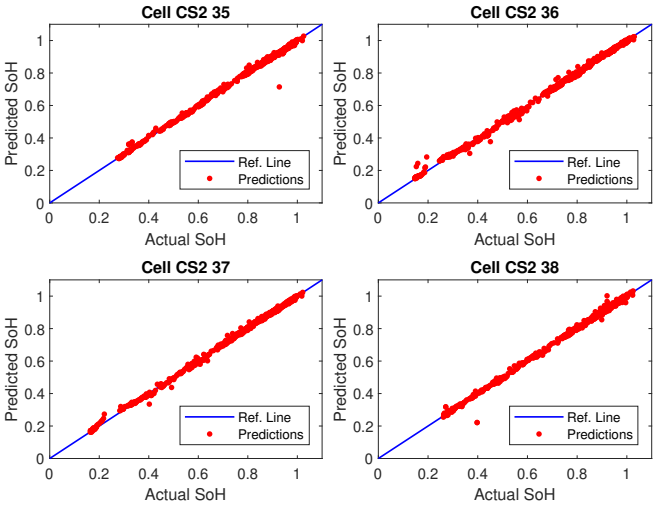

Figure 2. Predicted and actual SoH for CALCE cells.

5. Results

The results section is divided into SoH estimation and lifetime prediction experiments.

5.1. SoH estimation

The feature extraction was done in a similar way of (Richardson et al., 2019). We chose the lower and upper bound voltages as roughly 75% to 100% of the charging voltage span to represent a more realistic use case scenario as full-cycle charging are not always available. Specifically, we set $V_{\text{lb}} = 3.85$ and $V_{\text{ub}} = 4.2$ with $M = 4$ equispaced voltages. Therefore, the SoH model inputs shown in Eq. (10) is defined as:

$$\mathbf{x}^1_k = \begin{bmatrix} \tau_k(3.8500, 3.9375) \\ \tau_k(3.9375, 4.0250) \\ \tau_k(4.0250, 4.1125) \\ \tau_k(4.1125, 4.2000) \end{bmatrix}. \tag{14}$$

In this phase, no previous training was done and the SoH model learned online, as new data became available, how to predict the SoH from the inputs in Eq. (14). We set the model's fine tuning parameter to $r_1 = 0.3$. The predicted versus the actual SoH for both CALCE and NASA datasets are shown in Figure 2 and Figure 3. In the CALCE dataset the prediction far away from the reference line are outliers in the charge/discharge measurements from the Arbin testing system.

The SoH model predictions are also evaluated according to two metrics, namely Mean Absolute Percentage Error (MAPE) and Root Mean Squared Percentage Error (RMSPE), defined

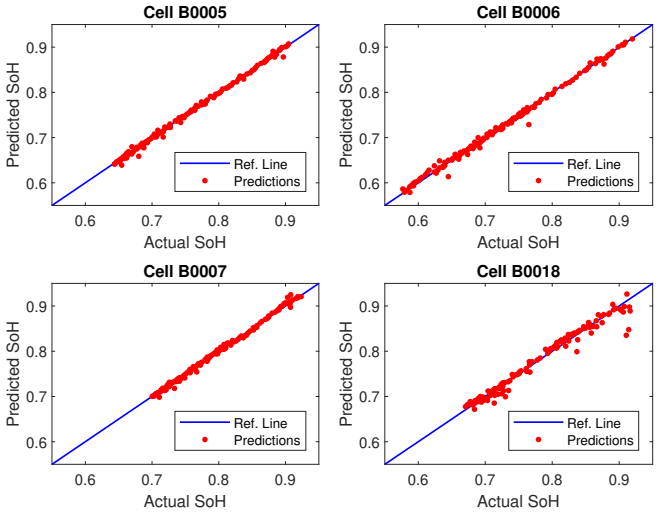

Figure 3. Predicted and actual SoH for NASA cells.

Table 1. MAPE and RMSPE values for the SoH model's predictions.

Cell ID	MAPE	RMSPE
CS2 35	0.7698%	1.5965%
CS2 36	1.2358%	3.4065%
CS2 37	0.9417%	1.8461%
CS2 38	0.8633%	2.3728%
B0005	0.3915%	0.6053%
B0006	0.5626%	0.9134%
B0007	0.3793%	0.5355%
B0018	0.9777%	1.5378%
Avg.	0.7652%	1.6017%

as:

$$\text{MAPE}(\mathbf{y}, \hat{\mathbf{y}}^1) = \frac{100}{N_T} \sum_{k=1}^{N_T} \left| \frac{y_k - \hat{y}_k^1}{y_k} \right| \tag{15}$$

$$\text{RMSPE}(\mathbf{y}, \hat{\mathbf{y}}^1) = \sqrt{\frac{1}{N_T} \sum_{k=1}^{N_T} \left(\frac{y_k - \hat{y}_k^1}{y_k} \right)^2} \tag{16}$$

where, $\mathbf{y}$ is the vector of true SoH values, $\hat{\mathbf{y}}^1$ is the vector of prediction made by the SoH model in Figure 1 and N_T is of samples. The results in Table 1 shows that the proposed model achieved a MAPE value of less than 1.5% for all tested cells with an average of 0.77%. These results indicate the competitiveness of the SoH model in comparison to more complex models reported in the literature.

5.2. Lifetime prediction

In the lifetime prediction task, we first run a cross validation procedure to find reasonable values of the parameters $\mathbf{\Omega}_2$. In order to solve the optimization problem in Eq. 12 we need to choose training and validation cells for both NASA and CALCE datasets. In this paper, this choice is done arbitrarily. For the NASA cells, we chose cells B0006 and B0018 as training and validation cells, respectively; for the CALCE cells, we chose cells CS2 35 and CS2 36 as training and validation cells, respectively.

Table 2. Successive lifetime prediction over multiple starting points for CALCE dataset.

k	CS2 37			CS2 38		
	r_k	$\hat{r}_k$	APE_k	r_k	$\hat{r}_k$	APE_k
100	542	394	27.31%	598	500	16.39%
150	492	418	15.04%	548	375	31.57%
200	442	407	7.92%	498	376	24.50%
250	392	350	10.71%	448	340	24.11%
300	342	294	14.04%	398	298	25.13%
350	292	230	21.23%	348	250	28.16%
400	242	197	18.60%	298	225	24.50%
450	192	200	4.17%	248	220	11.29%
500	142	170	19.72%	198	209	5.56%
550	92	84	8.70%	148	129	12.84%
600	42	32	23.81%	98	84	14.29%
Avg.			15.57%			19.85%

Moreover, we define the fault thresholds of NASA and CALCE cells as $\eta_{\text{NASA}} = 80\%$ and $\eta_{\text{CALCE}} = 70\%$, respectively, due to larger lifespan of CALCE cells. The cross validation procedure yielded the following parameters: $\mathbf{\Omega}_2^{\text{CALCE}} = \{r_2, L\} = \{0.3, 7\}$ and $\mathbf{\Omega}_2^{\text{NASA}} = \{r_2, L\} = \{0.45, 5\}$ for CALCE and NASA, respectively.

The lifetime prediction results for different prognostics starting time (t_P) are shown in Table 2 and Table 3. For each tested cell, there is the true RUL column (r_k), the value estimated by the lifetime model ($\hat{r}_k$) and the absolute percentage error (APE) [3]. Overall, the lifetime prediction errors for all cells did not exceed 25% and the results near the actual end of life of the tested cells do not always decrease, as is expected.

Another way to depict these results is through the $\alpha - \lambda$ plot, which is used to evaluate prognostics strategies since it shows whether the predicted RUL falls within a goal region around the true RUL given by $\pm(\alpha)(100)\%$ (Lall, Lowe, & Goebel, 2012). The $\alpha - \lambda$ plots for all tested cells are shown in Figure 4 using a goal region of $\alpha = 20\%$ to define the accuracy cone. Although the prediction error exceeds the threshold of 20%, it falls below the accuracy cone in almost all the times. This happens when the algorithm underestimates the actual Remaining Useful Life, which is a situation more tolerable than when it overestimates, due to safety reasons (Nectoux et al., 2012).

[3] This is a version of Eq. (15) without the averaging.

Table 3. Successive lifetime prediction over multiple starting points for NASA dataset.

k	B0005			B0007		
	r_k	$\hat{r}_k$	APE_k	r_k	$\hat{r}_k$	APE_k
20	19	21	10.53%	30	24	20.00%
22	17	19	11.76%	28	22	21.43%
24	15	17	13.33%	26	20	23.08%
26	13	14	7.69%	24	18	25.00%
28	11	12	9.09%	22	16	27.27%
30	9	10	11.11%	20	15	25.00%
32	7	7	0.00%	18	13	27.78%
34	5	5	0.00%	16	12	25.00%
36	3	3	0.00%	14	10	28.57%
38	1	1	0.00%	12	9	25.00%
Avg.			6.35%			24.81%

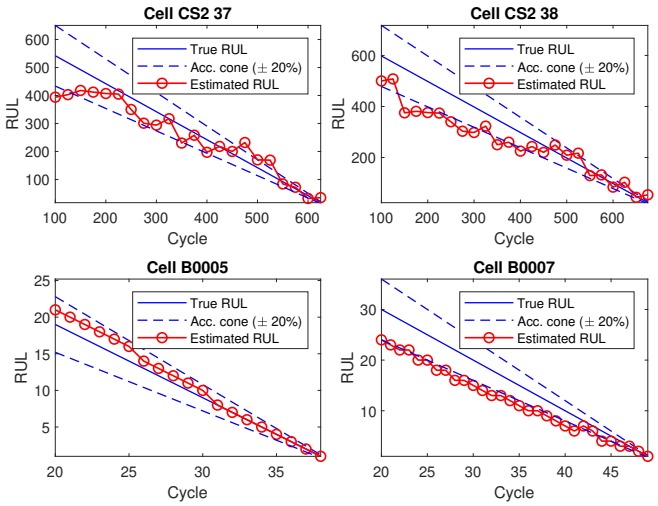

Figure 4. $\alpha - \lambda$ plot of the estimated RUL of all testing cells with goal region of $\alpha = 20\%$.

6. Conclusions

The proposed parallel architecture, composed of two self-learning incremental models, e.g., evolving TS, was capable of estimating both the SoH and the lifetime of different cells from two datasets. Moreover, the use of partial measurements of the charging voltage in the SoH model approximates us from more realistic use case scenarios, such as in electric vehicles where the charging cycles are rarely complete. Furthermore, the non-stationary behaviour seen in charge/discharge cycles can be accommodated through the operational ability to quickly change the model's parameters and structure as new data becomes available.

The technique reached an average percentage error below 1.5% in all tested cells, indicating its competitiveness concerning other models reported in the literature that are more complex and whose training phase happens offline. Self-learning incremental models are promising methods to deal with non-linear problems in non-stationary environments. Their structures are flexible, and their parameters can be updated recursively according to data stream changes.

Furthermore, the lifetime model managed to estimate the RUL for different cells with a low computational cost. Structural learning from scratch, quick recursive updates, and historical-data storage avoidance make self-learning incremental models quite suitable to be used in real-time prognostics systems. However, the results near the actual end of life of the tested cells do not always decrease, as is expected, indicating the proposed methodology can be improved to provide better long-term predictions. The proposed model have offered online condition monitoring and a way of fusing multivariate data streams describing the multiple-stage battery-degradation phenomenon.

Acknowledgment

This work was carried out with funding from the Faraday Institution (faraday.ac.uk; EP/S003053/1), grant number FIRG025.

References

Ahwiadi, M., & Wang, W. (2022). An adaptive evolving fuzzy technique for prognosis of dynamic systems. *IEEE Transactions on Fuzzy Systems*, *30*(3), 841-849. doi: 10.1109/TFUZZ.2021.3049916

Angelov, P. (2012). *Autonomous learning systems: from data streams to knowledge in real-time*. John Wiley & Sons.

Angelov, P., & Filev, D. (2004, February). An approach to online identification of takagi-sugeno fuzzy models. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *34*(1), 484–498. doi: 10.1109/tsmcb.2003.817053

Birkl, C. R., Roberts, M. R., McTurk, E., Bruce, P. G., & Howey, D. A. (2017, February). Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, *341*, 373–386. doi: 10.1016/j.jpowsour.2016.12.011

Cai, L., Lin, J., & Liao, X. (2022, July). A data-driven method for state of health prediction of lithium-ion batteries in a unified framework. *Journal of Energy Storage*, *51*, 104371. doi: 10.1016/j.est.2022.104371

Camargos, M., Bessa, I., Junior, L. A. Q. C., Coutinho, P., Leite, D. F., & Palhares, R. M. (2021). Evolving fuzzy system applied to battery charge capacity prediction for fault prognostics. In *Atlantis studies in uncertainty modelling*. Atlantis Press. doi: 10.2991/asum.k.210827.010

Camargos, M. O., Bessa, I., D'Angelo, M. F. S. V., Cosme, L. B., & Palhares, R. M. (2020, November). Data-driven prognostics of rolling element bearings using a novel error based evolving takagi–sugeno fuzzy model. *Applied Soft Computing*, *96*, 106628. doi: 10.1016/j.asoc.2020.106628

Chiachío, J., Chiachío, M., Sankararaman, S., Saxena, A., & Goebel, K. (2015). Condition-based prediction of time-dependent reliability in composites. *Reliab. Eng. Syst. Saf.*, *142*, 134–147. doi: 10.1016/j.ress.2015.04.018

Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, *2*(3), 267–278. doi: 10.3233/ifs-1994-2306

Cordovil, L. A. Q., Coutinho, P. H. S., Bessa, I., Peixoto, M. L. C., & Palhares, R. M. (2022, January). Learning event-triggered control based on evolving data-driven fuzzy granular models. *International Journal of Robust and Nonlinear Control*, *32*(5), 2805–2827.

Dong, G., Han, W., & Wang, Y. (2021, November). Dynamic bayesian network-based lithium-ion battery health prognosis for electric vehicles. *IEEE Transactions on Industrial Electronics*, *68*(11), 10949–10958. doi: 10.1109/tie.2020.3034855

Downey, A., Lui, Y.-H., Hu, C., Laflamme, S., & Hu, S. (2019, February). Physics-based prognostics of lithium-ion battery using non-linear least squares with dynamic bounds. *Reliability Engineering & System Safety*, *182*, 1–12. doi: 10.1016/j.ress.2018.09.018

Ge, M.-F., Liu, Y., Jiang, X., & Liu, J. (2021, April). A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries. *Measurement*, *174*, 109057. doi: 10.1016/j.measurement.2021.109057

He, W., Williard, N., Osterman, M., & Pecht, M. (2011, December). Prognostics of lithium-ion batteries based on dempster–shafer theory and the bayesian monte carlo method. *Journal of Power Sources*, *196*(23), 10314–10321. doi: 10.1016/j.jpowsour.2011.08.040

Khaleghi, S., Firouz, Y., Mierlo, J. V., & den Bossche, P. V. (2019, December). Developing a real-time data-driven battery health diagnosis method, using time and frequency domain condition indicators. *Applied Energy*, *255*, 113813. doi: 10.1016/j.apenergy.2019.113813

Lall, P., Lowe, R., & Goebel, K. (2012, November). Prognostics health management of electronic systems under mechanical shock and vibration using kalman filter models and metrics. *IEEE Transactions on Industrial Electronics*, *59*(11), 4301–4314. doi: 10.1109/tie.2012.2183834

Li, N., Lei, Y., Lin, J., & Ding, S. X. (2015). An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Ind. Electron.*, *62*(12), 7762–7773. doi: 10.1109/TIE.2015.2455055

Li, X., Wang, Z., & Yan, J. (2019). Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression. *J. Power Sources*, *421*(February), 56–67. doi: 10.1016/j.jpowsour.2019.03.008

Lui, Y. H., Li, M., Downey, A., Shen, S., Nemani, V. P., Ye, H., ... Hu, C. (2021, February). Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction. *Journal of Power Sources*, *485*, 229327. doi: 10.1016/j.jpowsour.2020.229327

Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., & Varnier, C. (2012, June). PRONOSTIA : An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12.* (Vol. sur CD ROM, p. 1-8). Denver, Colorado, United States: IEEE Catalog Number : CPF12PHM-CDR.

Nguyen, A. T., Taniguchi, T., Eciolaza, L., Campos, V., Palhares, R., & Sugeno, M. (2019). Fuzzy control systems: Past, present and future. *IEEE Comput. Intell. Mag*, *14*(1), 56–68. doi: 10.1109/MCI.2018.2881644

Omariba, Z., Zhang, L., & Sun, D. (2018, May). Review on health management system for lithium-ion batteries of electric vehicles. *Electronics*, *7*(5), 72. doi: 10.3390/electronics7050072

Qin, P., Zhao, L., & Liu, Z. (2022, March). State of health prediction for lithium-ion battery using a gradient boosting-based data-driven method. *Journal of Energy Storage*, *47*, 103644. doi: 10.1016/j.est.2021.103644

Richardson, R. R., Birkl, C. R., Osborne, M. A., & Howey, D. A. (2019, January). Gaussian process regression for *In Situ* capacity estimation of lithium-ion batteries. *IEEE Transactions on Industrial Informatics*, *15*(1), 127–138. doi: 10.1109/tii.2018.2794997

Saha, B., & Goebel, K. (2007). Battery data set. *NASA Ames Prognostics Data Repository*.

Saha, B., & Goebel, K. (2009). Modeling li-ion battery capacity depletion in a particle filtering framework. In *Proceedings of the annual conference of the prognostics and health management society* (pp. 2909–2924).

Severiano, C. A., de Lima e Silva, P. C., Cohen, M. W., & Guimarães, F. G. (2021, June). Evolving fuzzy time series for spatio-temporal forecasting in renewable energy systems. *Renewable Energy*, *171*, 764–783. doi: 10.1016/j.renene.2021.02.117

Shah, J., & Wang, W. (2021, May). An evolving neuro-fuzzy classifier for fault diagnosis of gear systems. *ISA Transactions*. doi: 10.1016/j.isatra.2021.05.019

Si, X.-S. (2015). An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data. *IEEE Transactions on Industrial Electronics*, *62*(8), 5082-5096. doi: 10.1109/TIE.2015.2393840

Škrjanc, I. (2021, December). An evolving concept in the identification of an interval fuzzy model of wiener-hammerstein nonlinear dynamic systems. *Information Sciences*, *581*, 73–87. doi: 10.1016/j.ins.2021.09.004

Soares, E., Angelov, P., & Gu, X. (2020, September). Autonomous learning multiple-model zero-order classifier

for heart sound classification. *Applied Soft Computing*, *94*, 106449. doi: 10.1016/j.asoc.2020.106449

Wang, L., Pan, C., Liu, L., Cheng, Y., & Zhao, X. (2016, April). On-board state of health estimation of $LiFePO_4$ battery pack through differential voltage analysis. *Applied Energy*, *168*, 465–472. doi: 10.1016/j.apenergy.2016.01.125

Wu, L., Fu, X., & Guan, Y. (2016, May). Review of the remaining useful life prognostics of vehicle lithium-ion batteries using data-driven methodologies. *Applied Sciences*, *6*(6), 166. doi: 10.3390/app6060166

Xing, Y., Ma, E. W., Tsui, K.-L., & Pecht, M. (2013, June). An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *Microelectronics Reliability*, *53*(6), 811–820. doi: 10.1016/j.microrel.2012.12.003

Zhang, J., Jiang, Y., Li, X., Huo, M., Luo, H., & Yin, S. (2022, June). An adaptive remaining useful life prediction approach for single battery with unlabeled small sample data and parameter uncertainty. *Reliability Engineering & System Safety*, *222*, 108357. doi: 10.1016/j.ress.2022.108357

Zhang, Q., Yang, L., Guo, W., Qiang, J., Peng, C., Li, Q., & Deng, Z. (2022, February). A deep learning method for lithium-ion battery remaining useful life prediction based on sparse segment data via cloud computing system. *Energy*, *241*, 122716. doi: 10.1016/j.energy.2021.122716

Zhang, Y., & Li, Y.-F. (2022, June). Prognostics and health management of lithium-ion battery using deep learning methods: A review. *Renewable and Sustainable Energy Reviews*, *161*, 112282. doi: 10.1016/j.rser.2022.112282