POLITECNICO DI TORINO
Repository ISTITUZIONALE

Battery state of health estimation with improved generalization using parallel layer extreme learning machine

(Article begins on next page)

18 October 2022

Article

Battery State of Health Estimation with Improved Generalization Using Parallel Layer Extreme Learning Machine

Ethelbert Ezemobi * , **Andrea Tonoli** and **Mario Silvagni**

Department of Mechanical and Aerospace Engineering, Politecnico di Torino, 10129 Torino, Italy;
andrea.tonoli@polito.it (A.T.); mario.silvagni@polito.it (M.S.)
* Correspondence: ethelbert.ezemobi@polito.it; Tel.: +39-388-095-8460

Abstract: The online estimation of battery state of health (SOH) is crucial to ensure the reliability of the energy supply in electric and hybrid vehicles. An approach for enhancing the generalization of SOH estimation using a parallel layer extreme learning machine (PL-ELM) algorithm is analyzed in this paper. The deterministic and stable PL-ELM model is designed to overcome the drift problem that is associated with some conventional machine learning algorithms; hence, extending the application of a single SOH estimation model over a large set of batteries of the same type. The PL-ELM model was trained with selected features that characterize the SOH. These features are acquired as the discrete variation of indicator variables including voltage, state of charge (SOC), and energy releasable by the battery. The model training was performed with an experimental battery dataset collected at room temperature under a constant current load condition at discharge phases. Model validation was performed with a dataset of other batteries of the same type that were aged under a constant load condition. An optimum performance with low error variance was obtained from the model result. The root mean square error (RMSE) of the validated model varies from 0.064% to 0.473%, and the mean absolute error (MAE) error from 0.034% to 0.355% for the battery sets tested. On the basis of performance, the model was compared with a deterministic extreme learning machine (ELM) and an incremental capacity analysis (ICA)-based scheme from the literature. The algorithm was tested on a Texas F28379D microcontroller unit (MCU) board with an average execution speed of 93 µs in real time, and 0.9305% CPU occupation. These results suggest that the model is suitable for online applications.

Keywords: state of health; parallel layer extreme learning machine; artificial intelligence; improved generalization; automotive; hybrid vehicles; electric vehicles; batteries; online application; energy reliability

1. Introduction

The choice of lithium-ion battery in several applications, especially in the automotive industry, is mainly influenced by the high energy density per weight. Moreover, the large number of charge/discharge cycles, and reduced memory effect influence this choice. Aging is a limiting factor in the operational life and the amount of energy that can be delivered by lithium-ion batteries, which has an impact on the reliability.

Battery aging reflects on the capacity loss affected by the loss of active electrode material, which is the loss in the storage medium. Greater loss is linked with the storage content, which is the loss in active lithium [1]. Graphite anode operates at a low potential of about 1.5 V, which is outside of the electrochemical stability window for the electrolyte component. As a result, reductive decomposition of the electrolyte occurs, accompanied by irreversible lithium-ion consumption [1,2]. The result of the decomposition is a film plating known as the solid electrolyte interface (SEI) at the interface between the electrolyte and the electrode. This plating protects the electrolytes against further decomposition and the charged electrode against corrosion [3]. The operational phase of the lithium battery is

associated with the intercalation and deintercalation of lithium into the electrode particle layer. This results in volume change, especially on the graphite anode (about 10%) [1]. The dilation of the graphite particle anode breaks the SEI, exposing a fresh graphite surface on which SEI is rebuilt. Consequently, the capacity of the battery is reduced. SEI formation is predominant at the beginning of cycling, but it is also seen as cycling progresses.

Battery energy reliability is a key concern in hybrid and electric vehicle applications; monitoring battery aging is crucial in such applications. Although different authors used different electrochemical indicators such as potential change and impedance rise for measuring battery aging, the state of health (SOH) indicator has remained the most commonly used. SOH is based on the capacity measurement over the life of a battery to account for the capacity fading and hence aging. Accurate estimation of SOH is essential for battery usage optimization and proper risk management.

Generally, SOH prediction strategies are broadly classified into model-based strategies and data-based strategies. The model-based designs often exist as electrochemical or empirical models [4,5]. The electrochemical model exploits the knowledge of the internal properties of the battery or the general intrinsic behavior to derive a white box that represents the battery [6]. The parameters and states of the model have a one-to-one relationship with the internal properties of the physical battery. On the other hand, the states and the parameters of the empirical model do not necessarily have a one-to-one relationship with the internal properties. The empirical models are associated with low computational complexity and perform satisfactorily, especially within the scope of the experiment applied for the parameter fitting. Equivalent circuit models are popular choices for empirical model, exploiting lumped-parameter circuit elements such as resistors, inductors, and capacitors to represent the battery dynamics [7]. Onboard performance of the empirical models is, however, often affected by cumulative measurement errors, capacity degradation over the usage life, environmental parameter variation, and sensitivity to initial conditions [8,9]. For SOH estimation, [10] proposes a model-based voltage-capacity approach that implements incremental capacity analysis (ICA). The application of capacity fading, internal battery internal resistance [11], physics-based modeling [12], dual adaptive H infinity filter (AHIF) combined with a strong tracking filter (STF) [13], and a recursive least square multi-timescale estimator [14] are some of the model-based approaches reported for SOH estimation.

On the other hand, artificial intelligence models are the prevalent conventional data-driven approaches for estimation problems. A host of methods that apply artificial intelligence for SOH estimation have been proposed in the literature. The authors in [15] estimate SOH with a back-propagation neural network classifier using variations in the voltage, state of charge (SOC), and energy state of the battery as feature variables. A time-series momentum back-propagation neural network algorithm is adopted with battery internal resistance and voltage variation as feature variables for SOH estimation in [16]. Estimation of SOH with a multilayer perceptron algorithm is investigated in [17]. Moreover, the authors in [18] estimate SOH using a group method for data handling (GMDH) polynomial neural network. Although artificial intelligence neural network models have numerous advantages, their solutions often lead to optimization problems, which can result in local minima. Models optimized at local minima are suboptimal.

A host of methods exist in the literature for extracting features that show a correlation with SOH. In [19], they characterize SOH using a exploited importance sampling (IS) strategy. The drawback is that it requires the battery charges or discharges to be sufficient for the feature variables to be extracted. ICA [20] and incremental voltage analysis (IVA) are other useful strategies, but their sensitivity to noise limits their application [4,21]. Real-life applicability is a concern for most of the approaches adopted in the literature. Other feature extraction strategies can be found in [10,18,22]. An alternative approach using a discrete variation of selected indicator variables as the feature variable is adopted in this work [15].

The production process of the lithium-ion battery is not fully mature. The cell-to-cell variation in the battery pack amounts to a variation in the distribution of cell parameter

variables [23]. This variation influences the pattern of aging. A model is required that, when trained with a set of battery dataset, will accurately predict the SOH of other batteries of the same type. However, it is often the case that many trained SOH prediction models experience deviation or drift when applied to unobserved batteries of the same type. In traditional machine learning algorithms, the training and test data are required to be identically distributed [19]. Such drift is often attributed to the wide variation in the distribution of cell parameter variables.

Some authors adopt transfer learning approaches for solving the drift problem [19]. Transfer learning via domain adaptation [24] transfers shared knowledge across the distribution of a known source battery and the distribution of a different but related target battery. The goal is for the source battery to acquire some knowledge from the solution task and use the knowledge to improve the estimation on the target battery. Minimization of the distance between these distributions, however, results in a high computational problem [24]. A procedure for SOH estimation using transfer component analysis (TCA) is demonstrated in [19]. Aside from the high model complexity and the high computational demand associated with domain adaptation approaches, the estimation model requires continuous learning each time a prediction is required. These drawbacks make it unsuitable for online automotive applications.

A less complex model based on ICA is proposed by [25] for enhancing generalization across batteries of the same type. The ICA model uses the concept of regional capacity and regional voltage to estimate SOH within the domain of incremental capacity analysis. A couple of papers report attempts to estimate SOH with extreme learning machine (ELM) [26,27]. The high prediction accuracy and learning speed of ELM makes it an elegant tool for applications where computational resources are a constraint. The ELM model can, however, be affected by the drift problem. The parallel layer extreme learning machine (PL-ELM) [28] proposed in this paper serves as a suitable alternative for SOH estimation with improved accuracy.

An extensive analysis and comparison between the original ELM and PL-ELM is reported in [28]. Some variants of the multiple layer ELM adopted for solving regression and classification problems include parallel layer perceptrons (PLP) ELM [29], and multiparallel extreme learning machine with excitatory and inhibitory neurons (MEI-ELM) [30]. Others are reported in [31,32].

Although PL-ELM is not a new concept, its advantages in terms of its low complexity and computational efficiency, especially for SOH estimation, can be useful for researchers. An improved version of PL-ELM obtained with deterministic parameters offers a mix of stability and good nonlinearity handling. As a contribution, this paper demonstrates the application of a deterministic PL-ELM algorithm as an alternative solution to the drift problem. A SOH estimation model developed with this approach showed an improved generalization across a large set of similar batteries. This algorithm was trained and validated with the NASA prognostic center of excellence (PCoE) dataset to demonstrate this application [33]. The results are compared with the ICA-based algorithm for a constant current profile.

2. Methodology

PL-ELM was developed from the building blocks of ELM. The concept of PL-ELM is that with two parallel layers, two independent projections to the feature space are developed. Each of the projections are nonlinearly activated and they combine through a product. This generates a more powerful nonlinear mapping than just using one activation function, and the prediction capacity is enhanced [28–31]. Since ELM is based on adjusting only the linear parameters using the least squares estimate (LSE), the PL-ELM network provides more freedom for proper adjustment [29]. By assigning the weights and biases deterministically, a stable PL-ELM model is obtained.

2.1. Deterministic Extreme Learning Machine

ELM is a training method for a single hidden layer feedforward artificial neural network [34]. Contrary to the traditional neural network that results in an optimization problem with a slow iterative training process, ELM is a fast noniteration trained model, with no issue of local minima as there is no optimization problem.

According to [35], the weights and biases of a single layer feedforward network (SLFN) can be assigned randomly given that the activation function is infinitely differentiable. ELM can also be used to train SLFN with nondifferentiable activation functions [36]. If the number of neurons on the hidden layer is less than the number of measurements, input weights can be randomly assigned. The output weights, which form the only unknown parameter, can then be determined analytically. ELM computes the estimated model output, y_k by solving Equation (1).

$$y_k = \sum_{j=1}^{m} \Phi_{j,k} \, g(\sum_{i=1}^{n} w_{i,j} x_i + b_j) \tag{1}$$

where $w_{i,j}$ and b_j are the input weight and bias, respectively, and i and n are the respective index and number of neurons on the input layer. The number of neurons corresponds to the number of feature variables, j and m are the respective index and number of neurons on the hidden layer, and k is the index of neurons on the output layer corresponding to the number of outputs. $g(\cdot)$ is a sigmoid activation function and $\Phi_{j,k}$ is the output weight.

Many activation functions including; hyperbolic and cosine activation functions exist in the literature. With an equal number of input neurons, a sigmoid activation function (SAF) returns an enhanced dataset discrepancy or separation when compared with a host of existing activation functions [36]. Also, increasing the number of input neurons (with independent characterizing feature variables) improves the discrepancy of the dataset. For the hidden layer output matrix represented as $H = g(w_{i,j} x_i + b_j)$, Equation (1) can be re-written as Equation (2).

$$H\Phi = \mathbf{y} \tag{2}$$

It is necessary to compute the unknown Φ such that an estimated value $\hat{y}$ approximates the true $\mathbf{y}$. The work of [36] highlights the variants of ELM defined on the basis of the output weight $\hat{\Phi}$, including backpropagation tuning ELM (tELM) and regression-based ELM (rELM). In this work, $\hat{\Phi}$ is developed based on the generalized Moore–Penrose inverse method.

$$\hat{\Phi} = H^+ \mathbf{y} \tag{3}$$

where H^+ is the generalized Moore–Penrose inverse matrix.

In the original ELM, the input weights and biases are assigned randomly. Although the estimated result may occasionally track the reference, the learning process is not deterministic and lacks repeatability. The resulting model can be unstable [36] under certain conditions. This is based on the fact that the randomly assigned input weights and biases have infinite possibilities. In addition, it has been shown that the input weights and the hidden layer biases do not necessarily have to be tuned when assigned and that the output layer matrix can remain unchanged [35]. Therefore, a deterministic ELM is designed by assigning the weights and biases using Equation (4). This increases the likelihood of having different values of H_j while limiting the weight within the distribution range of -1 to 1 [36]. The bias is distributed within the range of 0 to 1.

$$w_{i,j} = \frac{1}{2}[(-1 + \frac{2}{n}i) + (-1 + \frac{2}{m}j)]; \qquad b_j = \frac{j}{m} \tag{4}$$

2.2. Deterministic Parallel Layer Extreme Learning Machine

To develop the PL-ELM model, a second nonlinear layer is introduced in parallel with the hidden layer of the deterministic ELM. The additional parallel layer helps to improve the nonlinearity handling. Adopting PL-ELM is justifiable since it is possible to achieve model improvement in terms of accuracy with only a little increase in model complexity.

The weights of the second component of the PL-ELM are also assigned deterministically but with a low-discrepancy sequence (LDS) [28,37]. Some families of LDS have been developed to efficiently generate a sequence of points deterministically [28]. An LDS tries to maintain the discrepancy of the resulting points within $[0, 1]^n$ as small as possible, and provides a favorable asymptotical rate of convergence of the discrepancy itself. The weights and biases of the first component of PL-ELM are assigned with Equation (4), while the weight of the second component is assigned with Halton's LDS [28] but with zero bias. The structure of PL-ELM is shown in Figure 1.
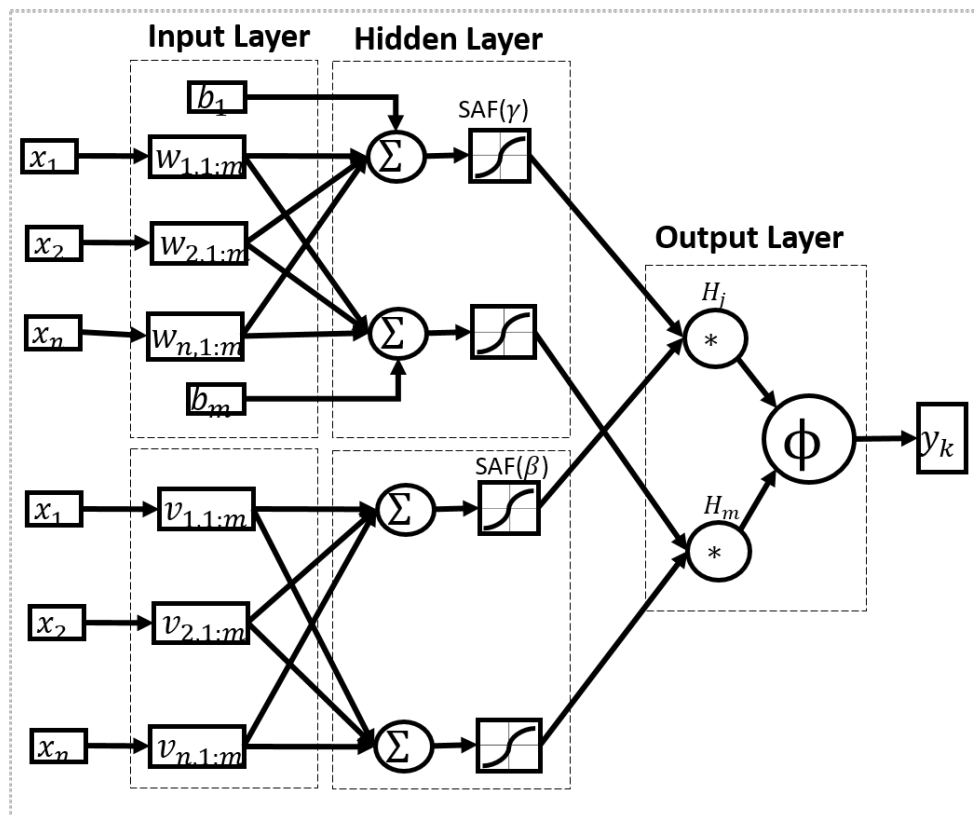

Figure 1. Parallel layer extreme learning machine architecture [28]. The weights and biases of the upper layer are assigned with Equation (4). The lower layer is weighted with Halton's low-discrepancy sequence (LDS) and zero biased.

β and γ are the nonlinear sigmoid activation function (SAF) of the two parallel ELM layers. The output weight, $\hat{\Phi}$ matrix is to be computed. The overall hidden layer output matrix, H (Equation (6)), is the Hadamard product (element-wise product) of the individual parallel hidden layer output matrices, H_1 (Equation (7)) and H_2 (Equation (8)) for N distinct samples (x_k, y_k); $x_k \in \mathbb{R}^n$. Considering SLFN, with m hidden neurons, the output of the network is evaluated as

$$y_k = \sum_{j=1}^{m} \Phi_j \, \beta(v_j x_k) \gamma(w_j x_k + b_j) \qquad for \; 1 \leq k \leq N \tag{5}$$

where v_j and w_j are the weight matrices for the respective parallel input layers and b_j is the bias. Equation (5) can then be written compactly in form of Equation (2), where

$$H = H_1 * H_2 \tag{6}$$

$$H_1 = \begin{pmatrix} \beta(v_1 x_1) & \dots & \beta(v_m x_1) \\ \vdots & \vdots & \vdots \\ \beta(v_1 x_N) & \dots & \beta(v_m x_N) \end{pmatrix} \qquad \in \mathbb{R}^{Nxm} \tag{7}$$

$$H_1 = \begin{pmatrix} \gamma(w_1 x_1 + b_1) & \dots & \gamma(w_m x_1 + b_m) \\ \vdots & \vdots & \vdots \\ \gamma(w_1 x_N + b_1) & \dots & \gamma(w_m x_N + b_m) \end{pmatrix} \quad \in \mathbb{R}^{N \times m} \qquad (8)$$

The objective is to obtain $\hat{\Phi}$ as in Equation (3) such that the error between the estimated and the measured output is minimum (Equation (9)). $\hat{\Phi}$ is obtained by the least square method.

$$\| H\hat{\Phi} - Y \| = \overset{min}{\underset{\Phi}{}} \| H\Phi - Y \| \qquad (9)$$

2.3. Experimental Dataset Description

The analysis was conducted with 2Ah capacity lithium-ion datasets provided by the prognostic center of excellence (PCoE) at NASA's Ames Research Centre for fault prediction and diagnostic studies [33]. The experiment was conducted at room temperature, completely cycling different batteries until their end-of-life (EOL). The term battery and cell are used interchangeably in this paper. Each cycle involved a complete charging and discharge of the battery. Charging was carried out in a constant current (CC) mode of 1.5 A (about 0.75 C-rate) until the battery voltage reached 4.2 V and then continued in a constant voltage (CV) mode until the charge current dropped to 20 mA. Discharge of batteries B0005, B0006, B0007, and B0018 was done at a constant current (CC) load of 2 A (that is a 1 C-rate) until the battery voltage reached the minimum allowable voltage.

2.4. State of Health (SOH)

In lithium-ion batteries, the performance is affected by both storage time and usage—phenomena known as calendar aging and cycle life aging, respectively. The cycle life is the maximum number of cycles or the maximum operating time for a specific set of operating conditions. The computation of SOH is based on capacity fading as in Equation (10). A fully charged battery has a releasable capacity, C_{max} that can be different from the rated capacity, C_{rate}. C_{max} decays with storage time and usage of the battery. These can be used for evaluating battery SOH [38].

$$SOH = \frac{C_{max}}{C_{rated}} 100\% \qquad (10)$$

Considering one of the datasets (B0005), the EOL of the battery was reached after 170 cycles. The SOH that describes the aging of the battery until the EOL is shown in Figure 2 for the charge and discharge phases. The good fit outcome of Figure 2 depicts the similarity of SOH in the charge and discharge phases. This suggests that the SOH estimation is comparable for the both phases.
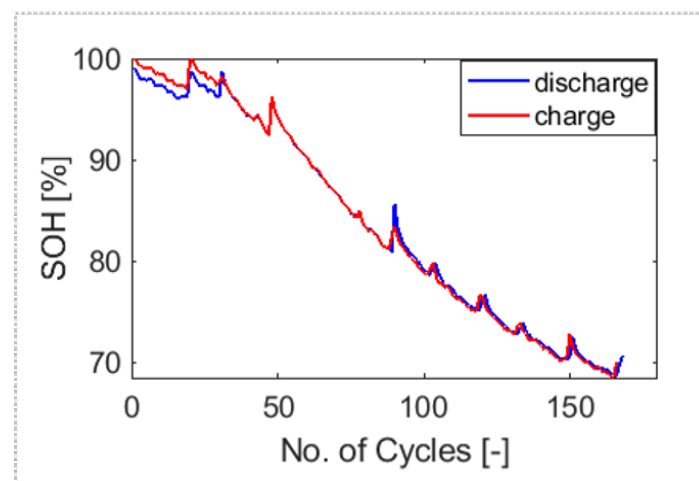

Figure 2. Similarity of state of health (SOH) in the charge and discharge phases. SOH of B0005 is shown to decrease with the number of cycles.

2.5. Characterization and Feature Selection

The training of the PL-ELM requires suitable inputs or feature variables obtained from the characterization of the SOH of the battery. The feature variables used in this model are the discrete variation of internal indicators and operational indicators of SOH. The internal indicators are characteristics that are inherent in the battery such as the SOC, the energy, and the capacity of a battery. The internal indicators are often measured offline. For online implementation, internal indicators can be reasonably estimated using suitable algorithms. Operational indicators such as terminal voltage, current, and temperature are readily available from the battery management system (BMS) and easily accessible at run-time.

The indicators used in this work consist of the combination of the internal and the operational indicators, and they include voltage [V], SOC [%], and energy [Wh] releasable by the battery, as shown in Table 1. The variation of the discrete points of the indicator variables were obtained at 90 s interval and were used as input to the SOH estimation model. This variation in the variables (ΔV, ΔSOC, ΔE) helps to capture the dynamics that exist in the indicator variables while training the SOH estimation algorithm. The choice of the discrete point variation interval was based on model tuning for accuracy.

Table 1. Model feature and output variables.

Indicator Variable	Model Train Features	Unit
Training Feature		
Voltage (V)	$\triangle$V	[V]
State of Charge (SOC)	$\triangle$SOC	[%]
Energy (E)	$\triangle$E	[Wh]
Model Output		
State of Health (SOH)		[%]

The experimental dataset consists of the measured battery terminal voltage and the corresponding current delivered based on request. The terminal voltage of the lithium-ion battery acquired directly from the experiment can be shown to be modeled according to the Thevenin equivalent model of Equation (11). The corresponding equivalent circuit model is shown in Figure 3.

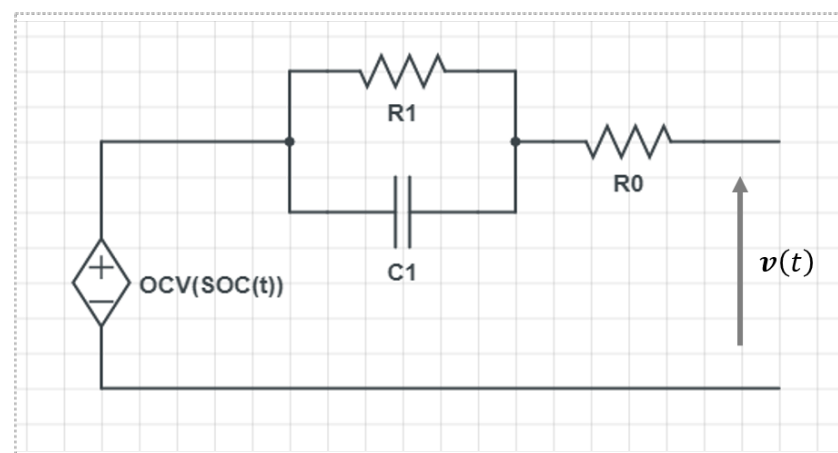$$v(t) = OCV(SOC(t)) - v_C(t) - i(t)R_0 \tag{11}$$

Figure 3. Equivalent circuit that describes the dynamic voltage of a lithium-ion battery [39].

This kind of resistor-capacitor (RC) circuit highlights the presence of a resistive component—the last term on the right-hand side of Equation (11). The presence of this component demonstrates that the energy state of the battery is not conserved. Hence, with

storage time and usage of a lithium-ion battery, capacity fading or aging occurs. The open circuit voltage *OCV* is the equilibrium voltage or the rest voltage of the battery, and it is intrinsically a function of SOC. v_C models the dynamic voltage, which is the voltage across C1 in the RC-circuit. To demonstrate the capacity fading/aging, a two dimensional voltage–capacity plot of the experimental dataset is shown in Figure 4.
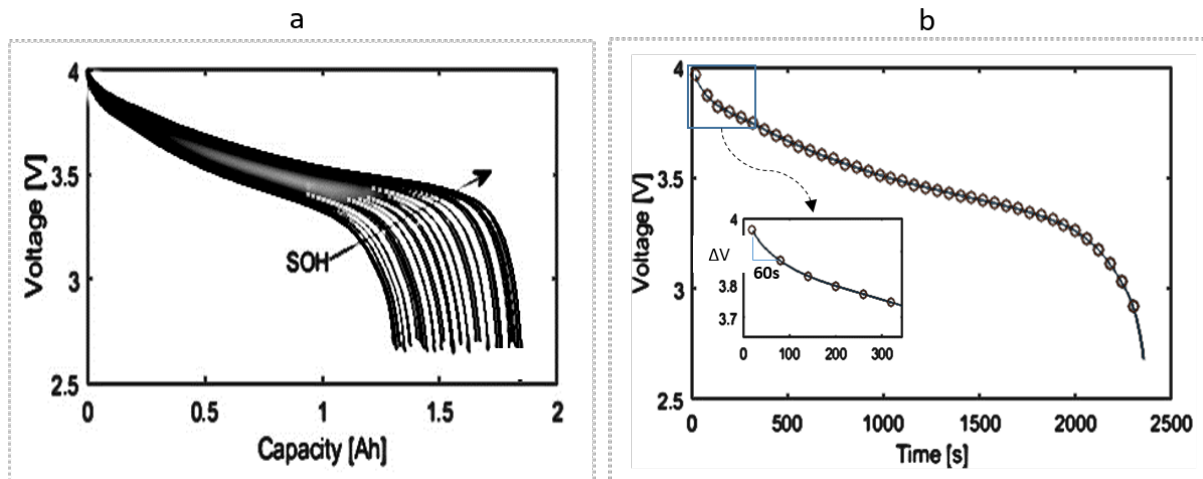

Figure 4. (**a**) Voltage vs. capacity plot: the output of a constant current discharge profile. Each curve represents a single cycle. The arrow is the reverse direction of aging across cycles. (**b**) The data points are taken at a 60 s interval to obtain the variation of voltage used as the feature variable.

As the lithium-ion battery ages, the maximum and minimum voltages of the battery are reached quicker during charge and discharge, respectively, as a consequence of the degraded charge/energy capacity. The degraded capacity results from the consumed active electrode and results in a faster voltage polarization. The battery *SOC* is an important battery parameter upon which many variables and other parameters are dependent. *SOC* is commonly defined as the percentage of the maximum possible charge that is present inside a rechargeable battery. Thus, the *SOC* serves as the fuel gauge for batteries. *SOC* can be expressed as the ratio of current capacity and the capacity releasable at the end of discharge, as in Equation (12). Since *SOC* is not directly measurable at run-time, it is estimated by a BMS algorithm [40].

$$SOC(t) = SOC(t_o) - \frac{\int_o^t I(t)}{C} dt \tag{12}$$

$SOC(t_o)$ is the initial *SOC* while *C* is the releasable capacity computed by coulomb count. Accurate measurement of *SOC* is necessary to minimize cumulative error. The algorithm demonstrated in [40] can be helpful to achieve this. Similar to the computation of battery releasable capacity by coulomb count, the deliverable energy is computed by integration of the product of voltage and current over time. For a constant discharge current, the equation for energy (*E*) computation, is shown in Equation (13).

$$E(t) = E(t_0) - \int_o^t I(t)v(t)dt \tag{13}$$

The releasable energy, *E*, provides information about the integrated power that can be released from the battery. Although *E* appears to be dependent on the other feature variables (voltage and SOC), it provides a unique feature as an integral of power. Figure 5a shows a three-dimensional plot of SOC, voltage, and energy, while Figure 5b shows the variations in SOC, voltage, and energy.
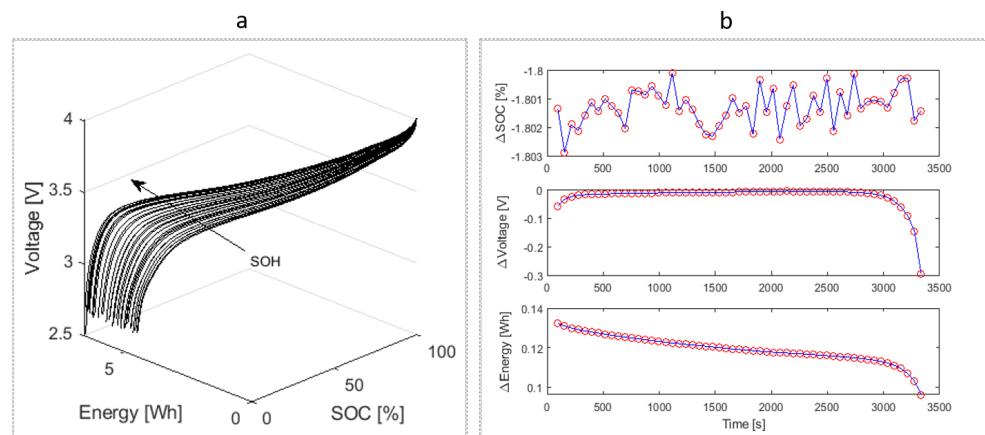
Figure 5. (**a**) A three-dimensional plot of SOC–capacity–voltage for a constant current discharge profile. Each curve represents a single cycle. The arrow is the reverse direction of aging across cycles. (**b**) The corresponding feature variables obtained as variations of SOC, capacity, and energy.

From Figure 5b, it can be seen that the variation of voltage between 300–3000 s (90–10% SOC) is almost constant. For this reason, SOH discrimination in this zone is poor with only voltage and SOC. On the other hand, the variation of energy does not show this behavior. The energy variable thus provides an additional feature for characterizing SOH.

2.6. Summary of Scheme Setup Procedure

The SOH estimation scheme using the PL-ELM algorithm as proposed in this paper is summarized below.

Model Input: [ΔV, ΔSOC, ΔE] $\in \mathbb{R}^{3 \times N}$
Model Output: SOH $\in \mathbb{R}^{1 \times N}$

- The dataset consists of voltage that is acquired by aging the battery with a constant current profile;
- SOC and energy are extracted from the dataset as described in Equations (12) and (13), respectively;
- The input of the PL-ELM model is computed as a discrete point variation of the voltage, SOC, and energy at a 90 s interval, as in Figure 4b;
- The discrete point output is computed as the mean SOH within the same interval;
- The discretized input and output data are used to train the model, defining the parameters (weight and bias) of the PL-ELM model;
- The number of hidden layer neurons is adjusted to obtain an optimum model, allowing for a compromise between performance and complexity;
- The designed SOH estimation model is then validated using the dataset of other batteries of the same type that have been aged under the same current load conditions as in the training.

---

3. Results and Discussion

3.1. PL-ELM Model Training

The results shown in this section demonstrate the estimation of SOH with deterministic PL-ELM to enhance generalization over a set of batteries of similar specifications. The result herein demonstrates that although each battery used in the test had a distinct distribution, the estimation algorithm performance was optimum over the entire set. In other words, the designed model is not affected significantly by the drift problem. The PL-ELM estimation model performance was compared with other methods, including the deterministic ELM algorithm and an ICA-based algorithm reported in [25]. The performance evaluation was computed with root mean square error (RMSE) and mean absolute error (MAE) indices, considering a 99% confidence interval of the errors. This consideration helps to provide information about the percentage of errors outside the defined error bound. The performance results were compared with other algorithms in the literature. The root

mean square error (RMSE) and mean average error (MAE) indices were computed using Equations (14) and (15), respectively.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{14}$$

where N is the total number of samples. The quadratic scoring RMSE is desirable to amplify large errors while diminishing small ones. A linear scoring MAE is defined as

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{15}$$

The inputs to the PL-ELM network or the model training features are the discrete variation of the indicator variables in Table 1. The structure of the PL-ELM was fixed by varying the number of neurons at the training phase. The number of neurons chosen was a compromise between model complexity and accuracy. To fix the model structure, the number of neurons was varied from 1 to 20, as seen in Figure 6c. It may be sufficient to use about twelve (12) neurons to obtain an optimal model in terms of complexity. However, by increasing the number of neurons sufficiently (20 neurons in this case), the estimation accuracy of the model was improved. The scheme was designed and simulated with Matlab software, and the total number of 5523 data samples from the B0007 dataset was used to train the model. Figure 6a shows the result of the training phase of PL-ELM while Figure 6b shows the normal distribution of the error. Figure 6c shows the performance based on model complexity when the number of neurons varied.
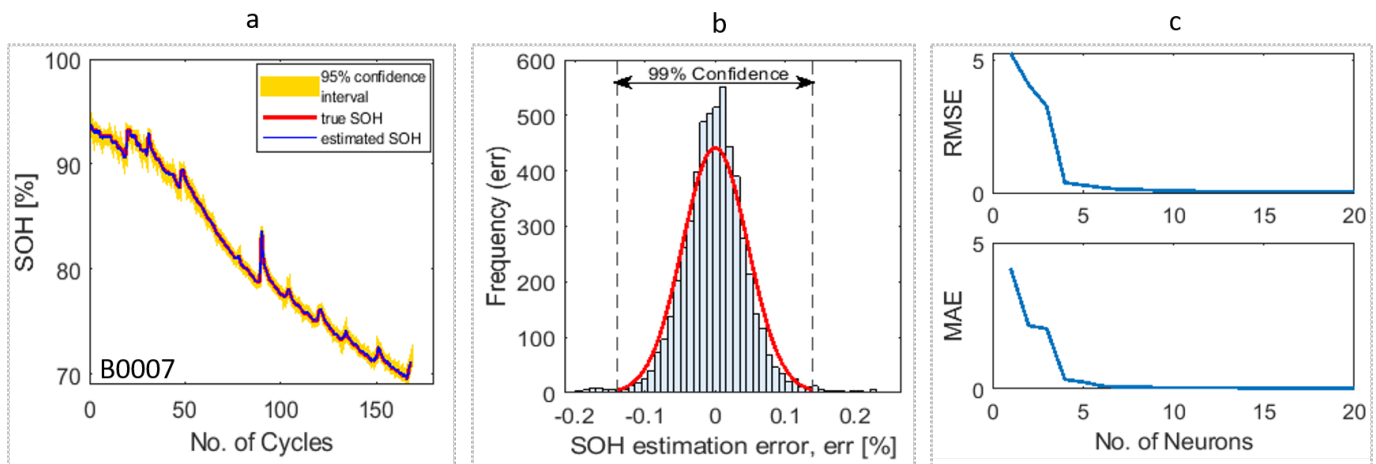

Figure 6. The result of parallel layer extreme learning machine (PL-ELM) model trained with B0007 dataset and consideration of a 99% confidence interval on the error. (**a**) The result of SOH estimation with 20 neurons. (**b**) Normal distribution of the error for 99% confidence interval (**c**) performance versus model complexity over 20 neurons.

Error estimation was obtained for a 99% confidence interval. The confidence interval is in the range provided by the error bound (EB). This provides information about the error deviation. The resulting model error is an approximation of natural distribution. The percentage out of bound error ($\%E_{OB}$) indicates the percentage of the error that is outside the 99% confidence interval. In the training phase, only about 1.57% of the training data points were outside the 99% confidence interval bound at $\pm0.14\%$ estimation error. The $\%E_{OB}$ for the validation phase is also listed in Table 2.

The RMSE of the trained model performance was computed as 0.046, while the mean average error (MAE) was 0.034 for twenty (20) neurons in the parallel layer network. From the performance results for 99% confidence interval shown in Table 2, the difference

between RMSE and MAE is useful for understanding the error variations in the predictions. The small difference between the RMSE and the MAE errors indicates low error variance in the prediction set. It can be seen that even with the low errors deviation in the EB, the $\%E_{OB}$ is minimal.

Table 2. SOH estimation performance comparison between PL-ELM and incremental capacity analysis (ICA) model in training and validation. The error bound (EB) and out of bound error ($\%E_{OB}$) information are also provided.

		PL-ELM				ICA Model
		RMSE	MAE	EB	$\%E_{OB}$	RMSE
Training	B0007	0.046	0.034	$(-0.14, 0.14)$	1.570	0.66
Validation	B0005	0.362	0.345	$(0.02, 0.67)$	2.037	0.87
	B0006	0.473	0.355	$(-0.62, 1.31)$	3.240	2.49
	B0018	0.170	0.158	$(-0.04, 0.36)$	0.250	-

The model performance is influenced by the choice of the discrete point variation interval of the input features. The choice of the interval was based on model tuning for accuracy. The results of the interval variation are not fully reported in this work. However, it can be shown that too small an interval results in a noisy feature, while too large an interval results in too much loss of information. In both extreme cases, the model accuracy is impacted adversely. In this work, a sampling interval of 90 s is believed to be optimum. Using a 60 s interval increased the training RMSE from 0.046% to 0.053%. On the other hand, with a 120 s interval, the training RMSE was reduced to 0.043%. However, the maximum validation RMSE increased from 0.473% to 1.089%.

3.2. PL-ELM Model Validation

The training of the estimation model was done with battery B0007 dataset. The model was validated with other batteries of the same type: B0005, B0006, and B0018. All the batteries were aged with a constant current profile of 2 A. The performance of the designed model was compared with the result obtained with the ICA-based model using the same dataset as reported in [25]. These performance results are shown in Table 2. From the comparison, the proposed model shows an improved relative performance.

The variation of RMSE, as shown in Table 2, is in the range of 0.170–0.4736 on the validation phase. Moreover, the variation of the MAE range is in the range of 0.158–0.355. Figure 7 shows the validation of the PL-ELM model with batteries B0005, B0006, and B00018.
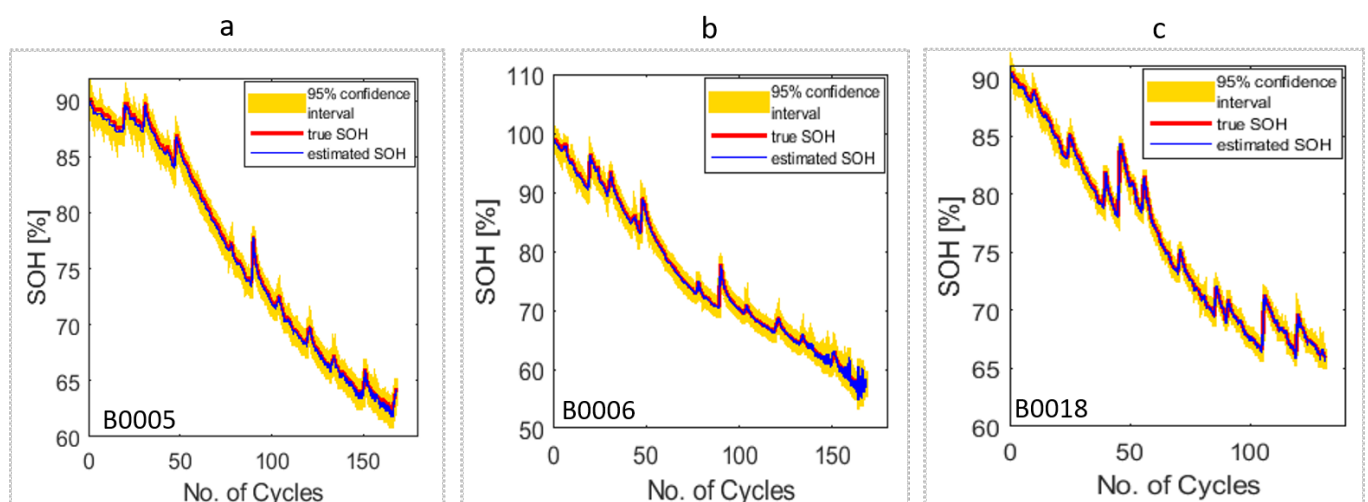

Figure 7. PL-ELM model validation: (**a–c**) model validated with B0005, B0006, and B0018 dataset, respectively.

It can be seen from Figures 6 and 7 that the aging pattern is different for the individual batteries despite the fact that they were subject to the same amount of load. The PL-ELM model can cope with the nonlinearity of SOH for all the batteries with reasonable accuracy. The nonlinearity handling strength is attributed to the two parallel (or independent) layers that are nonlinearly activated and projected into the feature space.

3.3. Model Comparison with Deterministic ELM and Demonstration of the Drift Problem

Although ELM is the building block of PL-ELM, it does not have as much nonlinearity handling capacity as the PL-ELM. ELM may be suitable for systems with fairly low nonlinear behavior. Like many machine learning algorithms, the ELM model often experiences a drift when the level of dissimilarity between the distribution of the training and the validation dataset is high. To compare results between ELM and PL-ELM, the experiment was repeated as summarized in Section 2.6, replacing the PL-ELM with the deterministic ELM single-layer model. The results of SOH estimation using the ELM model are shown in Figure 8.

As in the PL-ELM, the ELM model was trained with the battery B0007 dataset and validated with the B0005, B0006, and B0018 datasets. The training RMSE performance of the battery was 0.245, while the MAE was 0.191. In the validation, the RMSE was in the range of 0.501–1.563, while the MAE was in the range of 0.361–0.907. The ELM model performance is reported in Table 3. The model estimation tends to drift away from the true SOH across the entire cycle. Although the model results shown here were obtained for twenty (20) neurons, the drift problem never disappears regardless of the number of neurons used.
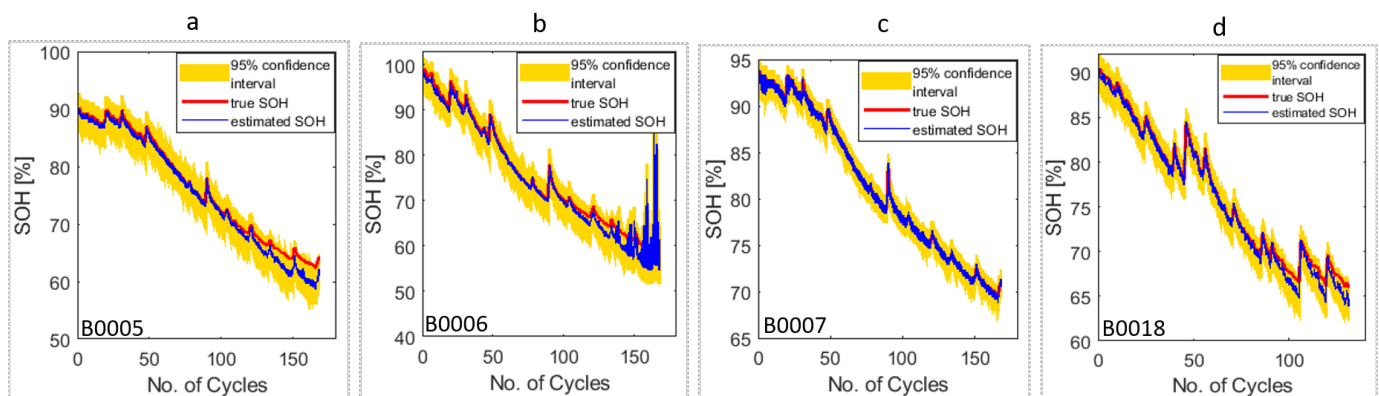

Figure 8. Extreme learning machine (ELM) model training and validation: (**a**) model trained with the B0007 dataset. (**b**–**d**) Model validated using the B0005, B0006, and B0018 datasets, respectively.

Table 3. SOH estimation performance of deterministic ELM model for training and validation. The EB and $\%E_{OB}$ information are also provided.

		RMSE	MAE	EB	$\%E_{OB}$
Training	B0007	0.245	0.191	$(-0.74, 0.74)$	0.615
Validation	B0005	1.117	0.762	$(-1.70, 3.22)$	1.24
	B0006	1.563	0.907	$(-3.18, 4.82)$	0.19
	B0018	0.501	0.361	$(-0.79, 1.46)$	0.89

From the results of Tables 2 and 3, PL-ELM and deterministic ELM can be compared based on performance. The training RMSE of the deterministic ELM increased from 0.046% to 0.245% when compared with that of the PL-ELM. The MAE increased from 0.034% to 0.191%. The maximum validation RMSE of the ELM increased from 0.473% to 1.563%, while the maximum MAE increased from 0.355% to 0.907%. These highlight the improvement in the performance of the PL-ELM as compared with the ELM. The poor performance of the

deterministic ELM model is evident in the SOH estimation results and in the large EB in the validation phase. The model drifting that affects the ELM becomes more pronounced as the battery ages.

It can be seen from Figures 7 and 8 that degradation/aging distribution varies across the individual cells, despite the fact that they were subjected to the same load condition and were of the same cell type. This variation in distribution results from cell-to-cell parameter variation [23]. In conventional machine learning models, the training and test datasets are required to be identically distributed [19] to avoid drift. Moreover, conventional neural network-based models are trained by optimization. The probable resulting models obtained at some local minima can be suboptimal and may experience drift on other batteries of the same type.

The model was set up, simulated, and validated using Matlab and Simulink software, version 2020b. To understand the real-time performance of the model in terms of computational cost and memory consumption, the algorithm was deployed to a Texas Instruments Delfino™ F28379D device. The 32 bit duo core MCU was connected to Simulink via serial connection and a processor in the loop (PIL) model simulation was performed. On the basis of the simulation, an average execution speed of 93 µs was recorded with 0.9305% average CPU utilization. This suggests that the model is suitable for online applications.

4. Summary and Conclusions

Reliability of energy supply is a key concern in lithium-ion battery applications, such as in electric vehicles. The reliability of a lithium-ion battery was analyzed on the basis of the SOH index. The SOH is computed based on relative capacity across the lifetime of a battery, otherwise known as capacity fading. In this work, SOH was experimentally characterized using the voltage, SOC, and energy releasable by the battery.

An approach using deterministic and stable PL-ELM was applied to develop SOH estimation models. This solution did not lead to an optimization problem. Moreover, the resulting model showed no drift when applied for the estimation of SOH on other batteries of the same type under equivalent load conditions.

The PL-ELM model was trained and validated using the discrete variation of the characterizing or indicator variables as an input to the model. The training and validation was experimentally possible with the help of the dataset from NASA PCoE. The RMSE of the validated model varies from 0.046% to 0.473% and the MAE error from 0.034% to 0.355% with a minimum error deviation beyond the 99% confidence interval for the battery sets tested. An improved relative performance was obtained for the proposed model in comparison with the other highlighted algorithms.

The model was tested on an MCU board using a PIL simulation with an execution time of 93 µs in real time and with 0.9305% CPU occupation. On the basis of the low model complexity, the performance, and the computation efficiency, the model is suggested as being suitable for online applications.

Author Contributions: Conceptualization, E.E.; Data curation, E.E.; Formal analysis, E.E.; Investigation, A.T. and M.S.; Methodology, E.E.; Project administration, A.T.; Resources, A.T.; Software, E.E.; Supervision, A.T.; Validation, E.E.; Visualization, M.S.; Writing—original draft, E.E.; Writing—review and editing, E.E., A.T. and M.S. All authors have read and agreed to the published version of the manuscript.

References

1. Gantenbein, S.; Schönleber, M.; Weiss, M.; Ivers-Tiffee, E. Capacity Fade in lithium-Ion Batteries and Cyclic Aging over Various State-of-Charge Ranges. *Sustainability* **2019**, *11*, 6697.
2. Vetter, J.; Novak, P.; Wagner, M.R.; Veit, C.; Moller, K.C.; Besenhard, J.O.; Winter, M.; Wohlfahrt-Mehrens, M.; Vogler, C.; Hammouche, A. Aging mechanisms in lithium-ion batteries. *J. Power Sources* **2005**, *147*, 269–281. [CrossRef]

3. Victor, A.; Jeffrey F. lithium-ion Battery Anode Aging Mechanisms. *Materials* **2013**, *6*, 1310–1325.
4. Groot, J. State-of-Health Estimation of lithium-ion Batteries: Cycle Life Test Methods. Master's Thesis, Chalmers University of Technology, Gothenburg, Sweden, 2012. Available online: http://publications.lib.chalmers.se/records/fulltext/154090/154090.pdf (accessed on 11 April 2021).
5. Ramadass, P.; Bala, H.; Ralph, W.; Branko, N.P. Mathematical modelling of the capacity fade of lithium-ion cells. *J. Power Sources* **2003**, *123*, 230–240. [CrossRef]
6. Paulo, K.; Shengbo, E.L.; Dongsuk, K. Simplification of pseudo two-dimensional battery model using dynamic profile of lithium concentration. *J. Power Sources* **2015**, *286*, 510–525.
7. Linlin, F.; Junqiu, L.; Bo, P. Online Estimation and Error Analysis of both SOC and SOH of Lithium-ion Battery based on DEKF Method. ICAE2018—The 10th International Conference on Applied Energy. *Energy Procedia* **2019**, *158*, 3008–3013.
8. Hansen, T.; Wang, C.J. Support vector-based battery state of charge estimator. *J. Power Sources* **2005**, *141*, 351–358. [CrossRef]
9. Wang, J.; Guo, J.; Ding, L. An adaptive Kalman filtering based State of Charge combined estimator for electric vehicle battery pack. *Energy Convers. Manag.* **2009**, *50*, 3182–3186.
10. He, J.; Wei, Z.; Bian, X.; Yan, F. State-of-Health Estimation of Lithium-Ion Batteries Using Incremental Capacity Analysis Based on Voltage–Capacity Model. *IEEE Trans. Transp. Electrif.* **2020**, *6*, 2. [CrossRef]
11. Arijit, G.; Amit, P. State of Health Estimation of Lithium-Ion Batteries Using Capacity Fade and Internal Resistance Growth Models. *IEEE Trans. Transp. Electrif.* **2018**, *4*, 1.
12. Christopher, R.; Osama, A.M. Adaptive Battery Management and Parameter Estimation Through Physics-Based Modeling and Experimental Verification. *IEEE Trans. Transp. Electrif.* **2016**, *2*, 4.
13. Liu, Z.; Dang, X. A New Method for State of Charge and Capacity Estimation of Lithium-Ion Battery Based on Dual Strong Tracking Adaptive H Infinity Filter. *Math. Probl. Eng.* **2018**, *5218205*, 18. [CrossRef]
14. Wei, Z.; Zhao, J.; Ji, D.; Tseng, K.J. A multi-time scale estimator for battery state of charge and capacity dual estimation based on an online identified model. *Appl. Energy* **2017**, *204*, 1264–1274. [CrossRef]
15. Bonfitto, A.; Ezemobi, E.; Feraco, S.; Tonoli, A.; Amati, N.; Hegde, S. State of Health Estimation of Lithium Batteries for Automotive Applications with Artificial Neural Networks. In Proceedings of the 2019 AEIT International Conference of Electrical and Electronic Technologies for Automotive, AEIT AUTOMOTIVE, Turin, Italy, 2–4 July 2019.
16. Chang, C.; Liu, Z.; Huang, Y.-W.; Wei, D.-Q.; Zhang, L. Estimation of Battery State of Health Using Back Propagation Neural Network. *Comput. Aided Draft. Des. Manuf.* **2014**, *24*, 60.
17. Kim, J.; Yu, J.; Kim, M.; Kim, K.; Han, S. Estimation of lithium-ion Battery State of Health based on Multilayer Perceptron: As an EV Application. *IFAC Pap.* **2018**, *51*, 392–397. [CrossRef]
18. Wu, J.; Wang, Y.; Zhang, X.; Chen, Z. A novel state of health estimation method of lithium-ion battery using group method of data handling. *J. Power Sources* **2016**, *327*, 457–464. [CrossRef]
19. Bowen, J.; Yong, G.; Lifeng, W. A State of Health Estimation Framework for Lithium-Ion Batteries Using Transfer Components Analysis. *Energies* **2019**, *12*, 2524.
20. Zhang, S.; Zhai, B.; Guo, X.; Wang, K.; Peng, N.; Zhang, X. Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks. *J. Energy Storage* **2019**, *26*, 100951. [CrossRef]
21. Weng, C.; Sun, J.; Peng, H. A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring. *J. Power Sources* **2014**, *258*, 228–237. [CrossRef]
22. Bian, X.; Wei, Z.; He, J.; Yan, F.; Liu, L. A Novel Model-based Voltage Construction Method for Robust State-of-health Estimation of Lithium-ion Batteries. *IEEE Trans. Ind. Electron.* **2020**. [CrossRef]
23. Santhanagopalan, S.; White, R.E. Quantifying Cell-to-Cell Variations in lithium-ion Batteries. *Int. J. Electrochem.* **2012**, *2012*, 395838. [CrossRef]
24. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain Adaptation via Transfer Component Analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 199–210. [CrossRef]
25. Tang, X.; Zou, C.; Yao, K.; Chen, G.; Liu, B.; He, Z.; Gao, F. A fast estimation algorithm for lithium-ion battery state of health. *J. Power Sources* **2018**, *396*, 453–458. [CrossRef]
26. Pan, H.; Lü, Z.; Wang, H.; Wei, H.; Chen, L. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. *Energy* **2018**, *160*, 466–477. [CrossRef]
27. Razavi-Far, R.; Chakrabarti, S.; Saif, M.; Zio, E.; Palade, V. Extreme Learning Machine Based Prognostics of Battery Life. *Int. J. Artif. Intell. Tools* **2018**, *27*, 1850036. [CrossRef]
28. Henríquez, P.A.; Ruz, G.A. Extreme learning machine with a deterministic assignment of hidden weights in two parallel layers. *Neurocomputing* **2017**, *226*, 109–116. [CrossRef]
29. Tavares, L.D.; Saldanha, R.R.; Vieira, D. Extreme learning machine with parallel layer perceptrons. *Neurocomputing* **2015**, *166*, 164–171. [CrossRef]
30. Li, G.; Zou, J. Multi-parallel Extreme Learning Machine with Excitatory and Inhibitory Neurons for Regression. *Neural Process. Lett.* **2019**, *51*, 1579–1597. [CrossRef]
31. Qu, B.Y.; Lang, B.F.; Liang, J.J.; Qin, A.K.; Crisalle, O.D. Two-hidden-layer extreme learning machine for regression and classification. *Neurocomputing* **2016**, *175*, 826–834. [CrossRef]

32. Tansel, D.; Ender, S. Evolutionary parallel extreme learning machines for the data classification problem. *Comput. Ind. Eng.* **2019**, *130*, 237–249.

33. Saha, B.; Goebel, K. *"Battery dataset"*, *NASA Ames Prognostics Data Repository*; NASA Ames Research Center: Moffett Field, CA, USA, 2007. Available online: http://ti.arc.nasa.gov/project/prognostic-data-repository (accessed on 11 April 2021).

34. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]

35. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of International Joint Conference on Neural Networks, IJCNN 2004, Budapest, Hungary, 25–29 July 2004.

36. Ömer, F.E.; Yılmaz, K. A detailed analysis on extreme learning machine and novel approaches based on ELM. *Am. J. Comput. Sci. Eng.* **2014**, *1*, 43–50. Available online: https://www.semanticscholar.org/paper/A-Detailed-Analysis-on-Extreme-Learning-Machine-and-Ertugrul-Kaya/08e6fcca9745ee3cc5e14b56d7472c068f5d94d8 (accessed on 11 April 2021).

37. Krömer, P.; Platoš, J.; Snášel, V. Differential evolution for the optimization of low-discrepancy generalized Halton sequences. *Swarm Evol. Comput.* **2020**, *54*, 100649. [CrossRef]

38. Martin, M.; Adel, G.A. *Closer Look at State of Charge (SOC) and State of Health (SOH) Estimation Techniques for Batteries*; TA15421-0-2/17; Analog Devices Inc.: Norwood, MA, USA, 2017.

39. Yang, J.; Yu, J.; Tang, D.; Dai, J. A closed-loop voltage prognosis for lithium-ion batteries under dynamic loads using an improved equivalent circuit model. *Microelectron. Reliab.* **2019**, *100–101*, 113459. [CrossRef]

40. Bonfitto, A.; Feraco, S.; Tonoli, A.; Amati, N.; Monti, F. Estimation accuracy and computational cost analysis of artificial neural networks for state of charge estimation in lithium batteries. *Batteries* **2019**, *5*, 47. [CrossRef]