Li-ion Battery Aging with Hybrid Physics-Informed Neural Networks and Fleet-wide Data

Renato G. Nascimento[1], and Matteo Corbetta[2], and Chetan S. Kulkarni[3], and Felipe A. C. Viana[4]

[1,4] *Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL, 32816, USA*
renato.gn@knights.ucf.edu
viana@ucf.edu
[2,3] *KBR LLC, NASA Ames Research Center, Moffett Field 94035 CA, USA*
matteo.corbetta@nasa.gov
chetan.s.kulkarni@nasa.gov

ABSTRACT

In this work, we propose a hybrid model for Li-ion battery discharge and aging prediction that leverages fleet-wide data to predict future capacity drops. The model is built upon an hybrid approach merging physics-based and empirical equations, as well as neural network models in a recurrent neural network cell. The hybrid physics-informed neural network can predict voltage discharge cycles given the loading profile, and estimate the used capacity of the battery under random-loading conditions by tracking aging parameters connected to the residual capacity of the battery. By merging information on the battery aging parameters with existing fleet-wide aging data, the model can predict the future residual capacity of the battery that is being monitored, and therefore enable predictions of voltage discharge curves far ahead in the battery life cycle. We validated the approach using the NASA Prognostics Data Repository Battery data-set, which contains experimental data on Li-ion batteries discharged at random loading conditions in a controlled environment. The approach also allows the identification of discrepancies between the battery aging trend and the trend observed at the fleet level, so that batteries behaving differently from the rest of the fleet can be subject to closer monitoring and further testing to refine predictions.

1. INTRODUCTION

Lithium-ion batteries are commonly used to power small and large electric vehicles, including both ground vehicles, like hybrid and electric cars driven by thousands of people every day, and future unmanned aircraft vehicles (UAVs) (Friedrich & Robertson, 2015; Madavan et al., 2016; Russell, 2019). Therefore, the ability to model both the state of charge as

well as battery health is very important for safe, reliable and affordable operation of vehicles fleets. Even though models based on first principles are accurate and trustworthy, the complex electro-chemistry that governs battery discharge and aging makes it hard to build and use such models for in-time monitoring of battery conditions. Moreover, the careful tuning or estimation of high-fidelity model parameters hampers the straightforward deployment in the field (Karthikeyan, Sikha, & White, 2008). Alternatively, reduced order models have the advantage of capturing the macroscopic dynamic behavior without the need for heavy computation, at the cost of some precision loss.

Reduced-order physics-based models are built by carefully simplifying the physics/chemistry such that computational cost is dramatically reduced while the overall behavior of the system is still captured. This approach can lead to a number of parameters to be estimated based on data as well as residual model-form uncertainty; a property shared with machine learning models. The latter are solely built on the basis of data, and can still capture unexpected non-linearities. The drawback is that, to generalize well, traditional machine learning tends to require large number of data points hard to retrieve in many scientific and engineering fields like, for example, the field of battery discharge and especially degradation prediction.

In this paper, we present a hybrid modeling approach for tracking and forecasting battery aging based on ''as-used'' conditions. Our approach directly implements a reduced-order model based on Nernst and Butler-Volmer equations within a deep neural network framework. While most of the input-output relationship is captured by reduced-order models, the data-driven kernels reduce the gap between predictions and observations. The hybrid model aims at estimating the overall battery discharge, and a multilayer perceptron (MLP) strategically placed within empirical equations models the battery internal voltage. Battery aging, resulting in a residual capacity

drop, can be described by an increase in internal resistance and a drop in the amount of available Li-ions. Thus, we built a model using MLPs to predict battery performance far ahead in the future based on the cumulative energy drawn from the battery over time. We address the issue of building and updating the aging model by reducing the need for reference discharge cycles, which would be beneficial to operators since it eliminates the need to take the batteries out of commission. We compensate for the lack of reference discharge cycles using a probabilistic model that leverages previously available fleet-wide information on the degradation of similar batteries.

We validate our approach using data publicly available through the NASA Prognostics Center of Excellence Data Repository (Bole, Kulkarni, & Daigle, 2014b). Results showed that our hybrid battery prognosis model can be successfully calibrated, even with a limited number of observations, and the model can help optimizing battery operations by offering long-term forecast of battery capacity. The construction of the model and all computations showed in this paper were performed using Python programming language and the deep learning frameworks Keras (Chollet, 2015) and Tensorflow (Abadi et al., 2015). Other libraries we utilized include Matplotlib (Hunter, 2007), Numpy and Scipy (Virtanen et al., 2020).

The remaining of the paper is organized as follows. Section 2 presents a brief review of the core reduced-order battery model and the data-set utilized. Section 3 details our hybrid physics-informed neural network model, and Section 4 shows the approach to predict future battery degradation based on fleet-wide data and the results of the numerical experiments. Section 5 concludes the paper.

2. REDUCED-ORDER MODEL AND AVAILABLE DATA

We employed a reduced-order model developed in (Daigle & Kulkarni, 2013) and further refined in (Bole, Kulkarni, & Daigle, 2014a). It is based on the work presented in (Karthikeyan et al., 2008). The main goal of this model is to work in real-time; it simplifies the complex battery electro-chemistry so that voltage predictions can be carried out using nonlinear ordinary differential equations rather than partial differential equations. We briefly summaries the main equations of the model hereafter and refer to (Daigle & Kulkarni, 2013) for a thorough description. The battery output voltage is defined by:
$$V = V_{U,p} - V_{U,n} - V_0 - V_{\eta,p} - V_{\eta,n} \, ,$$
which will serve as output of the physics-informed model. The model uses Nernst's equation for the equilibrium potential:
$$V_{U,i} = U_0 + \frac{R\,T}{m\,F} \ln \frac{1 - x_i}{x_i} + V_{INT,i} \, ,$$
where the electrode (negative or positive) is indicated by the subscript $i = \{n, p\}$; U_0 is the reference potential; R is the universal gas constant; T is the electrode temperature; m

is the number of electrons transferred in the reaction; F is the Faraday constant; x is the mole fraction for the Lithium-intercalated host material; and $V_{INT,i}$ is the internal voltage and activity correction term, null in ideal conditions. Details about V_{INT} will be provided hereafter.

The mole fraction is computed as the ratio between the amount of Li-ion q in electrode i, and the amount of available (moving) Li-ions q^{max}:
$$x_i = q_i/q^{max} \, , \qquad \text{and} \qquad q^{max} = q_n + q_p \, .$$

In order to accommodate the concentration gradient at the surface of the electrode, the total volume of the battery is split into two control volumes, bulk and surface, and the concentrations of Li-ions are calculated accordingly (Daigle & Kulkarni, 2013). The diffusion rate from the bulk to the surface is:
$$\dot{q}_{bs,i} = \frac{1}{D} \left(c_{b,i} - c_{s,i} \right) \, ,$$

where D is the diffusion constant and subscripts refer to bulk b, surface s, and negative or positive electrode i, respectively. One of the challenges of this reduced-order model is the description of the internal voltage $V_{INT,i}$. It was originally described by fitting experimental data to the the Redlich-Kister expansion (Karthikeyan et al., 2008):
$$\begin{aligned} V_{INT,i}(x_i;\, \mathbf{A}_i) = & \frac{1}{m\,F} \sum_{k=0}^{N_i} A_{k,i} \\ & \left((2x_i - 1)^{k+1} - \frac{2x_i k\,(1 - x_i)}{(2x_i - 1)^{1-k}} \right) . \end{aligned}$$

The mole fraction x_i is the independent variable, the coefficients $A_{k,i}$ are identified through data-fitting, and the number of elements in the sum N_i is empirically-derived.

The solid-phase Ohmic resistance, electrolyte Ohmic resistance, and current collector resistance can be lumped together into R_0 to calculate the voltage drop: $V_0 = i_{app}\,R_0$, where i_{app} is the applied current.

The amount of available Li-ions q^{max}, and the lumped internal resistance R_0 utilized to compute the voltage drop are directly tied to battery aging, as one decreases and the other increases with the usage of the battery. They will be utilized as proxies to predict the internal degradation of the battery, and for the sake of brevity, we will refer to q^{max} and R_0 as aging parameters henceforth.

2.1. Data Description

The Randomized Battery Usage Data-set (Bole et al., 2014b) we utilized for validation contains data from battery discharge experiments in a controller environment (including constant environment temperature). It has been widely utilized in the
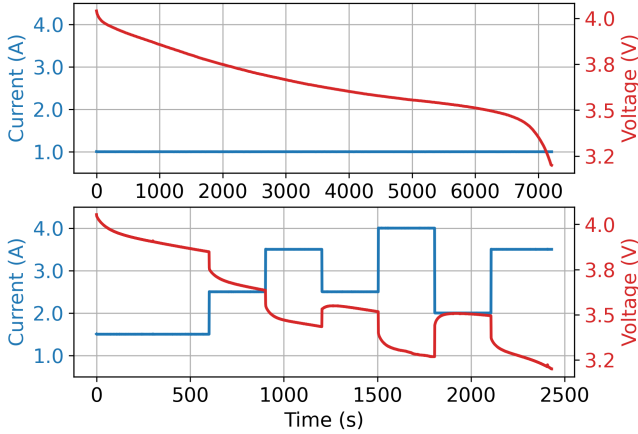
Figure 1. Example of constant (top panel) and randomized (bottom panel) loading discharge cycles from the NASA Prognostics Data Repository - Randomized Battery Usage Data-set.

PHM domain, and therefore, we report only a brief summary hereafter.

All batteries compose of a single cell, with a maximum voltage of approximately 4.2 V when fully charged, and all tests stopped when the output voltage reached 3.2 V. During the constant-loading tests, batteries were subject to 1 A current draw, while during the random-loading tests, the input current was randomized between 1 and 4 A, using a uniform distribution. Each sample was held constant for 300 seconds. So, $i = 1\,$A during the constant-loading experiments and $i \sim \mathcal{U}[1, 4]\,$A during the random-loading experiments. Figure 1 shows an example of constant (top) and random (bottom) loading discharge cycle.

3. HYBRID PHYSICS-INFORMED NEURAL NETWORKS: BATTERY CELL

We utilized a modeling approach proposed in past research (Nascimento & Viana, 2020; Nascimento, Fricke, & Viana, 2020; Viana, Nascimento, Dourado, & Yucesan, 2021), where the physics-based and empirical equations of the model are embedded within a recurrent neural network (RNN) cell (Goodfellow, Bengio, & Courville, 2016). The cell applies transformation to the state vector sequentially, similarly to a state-space formulation:

$$\begin{bmatrix} \mathbf{y}_t & \mathbf{h}_t \end{bmatrix}^\top = \mathbf{f}(\mathbf{u}_t, \mathbf{y}_{t-1}, \mathbf{h}_{t-1}) \,.$$

The subscript t represents the time discretization, $\mathbf{y} \in \mathbb{R}^{n_y}$ are the observable states, $\mathbf{h} \in \mathbb{R}^{n_h}$ are the internal states, $\mathbf{u} \in \mathbb{R}^{n_u}$ are input variables, and $\mathbf{f}(\cdot)$ defines the transitions between time steps (function of input variables and previous states) of hidden states and output. The hidden state vector $\mathbf{h}$ includes temperature, Li-ions available on each electrode (divided in surface and bulk), voltage drop and surface overpo-

tential. Input u is the required current i, and the output vector contains the output voltage V only:

$$\mathbf{h} = [T, V_0, V_{\eta,n}, V_{\eta,p}, q_{b,n}, q_{b,p}, q_{s,n}, q_{s,p}]^\top \,,$$
$$\mathbf{u} = u = i \,, \quad \text{and} \quad \mathbf{y} = y = V \,.$$

Figure 2 shows the RNN cell with physics-based and empirical equations from the original reduced-order model (blue blocks), MLPs utilized to model the interval voltage curve as a function of the mole fraction (green blocks), and aging parameters (the latter will also be tracked using MLPs that will be described later in the paper). The cell design resembles the one of the physics-based model, and the data-driven blocks (MLPs) are strategically placed within the model to mitigate the effect of model uncertainty by reducing the gap between between observed values and model predictions. The cell composes of two functions. The first is to estimate the hidden state of the battery $\mathbf{h}_t$ given $\mathbf{h}_{t-1}$ (called "Battery states" in Figure 2) and u_t ("Current"). The second function estimates the state-space output, i.e., output voltage V_t ("Voltage"), from the hidden state $\mathbf{h}_t$. The right panel of Figure 2 shows three voltage discharge cycles from randomized-input tests, and how the RNN cell predicts V_t given $\boldsymbol{h}_{t-1}, u_t$.

The data-driven blocks utilized to estimate the internal voltage on the positive p and negative n electrodes are two independent MLPs, receiving the mole fraction of the corresponding electrode as input:

$$V_{INT,p} = \text{MLP}_p(x_p; \mathbf{w}_p, \mathbf{b}_p) \,, \quad \text{and}$$
$$V_{INT,n} = \text{MLP}_n(x_n; \mathbf{w}_n, \mathbf{b}_n) \,.$$

Vectors $\mathbf{w}_\cdot, \mathbf{b}_\cdot$ define trainable weights and biases of the two data-driven models. From the insights obtained from the first model in (Daigle & Kulkarni, 2013) and (Bole et al., 2014a), we used a one neuron hidden layer with linear activation for the negative side of the electrode, thus assuming that $V_{INT,n}$ and x_n are (at most) linearly related. The positive side of the electrode, however, is characterized by a more complex relationship between input and output. Thus, we used a 2 hidden-layer network with 8 neurons in the first hidden layer, 4 in the second, and one neuron in the output layer. All hidden neurons have hyperbolic tangent activation function; see Table ?? for a summary of the MLP architectures. It should be noticed that trainable parameters as well as interval voltages are not observable, thus the training applies to elements deeply hidden in the model.

4. FLEET-WIDE APPROACH TO DEGRADATION MODELING

4.1. Tracking Aging of a Single Battery

We first observed that the relationship between the interval voltage $V_{INT,i}, \ i = \{n, p\}$ and the concentration overpotential $x_{s,i}, \ i = \{n, p\}$ is not affected by aging, or at least,
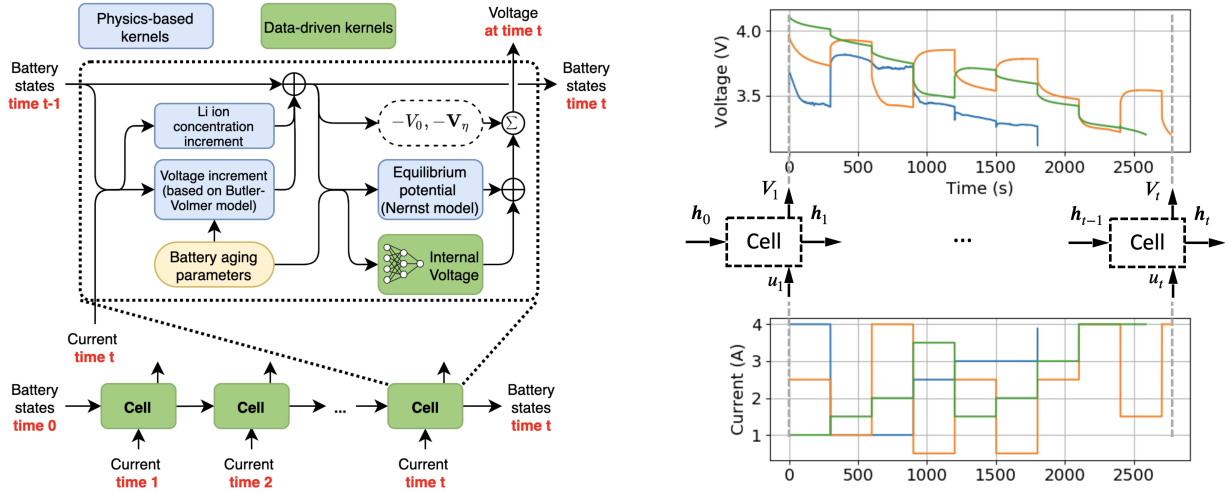
Figure 2. Battery hybrid RNN cell (left panel) and RNN implementation for time series prediction of output voltage given the input current (right panel).
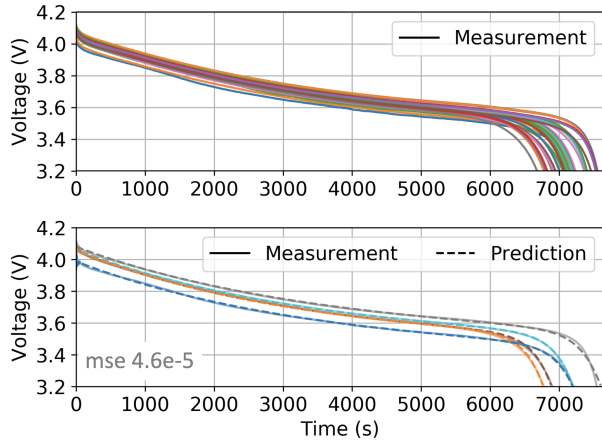

Figure 3. Constant-loading discharge curves used to train the internal voltage MLP models (top plot) and model predictions on a separate test set (bottom plot) with corresponding mean square error.

Layer	$V_{INT,p}$ (positive side)		$V_{INT,n}$ (negative side)	
	# neurons	activation	# neurons	activation
#1	8	tanh	1	linear
#2	4	tanh		
#3	1	tanh		

Table 1. Configuration of multilayer perceptrons used to approximate the behavior of the internal voltage ("tanh" and "lin" are the hyperbolic tangent and linear activation functions, respectively).

Layer	Mean		Variance	
	# neurons	Activation	# neurons	Activation
#1	4	elu	1	linear
#2	2	elu		
#3	1	linear		

Table 2. Configuration of variational multilayer perceptrons used for both q^{max} and R_0. ELU stands for exponential linear unit.

the effect appears negligible for the precision required for our life prediction purposes. Therefore, we utilized the first constant-loading discharge curves to train the interval voltage models. Table 1 details the design of each MLPs used for both $V_{INT,p}$ and $V_{INT,n}$. During this first training stage, the aging parameters of the battery were fixed to their default values obtained using reference discharge curves, assuming that the operator or manufacturer can run at least one reference discharge test per battery before its deployment in the field. Figure 3 shows the voltage discharge curves utilized for the training of $V_{INT,i}$, $i = \{n, p\}$ (top plot) and the predictions on a test set of the same type (bottom plot). The parameters of the internal voltage models $\mathbf{w}_i$, $\mathbf{b}_i$, $i = \{n, p\}$ are then fixed, so that their values will not change during usage.

The degradation of the battery will, instead, be tracked using the aging parameters (q^{max} and R_0). To appreciate the correlation of the selected aging parameters with the actual capacity of the battery (estimated from standard reference discharge cycles available in the data-set), Figure 4 shows capacity C and q^{max} as a function of E on a double y-axis[1]. The neural networks to track q^{max} and R_0 as a function of E are both variational MLPs (Graves, 2011; Kingma & Welling, 2014) and are further detailed in Table 2.

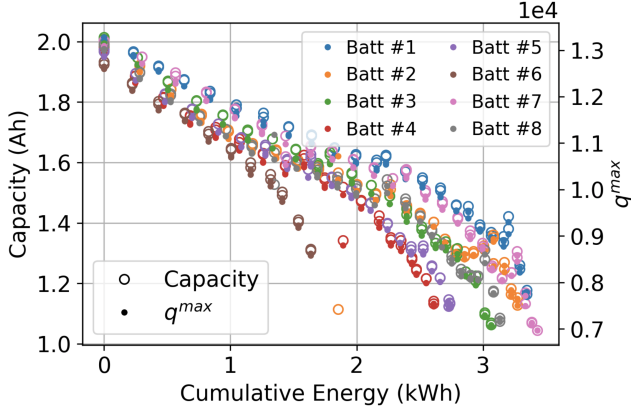[1] R_0 is also strongly correlated to C but it has not been shown for the sake of brevity.

Figure 4. Decrease of maximum capacity (left y-axis) and q^{max} (right y-axis) as a function of the cumulative energy drawn from the batteries.

Figure 5. Battery capacity used from random-loading discharge curves (small red dots) and reference discharge tests (large black dots)

4.2. Fleet-Based Aging Model

The goal of our fleet-wise approach is to exploit the correlation between C and $\{q^{max}, R_0\}$ to predict aging and future performance at any amount of cumulative energy E used, leveraging existing data from a fleet of similar battery types. To do so, we assume that aging data from a fleet of batteries have already been collected (in the form shown in Fig. 4). Thus, the first challenge lies in tuning a fleet-wide model to predict future C values for the battery that is being used, and such a prediction is updated as more usage data from the monitored battery are collected. However, batteries operating in the field are typically discharged with random loading conditions, while accurate values of q^{max}, R_0 can only be estimated using reference discharge cycles, where the applied loading is steady and very low. Data coming form the field cannot be directly utilized to estimate the trend of available q^{max} vs. E.

Figure 5 shows the used capacity for Battery #4 from the discharge curves under random loading conditions (small red dots), compared against the more reliable estimates of the actual residual capacity from the reference discharge tests at constant loading (large black dots). If we were to estimate aging by estimating q^{max} during random-loading cycles, there would be a large degree of uncertainty on how much capacity is left in the battery, which could be interpreted as uncertainty affecting the mole fraction value x on the positive electrode at the end of a full discharge cycle, as illustrated in Figure 6. By leveraging the assumption made above (that the battery is part of a fleet and aging data from those other batteries are available), we built aging models of all batteries in the data-set using variational inference, thus obtaining average and standard deviation of all weights in the MLPs mapping cumulative energy to the aging parameters (for more information about variational inference for neural networks, the reader is referred to (Graves, 2011)). Particularly, we built individual aging models for each battery, which correlates C to E, and
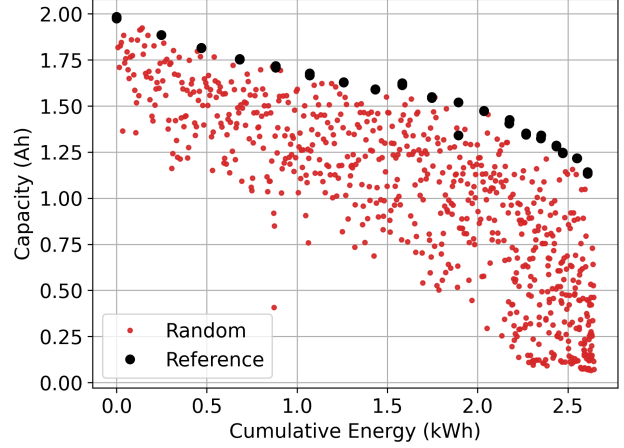
use ensemble averaging to predict the fleet behavior using future values for E. Afterwards, we used a scaling factor to estimate q^{max} from C, and then build a linear model, which we called γ model, to correlate q^{max} to R_0. For both models, the independent variable is the amount of energy drawn from the battery E, Figure 7. The models are described implicitly by the following equations:

$$q^{max}(E) = C(E)\,\alpha, \quad \text{and} \quad R_0(E) = \frac{q^{max}(E)}{\gamma(E)},$$

where $\gamma(E)$ is the linear model connecting the two aging parameters (Figure 7, right panel), and α is the scaling factor connecting $q^{max}(E)$ and $C(E)$. At this stage, we ignored the scatter of the data-points $\{E, q^{max}/R_0\}$, but future implementations may easily include the uncertainty caused by the linear approximation.

4.3. Updating Battery-specific Aging Model with Random Discharge Data

The next step of the proposed fleet-wide approach to discharge modeling is the filtering of the model forecast after the battery is deployed in the field and subject to random-loading discharge cycles. The goal is to merge the capacity estimate from the fleet-wide ensemble model, and the capacity estimates from the operation of the battery that is being monitored. To do so, we filter data at constant cumulative energy E, as summarized in Figure 8. The red dots represent the estimate of the used capacity during operations (as already observed in Figure 5). The green lines represent the predicted mean (thick) and confidence intervals (dashed) of the fleet-wide model, without information on the current battery capacity. The green area shows the fleet distribution at the specific cumulative energy value, and the red distribution represents the values of used
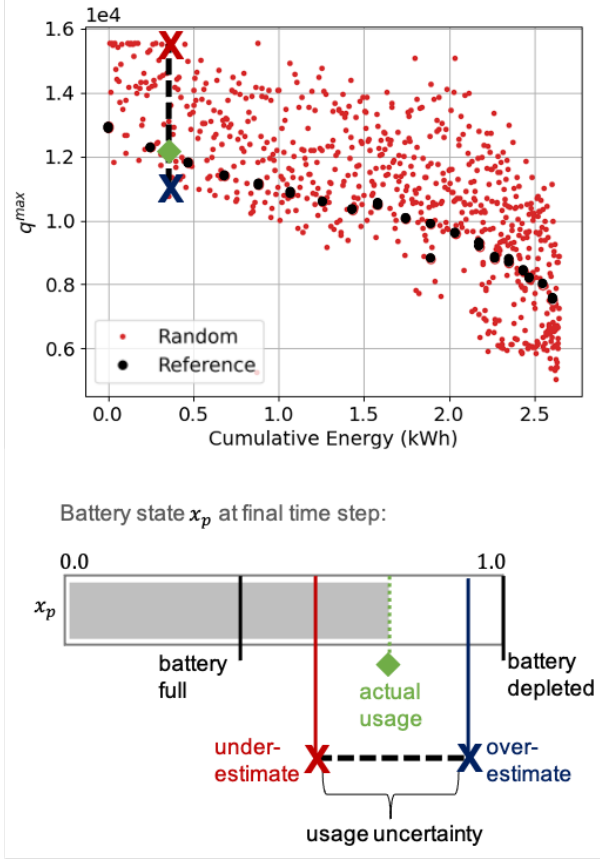
5

Figure 6. Random loading-induced uncertainty in capacity and x_p estimates. The scatter of used capacity under random loading conditions drives the scatter of q^{max} (top panel).

capacity collected up to that point. The purple distribution on the right panels, together with the purple lines, represent the filtered fleet distribution and the filtered prediction as described below.

We track the cumulative energy drawn from the battery during operation, and after a few discharge cycles have occurred, we compute the distribution of the battery ''used capacity,'' which is lower than or, at most, equal to, the available capacity of the battery (In the example case shown in Figure 8, we collected data up to $E = 0.5$ kWh to compute the used capacity distribution). Then, we compare such a distribution with the available capacity distribution from the fleet-wide ensemble model, at that same level of cumulative energy used. We use a sampling-based approach, drawing N samples from each distribution. Then, we filter the fleet-wide ensemble model distribution according to the following criterion: if the sample from the fleet-wide distribution is larger than the sample from the battery usage distribution, than keep the former; otherwise, discard it. The filtering can be summarized by this simple statement: if $C_b^{(i)} < C_f^{(i)}$, then accept $C_f^{(i)}$, and discard it otherwise. The subscripts b and f refer to battery and fleet distributions, respectively, and the superscript (i) indicates

the i-th sample, $\forall\, i = 1, \ldots, N$. The resulting, filtered distribution of available capacity is the one used to predict aging at later stages of the battery life (purple distribution in Figure 8), where we used $N = 10,000$.

This approach allows us to forecast the values of the aging parameters at higher values of cumulative energy, and thus make predictions of the future voltage discharge curves, as shown in Figure 9, where the forecast refers to $E = 2$ kWh for Battery # 4, which data were collected up to 0.5 kWh (blue distribution) and 1.5 kWh (red distribution). The figure was build upon the assumption that the loading conditions at 2 kWh (i.e., current draw profile) are perfectly known. In this particular case, the predicted average voltage drop does not change from 0.5 kWh to 1.5 kWh, however, the confidence intervals shrink significantly as more data have been recorded, and so there is less uncertainty on the future behavior of the battery at 2 kWh.

4.4. Detection of Anomalous Aging Patterns

Some batteries may experience different aging behavior with respect to the rest of the fleet. The reasons are multiple, such as, e.g., manufacturing variability of internal components and different environmental conditions the batteries are subject to during operation. This uncertainty can induce a different aging behavior since the very beginning of the battery life, as well as in later stages of the life cycles. In those situations, predictions from fleet ensemble-averaging are likely to be inaccurate. Battery # 6 in the data-set is just an example, as visible in Figure 4 and 7 (left panel). In order to monitor how well the battery aging aligns with the behavior observed at the fleet level, we built an indicator based on the battery delivered power.

The indicator is built on the area under the curve (AUC) of error between the predicted and the actual power delivered by the battery (estimated by multiplying the input current by the output voltage, $V\,I$). The resulting AUC is a measure of error in the discharge cycle energy, in Wh and while the model is predictive of the behavior of a given battery, error distribution is expected to be stationary. Therefore, as an indicator that the model is drifting, we monitor the changes in the energy error distribution as the battery accumulates usage.

Figure 10 shows the approach, where the changes in the Battery #4 power AUC with respect the first 0.25 kWh of cumulative operation. Figure 10a shows the error distribution for Battery # 4 in forms of slices at different cumulative energy (top panel), and as energy error vs. E (bottom panel). We further computed two distribution distance metrics, the Kolmogorov-Smirnov (KS) test and the Kullback-Leibler (KL) divergence to evaluate the discrepancy between Battery # 4 and the rest of the fleet, and the result is shown in Figure 10b.

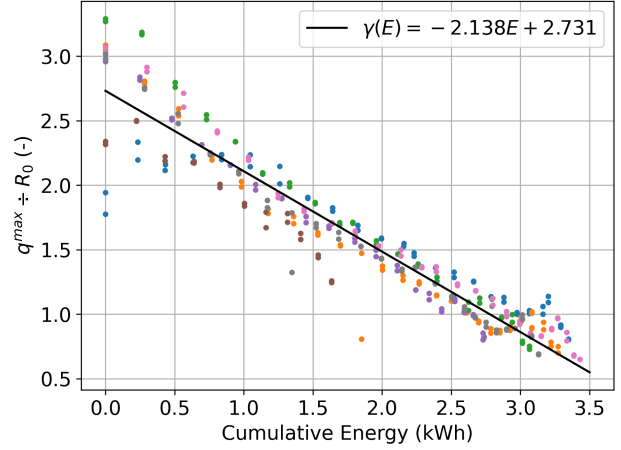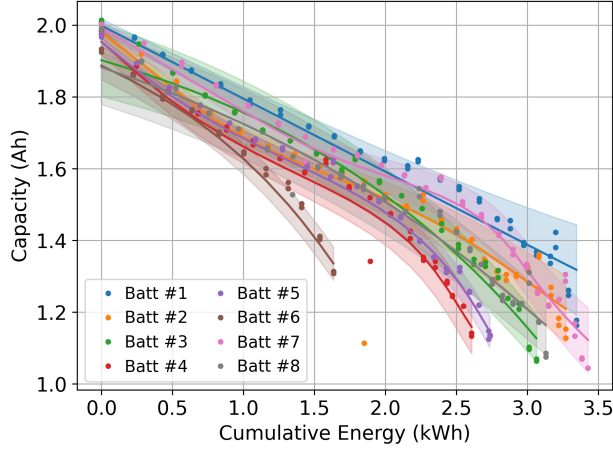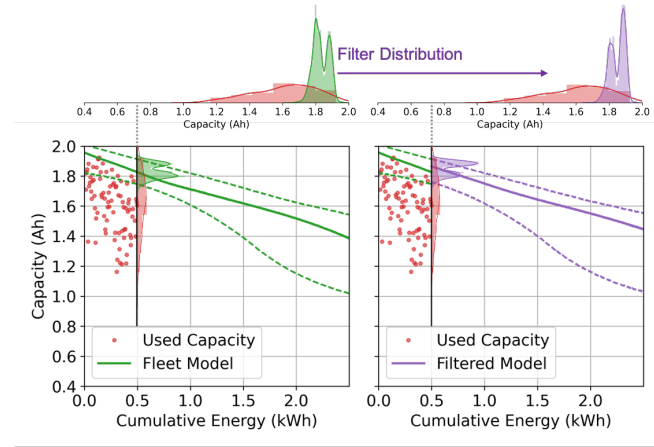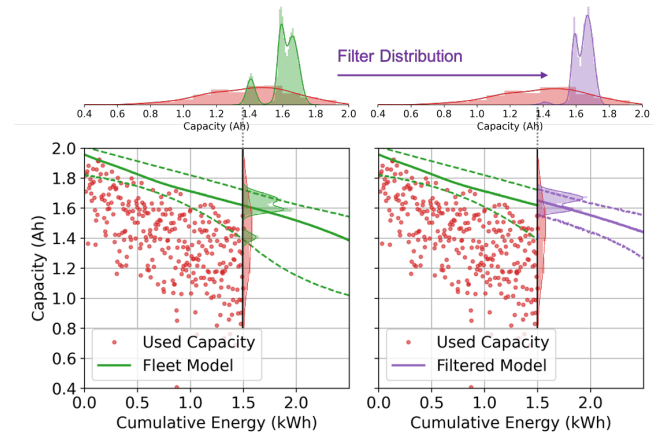The proposed approach leads to a model for Battery # 4 that

Figure 7. Variational model regressions for all 8 batteries in the data-set (left panel) and γ model to correlate R_0 to q^{max}.

(a) Filtering at 0.5 kWh.


(b) Filtering at 1.5 kWh.

Figure 8. Filtering process using the distribution of available capacity from the fleet-wide ensemble model. The two panels refers to distributions filtered at 0.5 kWh (a) and 1.5 kWh (b).

is predictive (meaning, small changes in energy error distribution) for much of the useful life, as indicated by both KS-statistic and KL-divergence. Both metrics do not show significant deviations from their reference values up to 2.0 kWh. After that, the model that used only a single reference discharge for calibration starts drifting significantly, indicating the need of additional reference discharge cycles for re-calibration of the model.

To validate the approach, we performed a cross-validation study of the KL-divergence[2] as a function of the cumulative energy for all 8 batteries in the fleet (i.e., the data-set). We first extracted Battery # 1 from the fleet, calculated the error distribution between the rest of the fleet and Battery # 1, and then computed the KL-divergence for the entire degradation profile. The process is repeated for all batteries in the set and results are shown in Figure 11. We can observe that for most batteries, KL-divergence values appear to be very consistent until around 2 kWh, where the nonlineary of the $\{E, C\}$ curves kicks in. This confirms the consistency in fleet behavior.

We also noticed that the approach can successfully detect batteries that are outliers with respect to the fleet. Battery # 6 shows large KL-divergence values early in its life, and that is expected by looking at the data-points collected from reference discharge cycles visible in Figure 4. This is also the case of Battery # 1, which shows discrepancies with the fleet behavior early in the process. The divergence in the energy error distribution happens since the early degradation data are used as reference to update the model; resulting into prediction errors in later stages of the degradation process. For these two batteries, the KL-divergence indicates the need of reference discharge cycles for re-calibration of models earlier than other

---

[2]We also cross-validated the results using the KS-statistic obtaining similar results, and therefore were not included in the paper for the sake of brevity.
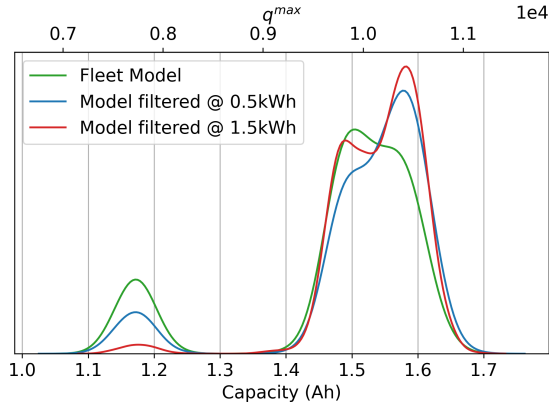
(a) Capacity and q^{max} distributions at 2.0 kWh.
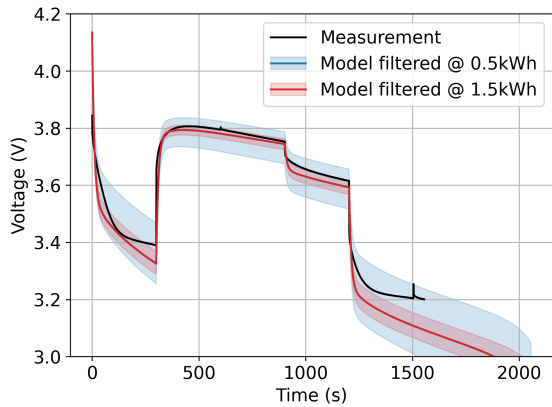
(b) Discharge cycle predictions at 2.0 kWh.

Figure 9. Aging and discharge forecast: capacity and q^{max} distributions (a), and prediction of voltage discharge cycle using model filtered with data up to 0.5 and 1.5 kWh (b).

batteries in the fleet.

4.5. Discussion on Assumptions and Drawbacks

The approach presented here relies on the assumption that a number of existing batteries have been monitored and their available capacity during operation was estimated accurately. If that were not the case, then the output of the fleet-wide model would be characterized by larger uncertainty, thus hindering the performance of the prediction.

Another important assumption regards the random loading conditions the monitored battery is subject to. The battery data-set utilized here contains discharge data at random current values, such that for a certain amount of cases, the used capacity is relatively close or approaches the available capacity. This can be observed in Figure 8, as some samples (red dots, used capacity of the battery) are relatively close to the average fleet-wide estimate of the available capacity. Instead, in case the battery were repeatedly subject to lower loading
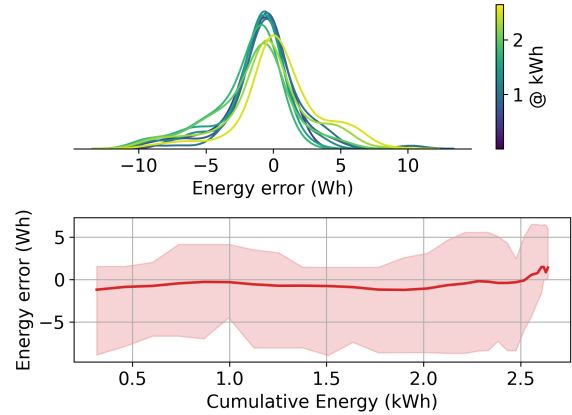
(a) Discharge cycles Energy (AUC of Power curve) error distributions (top) and evolution over Cumulative Energy (bottom).
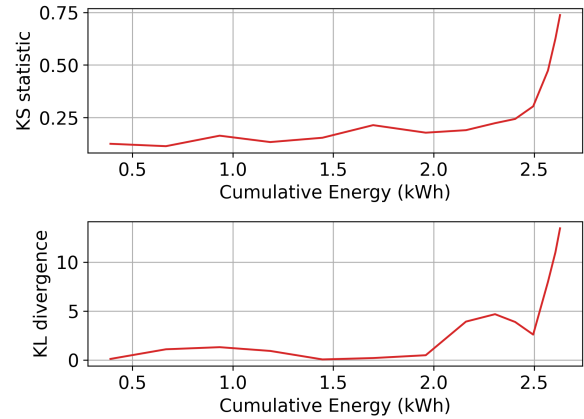
(b) Kolmogorov-Smirnov statistic (top) and Kullback–Leibler divergence (bottom).

Figure 10. Discharge cycles Energy (AUC of Power curve) error distributions over Cumulative Energy (a) and distance metrics (b).

conditions (for example, a UAV typically flying at speed and environmental conditions well below the capability of its battery), the fleet-wide distribution would remain unfiltered, and the predictions of future voltage discharge curve would be driven by the fleet information only.

Lastly, all batteries are subject to similar, although random, loading profiles. These discharge cycles were designed to bring the battery voltage down to 3.2 V (see Fig. 2). In the field, batteries could be recharged even if they were not fully or consistently depleted. This could introduce further errors in the modeling assumptions and contribute to premature divergence of models.

5. SUMMARY AND CLOSING REMARKS

In this paper, we demonstrated the use of a hybrid physics-informed neural network model to perform battery discharge
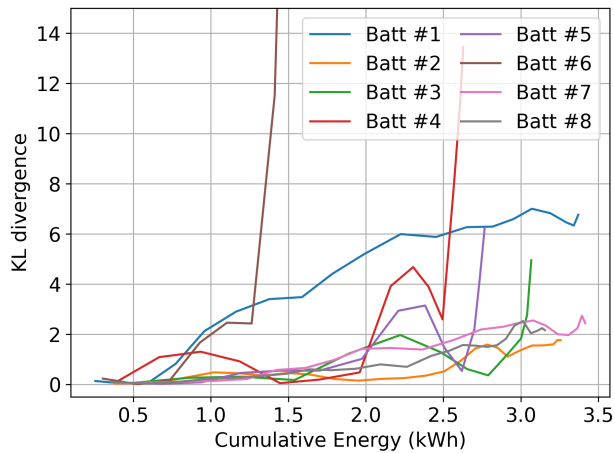
Figure 11. Kullback-Leibler (KL) divergence cross-validation.

and aging predictions using fleet-wide capacity data, thus reducing the need for reference discharge cycles required for precise estimates of battery residual capacity. The model leverages existing data from a fleet of similar batteries to accurately predict its aging behavior, without the need to decommission the battery for ad-hoc testing.

This work shows the potential of the hybrid model, where most of the model is driven by physics-based and empirical equations, while neural network blocks describes the relationships between hidden variables that are yet not well understood or characterized by large inter-specimen variability, and thus data-driven "blocks" can compensate for discrepancies between model predictions and observed behavior. The advantage of this hybrid model when compared to pure machine learning approaches, is that a limited amount of data is needed to train and calibrate the networks.

The aging modeling allows the prediction of the battery discharge behavior (in terms of voltage drop as a function of the input current) far-ahead in the future, thanks to the ensemble averaging of existing aging data from a fleet of similar batteries. The introduction of distance metrics like KS test and KL-divergence allows the detection of aging patterns start diverging (on a battery-specific basis), thus raising flags for ad hoc testing to refine aging predictions for those particular specimens.

The approach is easily scalable to more complex models, for example, of the entire UAV power-train composed of DC motor, electronic speed controller, and battery. Further steps of this research will aim at relaxing some of the hypothesis this approach is based upon already discussed in the previous section, and extending the model to other power-train elements and testing it on representative data from laboratory experiments.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from https://www.tensorflow.org/ (Software available from tensorflow.org)

Bole, B., Kulkarni, C., & Daigle, M. (2014b). Randomized battery usage data set. *NASA AMES prognostics data repository*, *70*.

Bole, B., Kulkarni, C. S., & Daigle, M. (2014a, September). Adaptation of an electrochemistry-based Li-ion battery model to account for deterioration observed under randomized use. In *Annual conference of the prognostics and health management society* (p. 9 pages). PHM Society.

Chollet, F. (2015). *Keras.* https://github.com/fchollet/keras. GitHub.

Daigle, M. J., & Kulkarni, C. S. (2013, September). Electrochemistry-based battery modeling for prognostics. In *Annual conference of the prognostics and health management society* (p. 13 pages). New Orleans, USA: PHM Society.

Friedrich, C., & Robertson, P. A. (2015). Hybrid-electric propulsion for aircraft. *Journal of Aircraft*, *52*(1), 176--189.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Retrieved from http://www.deeplearningbook.org

Graves, A. (2011). Practical variational inference for neural networks. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 24* (pp. 2348--2356). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/4329-practical-variational-inference-for-neural-networks.pdf

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90--95. doi: 10.1109/MCSE.2007.55

Karthikeyan, D. K., Sikha, G., & White, R. E. (2008). Thermodynamic model development for lithium intercalation electrodes. *Journal of Power Sources*, *185*(2), 1398--1407.

Kingma, D., & Welling, M. (2014). Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*.

Madavan, N., Heidmann, J., Bowman, C., Kascak, P.,

Jankovsky, A., & Jansen, R. (2016). A nasa perspective on electric propulsion technologies for commercial aviation. In *Proceedings of the workshop on technology roadmap for large electric machines, urbana-champaign, il, usa* (pp. 5--6).

Nascimento, R. G., Fricke, K., & Viana, F. A. C. (2020). A tutorial on solving ordinary differential equations using Python and hybrid physics-informed neural networks. *Engineering Applications of Artificial Intelligence*, *96*, 103996. Retrieved from `https://www.sciencedirect.com/science/article/pii/S095219762030292X` doi: 10.1016/j.engappai.2020.103996

Nascimento, R. G., & Viana, F. A. C. (2020). Cumulative damage modeling with recurrent neural networks. *AIAA Journal*. Retrieved from `https://arc.aiaa.org/doi/10.2514/1.J059250` doi: 10.2514/1.J059250

Russell, J. (2019, October). *NASA, U.S. industry aim to electrify commercial aviation.* Retrieved from `https://www.nasa.gov/feature/glenn/2019/nasa-us-industry-aim-to-electrify-commercial-aviation`

Viana, F. A. C., Nascimento, R. G., Dourado, A., & Yucesan, Y. A. (2021). Estimating model inadequacy in ordinary differential equations with physics-informed neural networks. *Computers and Structures*, *245*, 106458. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0045794920302613` doi: 10.1016/j.compstruc.2020.106458

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261--272. doi: https://doi.org/10.1038/s41592-019-0686-2