Exploring the Model Design Space for Battery Health Management

Bhaskar Saha¹, Patrick Quach², and Kai Goebel³

¹Mission Critical Technologies, Inc. (NASA ARC), El Segundo, CA 90245, USA
bhaskar.saha@nasa.gov

²NASA Langley Research Center, Hampton, VA 23681, USA
cuong.c.quach@nasa.gov

³NASA Ames Research Center, Moffett Field, CA 94035, USA
kai.goebel@nasa.gov

ABSTRACT

Battery Health Management (BHM) is a core enabling technology for the success and widespread adoption of the emerging electric vehicles of today. Although battery chemistries have been studied in detail in literature, an accurate run-time battery life prediction algorithm has eluded us. Current reliability-based techniques are insufficient to manage the use of such batteries when they are an active power source with frequently varying loads in uncertain environments. The amount of usable charge of a battery for a given discharge profile is not only dependent on the starting state-of-charge (SOC), but also other factors like battery health and the discharge or load profile imposed. This paper presents a Particle Filter (PF) based BHM framework with plug-and-play modules for battery models and uncertainty management. The batteries are modeled at three different levels of granularity with associated uncertainty distributions, encoding the basic electrochemical processes of a Lithium-polymer battery. The effects of different choices in the model design space are explored in the context of prediction performance in an electric unmanned aerial vehicle (UAV) application with emulated flight profiles.

1. INTRODUCTION

Battery-powered devices have become ubiquitous in the modern world, from tiny headsets to cameras, cell phones and laptops to hybrid and electric vehicles. Yet the battery is not a new invention. Battery artifacts date back to the early centuries A.D. (the Baghdad battery) and electric cars were favored over their gasoline counterparts in the late nineteenth century because of higher reliability. However, the uncertainty in determining battery life plagued electric vehicles then as it does now. A recent report by the Consumer Electronics Association, “Electric Vehicles: The Future of Driving”, indicates that although these vehicles are increasing in popularity, running out of battery power on the road is the top concern for consumers (71% of adults surveyed). Consequences of battery exhaustion may range from reduced performance to operational impairment and even to catastrophic failures, thus motivating the study of Battery Health Management (BHM).

One of the most critical applications of BHM technologies is in the field of electric vehicles (EVs). Usually combustion based powertrains run within narrow bands of RPMs (revolutions per minute) with metered fuel delivery. This combined with a known volume fuel tank allows reasonably accurate predictions of remaining use-time or travel distance. Batteries on the other hand, decrease in capacity with time and usage. Various factors like ambient storage temperatures and the state-of-charge (SOC) at which the battery was stored affects capacity fade. Additionally, the amount of usable charge of a battery for a given discharge profile is not only dependent on the starting SOC, but also other factors like battery health and the discharge or load profile imposed.

In this paper, the BHM problem is approached from the model-based point of view. The following sections will address the salient battery characteristics that need to be modeled, the BHM framework, explorations of the model design space, an electric unmanned aerial vehicle (UAV) application example, battery end-of-discharge (EOD) prediction results, and relevant conclusions.

2. BATTERY CHARACTERISTICS

Batteries are essentially energy storage devices that facilitate the conversion, or transduction, of chemical energy into electrical energy, and vice versa (Huggins,
2008). They consist of a pair of electrodes (anode and cathode) immersed in an electrolyte and sometimes separated by a separator. The chemical driving force across the cell is due to the difference in the chemical potentials of its two electrodes, which is determined by the difference between the standard Gibbs free energies the products of the reaction and the reactants. The theoretical open circuit voltage, $E^0$, of a battery is measured when all reactants are at 25°C and at 1M concentration or 1 atm pressure. However, this voltage is not available during use. This is due to the various passive components inside like the electrolyte, the separator, terminal leads, etc. The voltage drop due to these factors can be mainly categorized as follows.

**Ohmic Drop**

This refers to the diffusion process through which Li-ions migrate to the cathode via the electrolytic medium. The internal resistance to this ionic diffusion process is also referred to elsewhere as the IR drop. For a given load current this drop usually decreases with time due to the increase in internal temperature that results in increased ion mobility, and is henceforth referred to as $\Delta E_{IR}$.

**Activation Polarization**

Self-discharge is caused by the residual ionic and electronic flow through a cell even when there is no external current being drawn. The resulting drop in voltage has been modeled to represent the activation polarization of the battery, referred to from now on as $\Delta E_{AP}$. All chemical reactions have a certain activation barrier that must be overcome in order to proceed and the energy needed to overcome this barrier leads to the activation polarization voltage drop. The dynamics of this process is described by the Butler–Volmer equation. This process was represented by an exponential function in Saha and Goebel (2009). However, a log function is a more accurate representation, as abstracted from the Butler–Volmer equation.

**Concentration Polarization**

This process represents the voltage loss due to spatial variations in reactant concentration at the electrodes. This is mainly caused when the reactants are consumed by the electrochemical reaction faster than they can diffuse into the porous electrode, as well as due to variations in bulk flow composition. The consumption of Li-ions causes a drop in their concentration along the cell, between the electrodes, which causes a drop in the local potential near the cathode. This voltage loss is also referred to as concentration polarization, represented in this paper by the term $\Delta E_{CP}$. The value of this factor is low during the initial part of the discharge cycle and grows rapidly towards the end of the discharge or when the load current increases.

Figure 1 depicts the typical polarization curve of a battery with the contributions of all three of the above factors shown as a function of the current drawn from the cell. Since, these factors are current-dependent, i.e., they come into play only when some current is drawn from the battery, the voltage drop caused by them usually increases with increasing output current.

Since the output current plays such a big role in determining the losses inside a battery, it is an important parameter to consider when comparing battery performance. The term most often used to indicate the rate at which a battery is discharged is the C-Rate (Huggins, 2008). The discharge rate of a battery is expressed as $C/r$, where $r$ is the number of hours required to completely discharge its nominal capacity. So, a 2 Ah battery discharging at a rate of $C/10$ or $0.2$ A would last for 10 hours. The terminal voltage of a battery, as well as the charge delivered, can vary appreciably with changes in the C-Rate. Furthermore, the amount of energy supplied, related to the area under the discharge curve, is also strongly C-Rate dependent. Figure 2 shows the typical discharge of a battery and its variation with C-Rate. Each curve corresponds to a different C-Rate or $C/r$ value (the lower the $r$ the higher the current) and assumes constant temperature conditions.
3. Health Management Framework

Before investigating the issues with modeling the battery, this section takes a look at how the BHM framework is implemented using Particle Filters. The framework has been described before (Saha et al., 2009), however, some basic elements are reproduced below in order to set the context.

3.1 Particle Filter

The Particle Filter (PF) framework (Gordon et al., 1993) assumes that the state equations can be modeled as a first order Markov process with additive noise and conditionally independent outputs. Under these assumptions the state equations can be represented as:

\[ x_k = f_{k-1}(x_{k-1}) + w_{k-1} \]

and

\[ z_k = h_k(x_k) + v_k. \]

The filter approximates the posterior probability distribution denoted as \( p(x_k|Z_k) \), where \( Z_k = \{z_1, z_2, \ldots, z_k\} \) is the set of all measurements until \( t_k \), by a set of \( N \) weighted particles \( \{ (x^i_k, w^i_k) ; i = 1, \ldots, N \} \), such that \( \sum_i w^i_k = 1 \), and the posterior distribution can be approximated as:

\[ p(x_k|Z_k) \approx \sum_{i=1}^N w^i_k \delta(x_k - x^i_k). \]

Using the model in Eq. (1) the prior distribution going from \( t_{k-1} \) to \( t_k \) becomes:

\[ p(x_k|Z_{k-1}) \approx \sum_{i=1}^N w^i_{k-1} f_{k-1}(x^i_{k-1}). \]

The weights are updated according to the relation:

\[ \bar{w}^i_k = w^i_{k-1} \frac{p(z_k|x^i_k) f_{k-1}(x^i_{k-1})}{q(x^i_k|x^i_{k-1}, z_k)}, \]

where

\[ w^i_k = \frac{\bar{w}^i_k}{\sum_{j=1}^N \bar{w}^j_k}. \]

Resampling is used to avoid the problem of degeneracy of the PF algorithm, i.e., avoiding the situation that all but a few of the importance weights are close to zero. If the weights degenerate, we not only have a very poor representation of the system state, but we also spend valuable computing resources on unimportant calculations. More details on this are provided in Saha et al. (2009). The basic logical flowchart is shown in Figure 3.

During prognosis this tracking routine is run until a long-term prediction is required, say at time \( t_p \), at which point Eq. (4) will be used to propagate the posterior pdf (probability density function) given by \( \{ (x^i_k, w^i_k) ; i = 1, \ldots, N \} \) until \( x^i \) fails to meet the system specifications at time \( t_{EOL} \). The remaining useful life (RUL) pdf, i.e., the distribution of \( p(t_{EOL} - t_p) \), is given by the distribution of \( w^i_k \). Figure 4 shows the flow diagram of the prediction process.

3.2 Model Adaptation

One of the key motivating factors for using Particle Filters for prognostics is the ability to include model parameters as part of the state vector to be estimated. This performs model adaptation in conjunction with state tracking, and thus, produces a tuned model that can be used for long term predictions.

Assume that the system health state is 1-dimensional, given by \( x_t \), and the state evolution model \( f \) and the measurement model \( h \) are stationary in nature with known noise distributions \( \omega \) and \( v \) respectively. Additionally, we also assume that the parameter values of \( h \) are known. This assumption can be relaxed in a more generic approach.
Indeed, considering a non-stationary measurement model can be used to account for progressive degradation in sensors caused by corrosion, fatigue, wear, etc. The parameters of $f$, denoted by $\omega = \{a_j; j = 1, \ldots, n_t\}$, $n_t \in \mathbb{N}$, are combined with $x_k$ to give the state vector $x_k = [x_k \omega_k]^T$, where $T$ represents the transpose of a vector or matrix. Equations (1) and (2) can then be rewritten as:

$$x_k = f(x_{k-1}, a_{k-1}) + \omega_{k-1}$$

(7)

$$z_k = h(x_k) + v_k.$$  

(8)

The issue now is to formulate the state equations for $a_\omega$. One easy solution is to pick a Gaussian random walk such that:

$$a_{j,k} = a_{j,k-1} + \omega_{j,k-1}$$

(9)

where $\omega_{j,k-1}$ is drawn from a normal distribution, $\mathcal{N}(0, \sigma_j^2)$, with zero mean and variance $\sigma_j^2$. Given a suitable starting point $a_{j,0}$, and variance $\sigma_j^2$, the PF estimate will converge to the actual parameter value $a_j$, according to the law of large numbers.

It is not necessary to include all model parameters as part of the state to be estimated. In fact, the smaller the subset of parameters to be estimated, the faster the convergence since the state dimensionality is lower (Daum, 2003). However, this leads to the notion that the higher the model fidelity with respect to the real system, the lesser the number of parameters that need to be identified at run-time leading to better convergence properties.

4. MODEL DESIGN SPACE

The issue of modeling is paramount in any model-based algorithm like the PF. There can be many approaches to modeling, and for well studied systems like batteries the model design space is very large. There are several models that exist in literature at various levels of granularity and abstraction, like Gao, Liu, and Dougal (2002), Hartmann II (2008), Santhanagopalan, Zhang, Kumaresan, and White (2008), etc. Building these models require significant expenses in time and expertise. However, there are still issues with applicability in the field, since complex models need identification of several parameters, which might be impractical. Sometimes the models may be too complex to be run in real time.

For the purposes of the electric UAV BHM, we explore the model design space at a high level of abstraction of the underlying physics. It is desired to model the SOC of the battery in order to predict the EOD event as discussed below. In the results section the prediction performance for the different model choices are presented.

4.1 Model 1

For the empirical charge depletion model considered here, we express the output voltage $E(t_k)$ of the cell in terms of the effects of the changes in the internal parameters, as shown below:

$$E(t_k) = E^o - \Delta E_{IR}(t_k) - \Delta E_{AP}(t_k) - \Delta E_{CP}(t_k)$$

where $E^o$ is the Gibb’s free energy of the cell, $\Delta E_{IR}$ is the Ohmic drop, $\Delta E_{AP}$ is the drop due to activation polarizaton and $\Delta E_{CP}$ denotes the voltage drop due to concentration polarization. These individual effects are modeled as:

$$\Delta E_{IR}(t_k) = \Delta I_k R - \alpha_{1,k} t_k,$$

(11)

$$\Delta E_{AP}(t_k) = \alpha_{2,k} \exp(-\alpha_{3,k}/t_k),$$

(12)

$$\Delta E_{CP}(t_k) = \alpha_{4,k} \exp(\alpha_{5,k} t_k).$$

(13)

where $\Delta I_k$ is the change in current that flows through the internal resistance $R$ of the cell, and $a_{\omega} = \{a_j; j = 1, \ldots, 5\}$ represents the set of model parameters to be estimated.

The problem is to predict the time instant $t_{EOD}$ when the state $x$ denoting the cell voltage $E$ reaches the threshold level of 2.7 V. The PF representation of this problem is given by:

$$x_k = x_{k-1} - \left\{-a_{1,k-1} + \alpha_{2,k-1} \alpha_{3,k-1} \exp(-\alpha_{3,k-1}/t_{k-1}) \right\} t_{k-1} - \alpha_{4,k-1} \alpha_{5,k-1} \exp(\alpha_{5,k-1} t_{k-1}) (t_k - t_{k-1}) - \Delta I_k R + \omega_{k-1},$$

$$a_{1,k} = a_{1,k-1} + \omega_{1,k-1}$$

(14)

$$a_{2,k} = a_{2,k-1} + \omega_{2,k-1}$$

$$a_{3,k} = a_{3,k-1} + \omega_{3,k-1}$$

$$a_{4,k} = a_{4,k-1} + \omega_{4,k-1}$$

$$a_{5,k} = a_{5,k-1} + \omega_{5,k-1}$$

$$z_k = x_k + v_k.$$  

(15)

This is a 6 dimensional state vector with 1 dimension being the system health indicator (cell voltage) and the other dimensions coming from the model parameters. The term $\Delta I_k$ is the change in the load current at the time instant $t_k$.

4.2 Model 2

The model represented by Eqs. (14) – (15) does not represent the activation polarization process well. This is because the structure of the Butler Volmer equation is better
approximated by a log function rather than a negative exponential. Hence for Model 2, we change Eq. (12) to the following:

$$\Delta E_{AP}(t_k) = \alpha_{2,k} \ln(1 + \alpha_{3,k} t_k). \quad (16)$$

Correspondingly, Eq. (14) changes to:

$$x_k = x_{k-1} - \left\{ -\alpha_{4,k-1} \alpha_{5,k-1} \exp \left( \alpha_{5,k-1} t_{k-1} \right) (t_k - t_{k-1}) - \Delta I_k R + \omega_{h,k-1}, \quad (17) \right.$$

$$\alpha_{1,k} = \alpha_{1,k-1} + \omega_{h,k-1}
\alpha_{2,k} = \alpha_{2,k-1} + \omega_{h,k-1} 
\alpha_{3,k} = \alpha_{3,k-1} + \omega_{h,k-1} 
\alpha_{4,k} = \alpha_{4,k-1} + \omega_{h,k-1} 
\alpha_{5,k} = \alpha_{5,k-1} + \omega_{h,k-1}$$

The state vector is similar here as in Model 1. The level of granularity, indicating the different physical processes modeled, is the same although the abstraction of one of the processes has changed.

4.3 Model 3

It should be noted that for most batteries, the voltage as well as the charge delivered varies considerably with changes in $I$. This can be better represented by making two changes to the battery model described so far. Firstly, the parameters of the model must be load dependent. We model this by making $\alpha_3$ and $\alpha_4$ proportional to the load current $I$. Secondly, when we have step changes in the load, a higher load level followed by a lower one presents a period of relaxation for the battery. During this period the voltage does not immediately jump up but gradually rises which can be modeled by an exponential function. A similar effect can also be observed for a step increase in current level. These effects can be reconciled by considering the battery impedance as an RC equivalent circuit (Zhang, 2010). We can thus replace Eq. (11) by:

$$\Delta E_{IRC}(t_k) = \Delta I_k R \alpha_6 (1 - \exp(-\alpha_7 (t_k - t_{\Delta I})) - \alpha_8 t_k) \quad (18)$$

where $\Delta I_k$ is the step change in current at time $t_{\Delta I_k}$. The other processes are represented as:

$$\Delta E_{AP}(t_k) = \alpha_{2,k} \ln(1 + \alpha_{3,k} I_k t_k), \quad (21)$$

$$\Delta E_{CP}(t_k) = \alpha_{4,k} \exp(\alpha_{5,k} I_k t_k). \quad (22)$$

The filter equations can be derived out as before and are shown in Saha et al. (2011). Model 3 represents a higher level of granularity in the model design space since some additional battery behavior to changes in load is being taken into effect. This leads to higher accuracy in the model output as well as a corresponding increase in the number of parameters. To maintain a tolerable rate of convergence, all but the parameters $\alpha_3$ and $\alpha_4$ are learnt from training data, while $\alpha_3$ and $\alpha_4$ are estimated by the PF online.

5. Application Example

The test UAV platform for this research is a COTS 33% scale model of the Zivko Edge 540T. Details of this platform have been presented in Saha et al. (2011), but are also repeated here for the sake of readability. The UAV is powered by dual tandem mounted electric out-runner motors capable of moving the aircraft up to 85 knots using a 26 inch propeller. The gas engine in the original kit specification was replaced by two electric out runner motors which are mounted in tandem to power a single drive shaft. The motors are powered by a set of 4 Li-Poly rechargeable batteries. The batteries are each rated at 6000 mAh. The tandem motors are each controlled by separate motor controllers.

Testing on the Edge 540 UAV platform was carried out with the airframe restrained on the ground. The propeller was run through various RPM (revolutions per minute) regimes indicative of the intended flight profile (takeoff, climb, multiple cruise, turn and glide segments, descent and landing). Figure 5 shows the voltages during a typical flight. It is desired to predict when the battery will run out of charge, i.e., when the EOD event indicated by the end of the voltage plots after landing will occur.

![Figure 5. Battery voltages during a typical flight](image-url)
6. RESULTS

In order to evaluate the prognostic algorithm we make 7 predictions spaced 1 minute apart starting from 800 secs into the flight. It is not desired to make predictions till the end of the flight since there needs to be some time for the UAV pilot to land the aircraft with some safety margin on the remaining battery life. Figures 6 – 8 show sample predictions generated by the Models 1 – 3 respectively. The time instants when the predictions are made are shown in green vertical dashed lines, with lighter shades indicating earlier predictions. The corresponding EOD pdfs are shown in green patches on the 17.4 V EOD threshold voltage line (dashed gray). The pdfs themselves are given by the distribution \( \{ \left( t_{i,EO} - t_p, w_i \right); \ i = 1, \ldots, N \} \), where \( i \) is the particle index and \( t_{i,EO} \) is the predicted time where the \( i \)th particle trajectory crosses the EOD threshold. The real voltages are shown as red ×s, while the PF estimates are shown as blue dots. The large spread of the blue dots is caused by the bias errors and noise in the Hall effect current sensors used. Since this uncertainty has not been expressly modeled, the actual EOD can sometimes lie outside the predicted pdf as shown in Figures 6 – 8.

For statistical validation, we ran each model 100 times over the same data to generate the \( \alpha \)-\( \lambda \) metric plots as defined in Saxena et al. (2008). This prognostic metric measures whether RUL predictions continue improve in accuracy with time as more run-time data is made available, where \( t_{i,RUL} = t_{i,EO} - t_p \). It also enforces the notion that the prediction error needs to reduce as the prediction time instant approaches the end of life (EOD in this case) since there is less time to take corrective action. For these experiments, the \( \alpha \) value is chosen to be 0.1 and \( \lambda \) is chosen to be 0.5 (i.e. it is desired that the prediction trajectories be within 90% accuracy with 50% battery life left). Figures 9 – 11 show the \( \alpha \)-\( \lambda \) plots for Models 1 – 3 respectively for \( t_p = \{ 800, 860, 920, 980 \} \). Model 1 shows the worst performance, while Model 3 is the best as was expected from the model choices. The worsening performance of both Models 1 and 2 toward the end predictions is most likely due to the inability of these models to adapt to the low load glide modes as shown in Figure 5.
7. CONCLUSION

In summary, this paper investigates the battery life prediction performance that result from different choice points in the model design space. This is meant as a first step in formalizing the effect of model choices with the goal of ultimately parametrizing the model design space to analyze the tradeoffs involved. Higher granularity and lower levels of abstraction might generally give more accurate models, but that also results in larger parameter sets which may not have good convergence properties if included in the state vector. To mange such models, we would need to estimate most of the parameters from training data and choose only a few for online adaptation. This predicates a higher model development cost and computational complexity. A more formal analysis of these concepts will be presented in future publications.

ACKNOWLEDGEMENT

This work was a performed as a cross-center collaboration between NASA Ames and Langley Research Centers (ARC and LaRC) and Dryden Flight Research Center (DFRC). The authors would like to especially thank Edwin Koshimoto at DFRC, and Sixto L. Vazquez, Edward F. Hogge, Thomas H. Strom and Boyd L. Hill at LaRC for their contributions. The funding for this work was provided by the NASA Integrated Vehicle Health Management (IVHM) project under the Aviation Safety Program of the Aeronautics Research Mission Directorate (ARMD).

NOMENCLATURE

\( E \) = battery voltage
\( \Delta E \) = voltage drop
\( E^o \) = theoretical output voltage
\( x \) = state variable
\( y \) = measurement
\( t \) = time
\( \Delta t \) = time delay between consecutive time steps
\( \Delta I \) = change in load between consecutive time steps
\( \alpha \) = model parameter

REFERENCES


**Bhaskar Saha** received his Ph.D. from the School of Electrical and Computer Engineering at Georgia Institute of Technology, Atlanta, GA, USA in 2008. He received his M.S. also from the same school and his B. Tech. (Bachelor of Technology) degree from the Department of Electrical Engineering, Indian Institute of Technology, Kharagpur, India. He is currently a Research Scientist with Mission Critical Technologies at the Prognostics Center of Excellence, NASA Ames Research Center. His research is focused on applying various classification, regression and state estimation techniques for predicting remaining useful life of systems and their components, as well as developing hardware-in-the-loop testbeds and prognostic metrics to evaluate their performance. He has been an IEEE member since 2008 and has published several papers on these topics.

**Cuong C. Quach** got his M.S. from the School of Physics and Computer Sciences at Christopher Newport University in 1997. He is a staff researcher in the Safety Critical Avionics Systems Branch at NASA Langley Research Center. His research areas include development and testing of software for airframe diagnosis and strategic flight path conflict detection.

**Kai Goebel** received the degree of Diplom-Ingenieur from the Technische Universität München, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively. Dr. Goebel is a senior scientist at NASA Ames Research Center where he leads the Diagnostics & Prognostics groups in the Intelligent Systems division. In addition, he directs the Prognostics Center of Excellence and he is the Associate Principal Investigator for Prognostics of NASA’s Integrated Vehicle Health Management Program. He worked at General Electric’s Corporate Research Center in Niskayuna, NY from 1997 to 2006 as a senior research scientist. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He holds eleven patents and has published more than 100 papers in the area of systems health management.