A Simulation Engine for Predicting State-of-Charge and State-of-Health in Lithium-Ion Battery Packs of Electric Vehicles

Pablo A. Espinoza [1], Aramis Pérez [2], Marcos E. Orchard [3], Hugo F. Navarrete [4], and Daniel A. Pola [5]

[1,2,3,4,5] *Department of Electrical Engineering, University of Chile, Santiago, 8370451, Chile*
pablo.espinoza@ing.uchile.cl
aramis.perez@ing.uchile.cl
morchard@ing.uchile.cl
hnavarre@ing.uchile.cl
dpola@ing.uchile.cl

ABSTRACT

Recent developments in lithium-ion technology have enabled a revolution in the automotive industry. Fully electric vehicles (EVs) operate under distinctly variable conditions, requiring high-voltage battery packs to meet their torque/power demands. Our goal is to provide a simulation engine which, for a given battery pack size, determines when recharging or battery pack replacement are needed. To that end, we study both the State-of-Charge (SOC) and the State-of-Health (SOH) indicators, using discrete state space models for both. Predictions are based on a probabilistic characterization of EV usage profiles, which in turn are a function of generic user-input, such as mission maps, vehicle mechanical characteristics, driving schedules, and battery pack configuration. State space models benefit from the incorporation of meta-models for the ohmic internal resistance and the Coulomb efficiency of the pack. Both meta-models i) effectively introduce additional phenomenology –such as dependency on the magnitude of discharged current and depth of discharge (DoD)–, and ii) provide a link between SOC/SOH and how each discharge cycle affects the health status of the battery pack as a whole. The approach for the simulation engine presented here is stochastic in nature, meaning that prognostics for the SOC and SOH are generated in a particle filter-based scheme. Thus risk and confidence intervals can be obtained for the end-of-discharge and end-of-life respectively

1. INTRODUCTION

Today, the enhance of greenhouse gases in the atmosphere is one of the main contributors to global warming (Bose, 2010). This scenario has forced us to look for alternative and environmentally friendly power sources. An example of this new paradigm comes from the electrification of transportation. Although internal combustion engines are still the standard, in the last decade several versions of hybrid and electric vehicles (EVs) have become a real alternative to the public worldwide (Macharis, Lebeau, Mierlo, & Lebeau, 2013). Companies like Nissan, Renault, and Tesla, often in cooperation with their respective national governments, have developed long range and clean transportation accessible to wide segments of the market (Wansart & Schneider, 2010).

However, there are still significant challenges to meet. A basic example relates to *range anxiety* (Yang et al., 2015), i.e. the fear that a vehicle has insufficient energy to reach its destination, stranding its passengers. Several strategies are currently being used to alleviate it, such as the deployment of extensive charging infrastructure, precise navigation, and accurate range prediction. But the issue of being unable to assign risk or a confidence interval to *when* the energy storage system (ESS) will run out of power, persists.

Another important ESS challenge involves the design and sizing of the battery pack. Price, together with considerations on volume and weight, argues for a limited size or number of modules (Wansart & Schneider, 2010). Even more importantly, the future *usage profile* of an EV is an unknown piece of information, and thus the manufacturer needs to plan for different contingencies and always aim for robustness. Under normal usage, the battery pack of the Nissan Leaf is expected to have a lifespan of approximately ten years (Zeff, 2016). However, *normal usage* may require rather specific conditions, both environmental and driver-dependent, that are not well-defined. What happens if the EV is driven aggressively, or charged erratically?, Or even operated consistently under hostile weather? We are interested in determining when a replacement will be required, or equivalently, to predict battery pack remaining life as a function of future usage profile.

Dealing with these problems means we need to model, esti-

mate, and make predictions on both the SOC and the SOH (Xiong, He, Sun, & Zhao, 2013; Pattipati, Sankavaram, & Pattipati, 2011; Han, Xu, Yuan, & Shen, 2014). Various approaches to study them have been employed, such as neural networks, fuzzy logic, or Bayesian methods. Within these, models ranging from electrochemical to circuit-based and empirical have been used (Charkhgard & Farrokhi, 2010; Salkind, Fennie, Singh, Atwater, & Reisner, 1999; Saha, Goebel, Poll, & Christophersen, 2009; Orchard, Tang, Saha, Goebel, & Vachtsevanos, 2010). In this work we employ Bayesian processors on empirical state-space models of the SOC and SOH. The family of sequential Monte Carlo (SMC) methods is effective when dealing with nonlinear phenomena, and as such, long-term predictions are accordingly generated within a particle filter (PF)-based scheme.

In this paper we build a simulation engine that delivers stochastic predictions for i) battery pack autonomy, and ii) battery pack remaining life. How aggressively the vehicle is driven, its mission, and the configuration of the battery pack are all generic inputs in our simulation. The stochastic nature of our formulation allows to obtain associated measures of risk and confidence intervals for the end-of-discharge and end-of-life. We include lithium-ion battery phenomenology (Urbain, Rael, Davat, & Desprez, 2008; Miles, 2001; Penna, Nascimento, & Rodrigues, 2012), and how operational conditions in each discharge cycle affect both the SOC/SOH indicators. We aim to quantify how these conditions significantly influence long-term evolution, especially for the SOH.

The article is organized as follows: Section 2 describes the main mathematical tools used in this work. Section 3 presents the basis of the electrical/mechanical modeling used to generate usage profiles. Section 4 deals with the battery pack autonomy problem through SOC prediction. The battery pack remaining life estimation is tackled in Section 5. In Section 6, we summarize our conclusions.

2. THEORETICAL FRAMEWORK

Estimation of SOC/SOH (Xiong et al., 2013; Pattipati et al., 2011) is important, but their future behavior is critical to battery autonomy and remaining life. End-of-life indicators allow us to take informed decisions on when we need a recharge or replacement (Han et al., 2014). These indicators, however, are not directly observable. We need to perform statistical inferences, based on variables we are capable to sense or evaluate. We begin this section by exploring Bayesian inference in the context of sequential Monte Carlo (SMC) methods. Then we introduce a PF-based prognosis scheme, and finish with a characterization of end-of-life indicators.

2.1. Bayesian inference

The evolution of dynamical systems can be described by state-space models, whether they contain differential or difference equations:

$$x_k = f_k(x_{k-1}, v_{k-1}) \quad (1)$$

$$z_k = h_k(x_k, n_k), \quad (2)$$

where Eqns. (1) and (2) correspond to the evolution of the state vector $\{x_k, k \in \mathbb{N}\}$ and measurement vector $\{z_k, k \in \mathbb{N}\}$. In this formulation, $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_x}$ and $h_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_n} \rightarrow \mathbb{R}^{n_z}$ are non-linear functions, with $\{v_{k-1}, k \in \mathbb{N}\}$ and $\{n_k, k \in \mathbb{N}\}$ corresponding to independent and identically distributed process and measurement noises, respectively.

The Bayesian filtering problem deals with obtaining information about the state x_k recursively, using the available measurements $z_{1:k}$ until time k. Thus we seek the probability density function (pdf) $p(x_k|z_{1:k})$, also known as the *posterior* distribution of the state vector x_k. Let us assume the pdf at time $k-1$, $p(x_{k-1}|z_{1:k-1})$, is available. Through the prediction and update steps of the filter we can obtain the *posterior* pdf at time k, by means of a new observation arriving at time k and Bayes theorem (Gregory, 2005):

$$\overbrace{p(x_k|z_{1:k})}^{Posterior} = \frac{\overbrace{p(z_k|x_k)}^{Likelihood}\overbrace{p(x_k|z_{1:k-1})}^{Prior}}{\underbrace{p(z_k|z_{1:k-1})}_{Evidence}}. \quad (3)$$

Eq. (3) corresponds to the theoretical solution for the optimal Bayesian estimation. However, an analytical form can not be computed generally. If the state space model is linear and affected by Gaussian noises, optimal analytical solutions (in the mean square error sense) can be obtained. These include the Kalman filter (Kalman, 1960), and grid-based filters (Arulampalam, Maskell, Gordon, & Clapp, 2002).

2.2. Particle filters (PF)

When models are non-linear and noises non-Gaussian, we resort to sub-optimal methods (Creal, 2012). Here the posterior pdf is represented by a set of random samples, with associated weights (particles). Let $\{x^i_{0:k}, w^i_k\}_{i=1}^{N_p}$ be a set of N_p particles that characterizes the posterior pdf $p(x_{0:k}|z_{1:k})$, where $\{x^i_{0:k}, i = 1, .., N_p\}$ is a set of support mass points, with associated weights $\{w^i_k, i = 1, .., N_p\}$, and $x_{0:k} = \{x_j, j = 0, .., k\}$ corresponds to all the states at time k. The weights are normalized, such that $\sum_{i=1}^{N_p} w^i_k = 1$. The posterior pdf at time k can then be approximated by:

$$p(x_{0:k}|z_{1:k}) \approx \sum_{i=1}^{N_p} w^i_k \delta(x_{0:k} - x^i_{0:k}). \quad (4)$$

The expression in Eq. (4) is an empirical distribution that

approximates the real posterior pdf, $p(x_{0:k}|z_{1:k})$, and which weights are chosen through the *importance sampling* principle (Doucet & Johansen, 2008).

As the number of particles $N_p \rightarrow \infty$, the approximation (see Eq. (4)) converges to the true posterior pdf $p(x_k|z_{1:k})$ (Doucet, Godsill, & Andrieu, 2000; Creal, 2012; Doucet & Johansen, 2008). In this way, the importance sampling method recursively propagates the samples set, along with its associated weights, whenever a new observation arrives sequentially. However, this algorithm is often vulnerable to the degeneracy phenomenon (Arulampalam et al., 2002). A solution to this problem is given by re-sampling. When re-sampling operates, a new set of samples $\{x_k^{i*}\}$ is generated, through N_p re-samples (with replacement), such that:

$$\mathbb{P}(x_k^{i*} = x_k^j) = w_k^j. \tag{5}$$

Although re-sampling reduces degeneracy, it also introduces negative effects (Doucet et al., 2000; Doucet & Johansen, 2008). The most significant particles, with large associated weights w_k^i, are sampled many times. This yield to a lack of diversity amongst particles, with a population that is composed largely by a repetition of a few of them.

2.3. Regularized particle filter (RPF)

A solution to the problem of lack of diversity is given by the regularized particle filter (Musso, Oudjane, & Gland, 2001). In the re-sampling stage of the sequential importance sampling algorithm, particles are sampled from a discrete approximation. For that reason, it is possible that over time this would lead to severe *particle collapse*, where all of them occupy the same position in state space. To avoid that, the regularized PF samples from a continuous distribution instead:

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N_p} w_k^i K_h(x_k - x_k^i), \tag{6}$$

where

$$K_h = \frac{1}{h^{n_x}} K\left(\frac{x}{h}\right) \tag{7}$$

is the rescaled kernel density $K(\cdot)$, $h > 0$ is the kernel bandwidth, and n_x is the dimension of the state vector. The kernel and bandwidth are chosen to minimize the MSE between the true posterior density and the regularized empirical representation in Eq. (6). In the special case of all particles having the same weight, the optimal choice for the kernel is the Epanechnikov kernel:

$$K(x) = \begin{cases} \frac{n_x+4}{2c_{n_x}}(1-|x|) & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{8}$$

where c_{n_x} is the volume of the unit sphere in $\mathbb{R}^{n_x}$. When the underlying density is Gaussian with a unit covariance matrix, the optimal choice for the bandwidth is:

$$h_{opt} = A \cdot N^{-\frac{1}{n_x+4}}, \tag{9}$$

$$A = \left(8c_{n_x}^{-1} \cdot (n_x+4) \cdot (2\sqrt{\pi})^{n_x}\right)^{\frac{1}{n_x+4}}. \tag{10}$$

2.4. PF-based prognosis scheme for dynamic systems

Prognosis schemes can be understood as long-term predictions on a system. Through state-space modeling, it is possible to describe an indicator's evolution through time. Let us consider the n-step prediction for the conditional pdf $\hat{p}(x_{k+\tau}^i|x_{k+\tau-1}^i)$. This describes the state evolution in the future time $k+\tau$, $(\tau = 1, .., n)$ when the particle $x_{k+\tau-1}^i$ is used as initial condition. Assuming current weights $\{w_t^i\}_{i=1}^{N_p}$ are a good representation of the state pdf at time k, then the state pdf at time $k+\tau$ can be approximated by the law of total probabilities and weights at time $k+\tau-1$, as shown in Eq. (11):

$$\hat{p}(x_{k+\tau}|\hat{x}_{1:k+\tau-1}) \approx \sum_{i=1}^{N_p} w_{k+\tau-1}^i \cdot \hat{p}(x_{k+\tau}^i|\hat{x}_{k+\tau-1}^i), \forall \tau \in \{1, .., n\}. \tag{11}$$

The weight of each particle must be modified in each prediction step τ to evaluate Eq. (11). Thus, an update of particle weights is necessary, but it can not depend on the arrival of new measurements. One approach to circumvent this issue, and that has proven particularly useful in large prediction horizons (Orchard & Vachtsevanos, 2009), is based on the regularized PF algorithm (Musso et al., 2001). Instead of updating particle weights in each prediction step, the uncertainty is represented by a re-sampling of the predicted state pdf (see Eq. (11)). In this way, the state pdf in future time steps is characterized by the position of particles, and not by their associated weights. Consider the discrete approximation (see Eq. (12)) for the predicted state pdf:

$$\hat{p}(x_{t+k|\hat{x}_{1:t+k-1}}) \approx \sum_{i=1}^{N_p} w_{t+k-1}^i \cdot K_h(x_{t+k} - \mathbb{E}\{x_{t+k}^i|\hat{x}_{t+k-1}^i\}), \tag{12}$$

where $K(\cdot)$ is a kernel density function (see the Epanechnikov kernel in Eq. (8)). This method proposes a computationally affordable solution based on the assumptions of uncorrelated process noise and the use of kernel transitions to describe the state PDF before the re-sampling step. To avoid a loss of diversity in the particle population we assume the state covariance matrix $\hat{S}_{k+\tau}$ is equal to the empirical covariance matrix $\hat{x}_{k+\tau}$, and that a set of particles with uniform weights $\hat{x}_{k+\tau-1}$ is available.

2.5. Statistical characterization of the End-of-Discharge and End-of-Life

The main objective of the prognosis algorithm is to characterize the End-of-Discharge (EOD) and End-of-Life (EOL) pdfs, for the SOC and SOH schemes respectively. The resulting predicted pdfs contains critical information about the evolution of the fault indicator over time. In the automotive context, the EOD pertains to when a recharge is needed, while the EOL relates to the concept of RUL, and when the battery pack is set to be replaced.

These probabilities depend on both long-term predictions and empirical knowledge about critical conditions for the system, usually given as thresholds known as *hazard zones* (Orchard, Tang, Goebel, & Vachtsevanos, 2009). Therefore, the probability of failure at any future time instant $k = \{eod, eol\}$ is given by the law of total probabilities:

$$\mathbb{P}(EOD = eod) = \sum_{i=1}^{N_p} \mathbb{P}(Failure|X = \hat{x}^i_{eod}) \cdot w^i_{eod} \quad (13)$$

$$\mathbb{P}(EOL = eol) = \sum_{i=1}^{N_p} \mathbb{P}(Failure|X = \hat{x}^i_{eol}) \cdot w^i_{eol}. \quad (14)$$

For SOC prognosis, critical conditions are met when the voltage measured at battery pack terminals falls below a certain threshold. In the case of SOH, failure prognosis occurs at the cycle when the battery pack capacity drops to a fraction of its nominal value (in this work we consider a 75% of the initial capacity).

One way to represent this information is through the computation of statistics, e.g. expectation values or 95% confidence intervals, for either the EOD or EOL pdfs. We also incorporate a measure of *risk* (Orchard, Tang, et al., 2009) given by the just-in-time point or $JITP_{\gamma\%}$. This metric specifies the time instant where the probability of failure reaches a specified threshold γ. Both expectation and $JITP_{\gamma\%}$ values for the EOD are given by the following expressions:

$$\widehat{EOD} = \mathbb{E}\{EOD\} = \mathbb{E}\{k \,|\, \mathbb{E}\{x(k)\} = f_{th}\}, \quad (15)$$

$$JITP_{\gamma\%} = \arg\min_{eod} \{\mathbb{P}(EOD \leq eod) \geq \gamma\%\}, \quad (16)$$

where γ takes values of $\{5\%, 15\%\}$ to capture the behavior of the pdf tail, and $f_{th} = 5\%$ of the energy remaining in the pack. Eqns. (15) and (16) are analogous for the EOL case, but now f_{th} corresponds to the 75% of the initial capacity of the battery pack.

3. SIMULATING USAGE PROFILES FOR ELECTRIC VEHICLES

We now explain the simulation procedure to obtain usage profiles as a function of generic inputs. First, we describe the battery pack (Section 3.1) and EV mechanical parameters (Section 3.2). This information coupled with the mission map (Section 3.3), i.e. the terrain characteristics, and the driving schedule (or how the vehicle will be driven, Section 3.4), allow us to obtain the power demand. Power demand (Section 3.5) combined with the battery pack electrical description allow us to reach our output: a usage profile or time-series of current consumption that the ESS must deliver to fulfill the given mission (Section 3.6). Figure 1 summarizes this process.

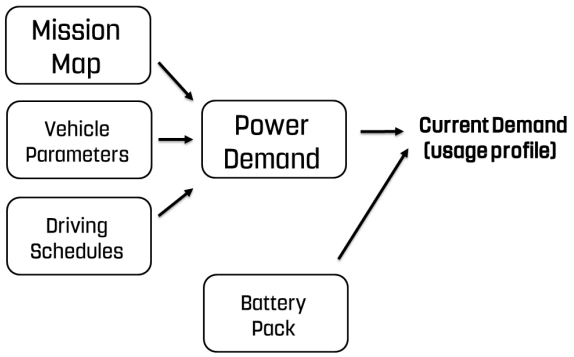


Figure 1. A progression scheme to simulating usage profiles for EVs.

We describe the simulation with inputs as realistic as possible. Along that line, we exemplify for the Nissan Leaf case. All the complexity and large amount of parameters required to represent the EV, the ESS, its mission, and driving schedules is reduced to a current demand, as explained below.

3.1. Battery pack: Electrical characterization

The schematics of a generic ESS consist on the aggregation of cells into modules, and modules into a pack. Detailed data on battery pack characteristics is not abundant. However, datasheets with rated (nominal) quantities are available, and in the case of the Nissan Leaf there is additional information coming from the cell manufacturer, Automotive Energy Supply Corporation (Ikezoe, Hirata, Amemiya, & Miyamoto, 2012) and from the U.S. Department of Energy and their Advanced Vehicles Testing Activity Program[1] (AVT).

The Leaf's 24 kWh battery pack consists of 48 modules and each module contains four battery cells, for a total of 192 cells, in a 96 series of two parallel strings configuration. Each module contains four lithium-ion battery cells. At a 66.2 Ah rated capacity, the ESS is expected to have a lifespan of ap-

[1] See https://avt.inl.gov/

Table 1. Battery pack characterization: input example for the 2015 Nissan Leaf. Spatial dimensions and weights are given for completeness, but don't affect our modeling.

Pack	**from AVT**	**Description**
v_{max}	403.2	Maximum Rated Voltage (V)
v_{nom}	360	Nominal Rated Voltage (V)
E_{nom}	24	Rated Energy (kWh)
C_{nom}	66.2	Rated Capacity (Ah)
P_{out}	> 90	Power Output (kW)
ρ_E	140	Energy density (Wh/kg)
l	1.57	Length (m)
w	1.19	Width (m)
h	0.26	Height (m)
M	294	Weight (kg)
Module	**(from AESC)**	**Description**
l	0.30	Length (m)
w	0.22	Width (m)
h	0.04	Height (m)
M	3.8	Weight (kg)
Cell	**(from AESC)**	**Description**
v_{max}	4.2	Maximum Rated Voltage (V)
v_{nom}	3.8	Nominal Rated Voltage (V)
C_{nom}	33.1	Rated Capacity (Ah)
ρ_E	157	Energy density (Wh/kg)
l	0.29	Length (m)
w	0.22	Width (m)
h	0.007	Height (m)
M	0.8	Weight (kg)

proximately 8 years (i.e. to retain 70% of its capacity for a decade) (Zeff, 2016). For any given cycle, it can be charged from total discharge to 80% capacity in about 30 minutes using a DC fast charger.

We list the input data necessary for battery pack characterization in Table 1, which sums up the values obtained from AESC and the AVT. In addition to this, it is important to remark that for a detailed modeling we also need the *functional forms* (curves or lookup tables) for discharges of both individual cells and the battery pack.

3.2. Vehicle mechanical parameters

To obtain the mechanical power required to complete a mission, we need to characterize dynamics. And to characterize dynamics, our simulation requires the mechanical parameters of the considered EV. These parameters are shown in the left column of Table 2, and can be specified for any vehicle the user chooses.

The parameters in the first part of the table account for the EV motion through the air flow, and therefore will be necessary to model dissipative forces acting against the direction of motion (aerodynamic and drag). The second part of the table mostly accounts for structural and engine parameters that become relevant for cruise control schemes. Additional environmental parameters, which we have considered fixed in our simulation are listed in Table 3:

Table 2. Mechanical parameters: input example for 2015 Nissan Leaf.

Parameter	**Value**	**Description**
m	1525	EV+ driver mass (kg)
C_d	0.29	Drag coefficient (dimensionless)
A	2.27	Surface frontal area (m^2)
c_{r1}	0.01	Static drag coeff. (dimensionless)
c_{r2}	1.789e-4	Dynamic drag coefficient (s/m)
n_w	4	Number of wheels (dimensionless)
I_w	0.0001	Wheel moment of inertia ($kg\ m^2$)
r_w	0.4064	Radius of a wheel (m)
m_{eq}	$m + \frac{n_w I_w}{r_w^2}$	Equivalent mass (kg)
T_m	280	Maximum torque ($N\ m$)
w_m	285.88	Maximum engine speed (rad/s)

Table 3. Environmental parameters (fixed).

Parameter	**Value**	**Description**
u_w	0	wind velocity (m/s)
ρ_a	1.204	air density (kg/m^3)
g	9.81	gravitational acceleration (m/s^2)

3.3. Mission Map: Terrain characteristics

The next input of our simulation is the trajectory. We parameterize it by two quantities: distance, and elevation profile. The first refers to the cumulative length traveled in a given discharge cycle, while the elevation profile describes the slopes the vehicle must go through. The elevation changes are the major contributor to the ESS energy budget when compared to the load of lateral accelerations. Our simulation accepts arbitrary entries for the vehicle's trajectory, and it requires the user only to define an starting and finishing point in Google Maps. The slope angle -that describes the elevation changes through the trajectory- is: $\theta(r) = \frac{dh}{dr}|_r$. The elevation profile of the route is defined as $h(r)$.

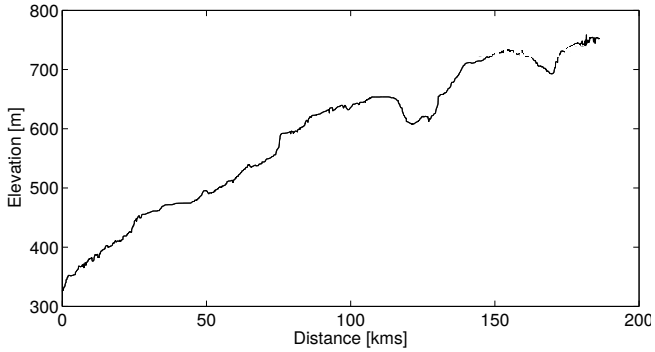


Figure 2. Elevation profile $h(r)$ in meters of the route between Tucson and Phoenix, AZ, North-South direction.

Figure 2 shows an example for this module based on the route between two cities in Arizona, USA: Tucson and Phoenix.

The right panel shows the elevation profile $h(r)$ as a function of traveled distance (in the North-South direction). Over 191 kilometers -which results in a larger energy requirement than a full discharge of a new ESS in the Nissan Leaf- the height change is almost 400 meters. This results in somewhat smooth slope changes, with maximum local variations of $\frac{dh}{dr}|_r \sim \pm 10^\circ$.

3.4. Driving schedules

Now the user of the simulation engine needs to specify how the vehicle is driven. The U.S. Environmental Protection Agency data provide public chassis dynamometer driving schedules[2] that we will use as strongly differentiated driving styles in our simulation (Kim, Rousseau, & Rask, 2016). A summary of the available options for user input can be found in Table 4 and Figure 3.

Table 4. EPA vehicle chassis dynamometer driving schedules.

Code	Description
US06	High acceleration, aggressive driving
HWFET	Highway conditions under 60 mph
NYCC	Low speed stop-and-go traffic
UDDS	Represents city driving conditions
FTP	UDDS and then its first 505 s again
HDUDDS	UDSS for heavy duty vehicle testing

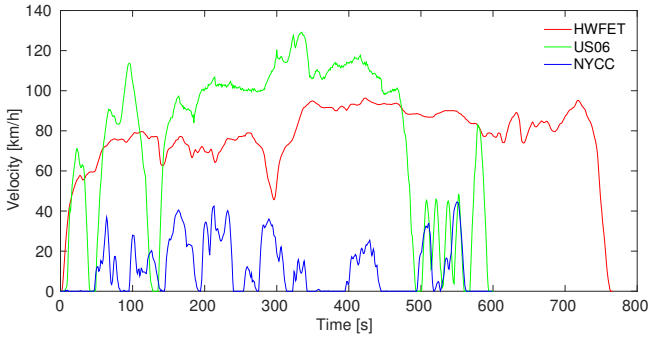
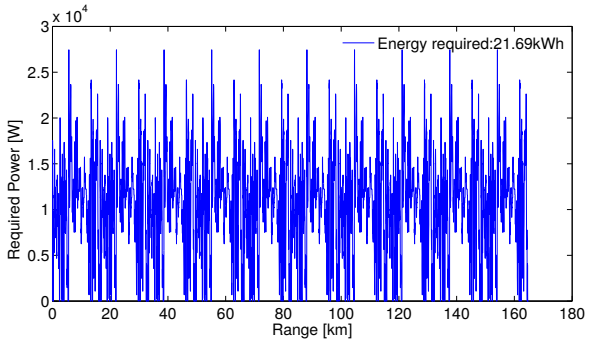


Figure 3. Examples of different EPA vehicle dynamometer driving schedules: Red, green and blue lines represent the HWFET, US06, and NYCC driving styles respectively.

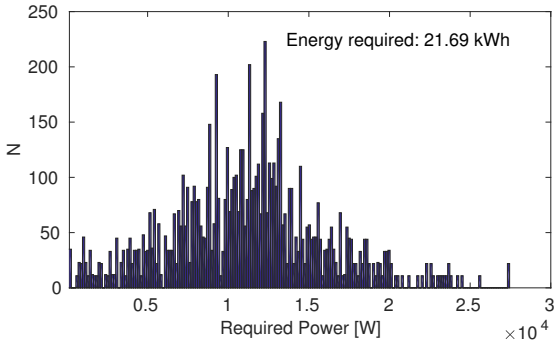
3.5. Power demand for the EV mission

We can now bring together most of the modules described in this section. The EV mechanical parameters are the ones listed in Table 2, the environmental variables are those from Table 3, the mission map will have a constant elevation profile, and we will use the HWFET schedule as our velocity profile $u(t)$. Under these conditions, the distance traveled by the EV is 164.92 km, with a mission time of 7,659 s and mechanical energetic requirement of 21.69 kWh. Since we use a constant $\alpha_{eff} = 0.9$ for the conversion of mechanical into

[2]See https://www.epa.gov/vehicle-and-fuel-emissions-testing/



(a)



(b)

Figure 4. (a) Required Power P_{req} to keep the EV at HWFET driving schedule speeds. (b) Histogram of P_{req} (see text for details).

electrical energy, the mission requires approximately all the capacity of the battery pack, i.e. the rated 24 kWh.

For physical modeling (Young, Wang, Wang, & Strunz, 2013) we include forces such as gravity (f_g), aerodynamic (f_{aero}) and drag (both static f_{r-sta}, and dynamical f_{r-dyn}):

$$f_g = m \cdot g \cdot sin(\theta) \tag{17}$$

$$f_{aero} = (1/2) \cdot C_d \cdot \rho_a \cdot A \cdot (u(t) + u_w)^2 \tag{18}$$

$$f_{r-sta} = m \cdot g \cdot cos(\theta) \cdot c_{r1} \tag{19}$$

$$f_{r-dyn} = m \cdot g \cdot cos(\theta) \cdot c_{r2} \cdot u(t) \tag{20}$$

with all the symbols previously defined. Being non-conservative forces, we have the equilibrium equation:

$$dK(t)/dt = P_{req}(t) - P_{dis}(t), \tag{21}$$

where K is the kinetic energy, P_{req} is the required power the engine shall exert to keep the EV moving at $u(t)$, and P_{dis} is the power dissipated by non-conservative forces.

Figure 4 shows the required mechanical power the battery pack and AC motor combination must provide to keep the EV on the HWFET driving schedule. The Figure also displays a histogram for the P_{req}, where we can observe that the mean consumption falls around 12000 W. This value is consistent with highway measurements done with the Nissan

Leaf (Wishart, Carlson, Chambon, & Gray, 2013), and thus validate our simple dynamical modeling.

3.6. EV current demand: A deterministic usage profile

Finally we obtain the current demand during a given mission. This is a time series we call *usage profile*. To do this we begin by building an approximation for both the open-circuit voltage (OCV) (Snihir, Rey, Verbitskiy, Belfadhel-Ayeb, & Notten, 2006; Weng, Sun, & Peng, 2013) and the internal ohmic resistance of the battery pack, from AESC and AVT data.

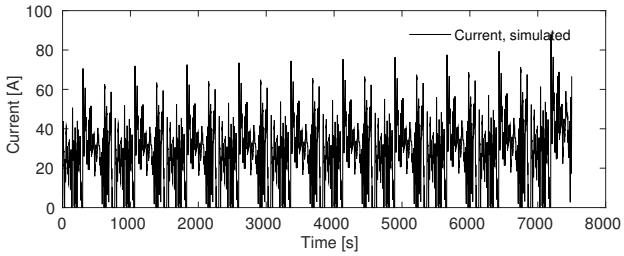


Figure 5. EV current demand: Simulated for a full discharge of the battery pack (see text for details).

Then we combine it with the known power demand the ESS fulfills in a given mission:

$$P_{req}(n) = \alpha_{eff} \cdot v(n) \cdot i(n), \tag{22}$$

where n represents time steps, v is the voltage, and i is the current. We can combine the OCV and Eq. (22), so the problem reduces to iteratively solve a system of two equations and two unknowns. We prove that all the solutions for $i(n)$ are complex conjugate, except for the one shown in Figure 5. This current demand is the fundamental building block to tackle the battery pack SOC/SOH problems.

4. BATTERY PACK AUTONOMY: FILTERING AND PROGNOSIS SCHEMES FOR THE STATE-OF-CHARGE

In the stochastic formulation of the battery autonomy problem, we need to provide confidence intervals for the end-of-discharge (EOD). To achieve the goal we need to consider operational conditions in a given cycle, and how these drive the n-step SOC prediction into the future.

An example comes from the internal resistance: we need to quantify how operational conditions affect it to introduce it effectively into a SOC prediction scheme (Section 4.1). Estimation of the SOC (Section 4.3) is not our objective, but it provides an opportunity to test the model and the simulated deterministic profiles obtained in Section 3. A more challenging issue is the prognosis problem (Section 4.5), where we predict future values of the SOC to plan and take decisions on impending device operation. The difficulty arises from the ignorance about the usage profile on the prognosis stage. Dealing with that uncertainty means resorting to probabilistic characterizations of the EV usage profile (Section 4.4).

4.1. Metamodel A: Internal ohmic resistance

A metamodel can be defined as a surrogate model. It defines the components of a conceptual process, by the development of the constraints and theories applicable to a class of problems (Friedman, 1996). We apply this concept to assess the internal impedance ($|z_{int}|$, or resistance R_{int}) of the EV battery back.

Resistance may depend on many variables, such as temperature, energy stored in the battery, and current demand. We seek to *quantify* R_{int}, given the available information from a single discharge in a resistance AVT test (Espinoza, 2017). The first thing to notice is the obvious dependence of the internal resistance with SOC. We can fit a polynomial function as:

$$\begin{aligned}|z_{int}|(SOC) = 8.14e^{-9} \cdot SOC^4 - 1.94e^{-6} \cdot SOC^3 ... \\ + 1.71e^{-4} \cdot SOC^2 - 0.007 \cdot SOC + 0.226,\end{aligned} \tag{23}$$

for SOC between 0 and 100%, and $|z_{int}|$ measured in Ω.

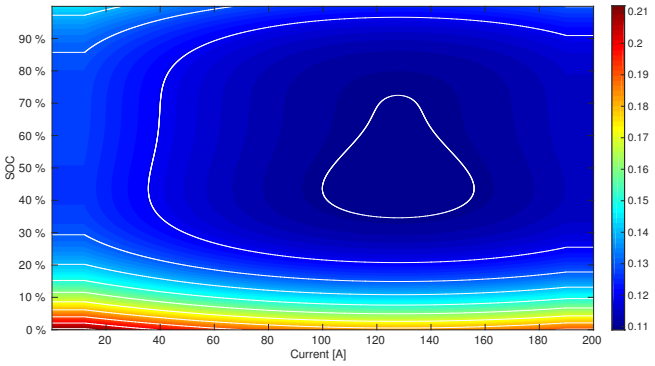


Figure 6. Graphical representation of the obtained model for the internal resistance of the pack. This 2-D lookup table is referred as $|z_{int}|(SOC, I)$ from now on. Color-bar shows its numerical values in Ω.

AESC data give us voltage vs. SOC curves for constant discharge rates: $C/3$, C, 1.84 C, and 2.78 C. Since these refer to new cells, we can aggregate them into a pack and correct for aging effects (Espinoza, 2017). Then we compare to AVT data. If we attribute the different scaling of AESC-based curves purely to the different current demands, we can derive a dependency on current for the internal resistance of the pack. This can be written as a polynomial fit of the form:

$$|z_{int}|(I) = -1.87e^{-9} \cdot I^3 - 4.84e^{-7} \cdot I^2 + 1.3e^{-4} \cdot I + 0.1518, \tag{24}$$

where the current I is saturated below and above 15 and 190 A respectively, and $|z_{int}|$ is measured in Ω. This de-

pendency is much weaker than the SOC one, but still considerable if we take into account that small changes lead to large discrepancies when doing n-step predictions for the SOC. Combining both Eqns. (23) and (24), and assuming separation of variables, we obtain a model (or 2-D lookup table) for the internal resistance as a function of SOC and current. A graphical representation is illustrated in Figure 6 for $|z_{int}|(SOC, I)$. It should be noted that environmental variables are excluded since we don't have priors on them.

4.2. State space model for the SOC

The state space (SS) description for the SOC corresponds to an empirical model based on the reduction of the battery pack operation into an equivalent circuit (Pola et al., 2015; Rahmoun, Biechl, & Rosin, 2012). It is discrete in time and has only two states, $x_1(k)$, an unknown model parameter, and $x_2(k)$, the SOC. The output is given by a voltage signal $v(k)$. We present the model in Eqns. (25), (26), and (27).

State transition equations

$$\begin{aligned} x_1(k+1) &= \varepsilon \cdot x_1(k) + \omega_1(k) \\ \varepsilon_{filtering} = 1, \;\; \varepsilon_{prognosis} &= \frac{|z_{int}|(x_2(k), i(k))}{x_1(k)} \end{aligned} \tag{25}$$

$$\begin{aligned} x_2(k+1) = x_2(k) - \Big[V_L + (V_0 - V_L) \cdot \mathrm{e}^{\gamma \cdot (x_2(k)-1)} + \ldots \\ \alpha V_L (x_2(k) - 1) + (1-\alpha) V_L \cdot (\mathrm{e}^{-\beta} - \mathrm{e}^{-\beta \cdot \sqrt{x_2(k)+\zeta}}) - \\ i(k) \cdot x_1(k) \Big] \cdot i(k) \cdot \Delta t \cdot E_{crit}^{-1} + \omega_2(k) \end{aligned} \tag{26}$$

Measurement equation

$$\begin{aligned} v(k) = V_L + (V_0 - V_L) \cdot \mathrm{e}^{\gamma \cdot (x_2(k)-1)} + \alpha V_L \cdot (x_2(k) + \ldots \\ -1) + (1-\alpha) \cdot V_L \cdot (\mathrm{e}^{-\beta} - \mathrm{e}^{-\beta \cdot \sqrt{x_2(k)+\zeta}}) - \ldots \\ i(k) \cdot x_1(k) + \eta(k). \end{aligned} \tag{27}$$

where the current $i(k)$ $[A]$ and the sample time $\Delta t[s]$ are input variables, and the battery voltage $v(k)$ $[V]$ is the system output. The quantities V_0, V_L, α, β, γ, and ζ completely capture the OCV non-linear behavior, and are to be estimated off-line (see Table 5). E_{crit} is the expected total energy delivered by the pack (that could be inferred from the nominal capacity or the integration of a discharge capacity curves). Process (ω_1 and ω_2) and measurement (η) noises are assumed Gaussian. As we mentioned, x_1 is an unknown model parameter. We associate this state with the instantaneous value of the module of the pack impedance, $|z_{int}|$ (or resistance R_{int}).

A notable feature of this SS model that deviates from (Pola et al., 2015) is the inclusion of an ε factor in the x_1 equation that enriches the phenomenology of the model. While on a filtering stage $\varepsilon = 1$ and we have artificial evolution for the state (Liu & West, 2001); during prognosis we make use of all prior knowledge and x_1 obeys metamodel A.

Table 5. Off-line parameters of the measurement equation of the SOC SS model

Parameter	Value
α	0.0995
β	9.2501
γ	22.5058
ζ	0.0125
E_{crit}	76098121.9823 $[V \cdot A \cdot s]$
V_0	404.0559 [V]
V_L	394.7449 [V]

4.3. Regularized particle filter applied to the SOC SS model

Although not the main objective in this section, estimation is interesting in its own right. For EVs, reliable information about the remaining energy on the ESS is critical. Given that the SOC SS model is non-linear, we resort to a sub-optimal solution which may also be implemented on-line: the PF (Orchard & Vachtsevanos, 2009; Creal, 2012). To exemplify the estimation of both the SOC and the internal resistance during a battery pack discharge, we will rely only on current and voltage measurements. These "measurements" are the profiles we have simulated in Section 3, specifically as shown in Figure 5.

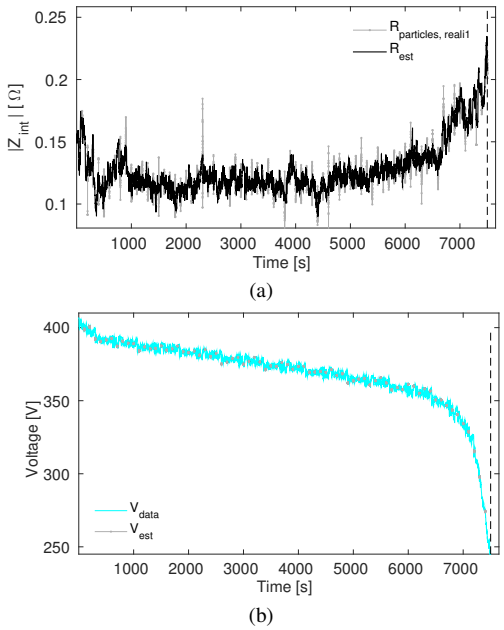


Figure 7. (a) Internal resistance of the battery pack as a function of time, with particles represented as gray points. Their weighted mean average is shown as a black line. (b) Voltage measurement for the battery pack as a function of time (cyan), with particles represented as gray points.

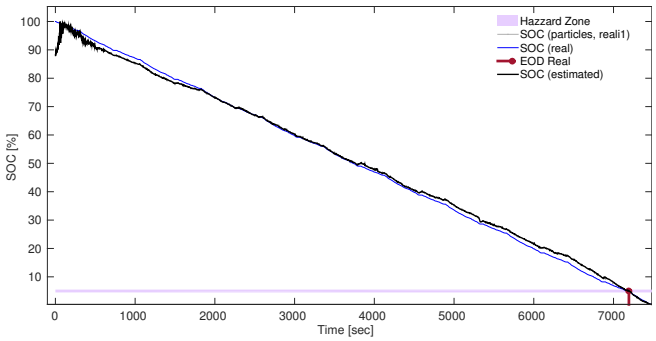


Figure 8. SOC estimation (black curve) rapidly converges to the real value, from erroneous initial conditions. SOC estimation follows the real value (blue line) for the discharge, with deviations $< 2.5\%$

Estimation is performed with unknown initial conditions associated to the SOC. In practice, discharges won't always start from a fully charged state. A uniform distribution between 80% and 90% is assumed as initial condition for x_2. Convergence to the real value is achieved by an adaptive learning strategy (Orchard, Tobar, & Vachtsevanos, 2009).

Results of the estimation for the pack internal resistance and SOC are shown in Figure 7 and 8 respectively. In these we show the propagating pdfs, which are approximated by 40 particles. The internal resistance shows its characteristic functional form, showing that artificial evolution is an effective technique at the filtering stage. There is a clear dependence with SOC even though this wasn't included in the SS model. Accordingly, voltage is also estimated correctly, with very small spread of the particles around the simulated data. SOC estimation shows to rapidly converge to the real value, from a deliberately poor guess for its initial condition. The estimation (black curve) closely follows the ground truth (or real value for the discharge), with deviations $< 2.5\%$.

4.4. Probabilistic characterization of usage profiles based on Markov chains

When dealing with prognosis, a simplified approach is to perform long-term predictions deterministically. The alternative considered here is to define usage profiles in a future horizon, with probabilistic characterizations of them. This is a task plagued with difficulties but potentially more rewarding in terms of accuracy and exactitude of the predictions.

The probabilistic characterization of the usage profile is basically a model that describes how the vehicle is used. This is set by the mission map, EV mechanical parameters and the driving schedule. They are all embedded in the construction of the deterministic usage profile we obtained from Section 3. However, to tackle the n-step prediction problem, we go from deterministic to statistical characterization of the usage profile. For this, we follow (Navarrete, 2014; Pola et al., 2015; Espinoza, 2017): these propose a general methodology to model real-world signals with homogeneous Markov chains. It is important to note that this characterization follows an optimal approach that maximizes the number of states, and minimizes the estimation error of transition probabilities, considering that only finite data is available.

4.5. Prognosis scheme applied to the SOC SS model

When no measurements are acquired, the uncertainty of a current state must be propagated using a regularized PF-based approach, as described in Section 2.4. The goal is to statistically characterize the EOD through an expectation value, 95% confidence intervals, and JITPs (Section 2.5). All this, starting only from the known initial condition at $k_0 = 0$ of a fully charged battery pack, SOC SS model knowledge, and a probabilistic characterization of EV usage profiles via a Markov chain model. Due to the stochastic nature of PF-based predictions, independent randomized realizations are necessary for the statistical characterization of the EOD.

Previous works (Pola et al., 2015; Olivares, Cerda, Orchard, & Silva, 2013) have tested what are the appropriate parameters for the regularized PF-based prognosis algorithm in the case of lithium-ion batteries. These parameters are i) the number of particles to portray the states pdf satisfactorily in each prediction step, ii) the number of iterations of the prognosis to estimate the EOD accurately, and iii) the number of Markov chain realizations to represent uncertainty precisely. By means of a simplified SS model, and direct comparison between an optimal approach (Kalman filter) and the PF, they settled for 40 particles and one iteration of the prognosis algorithm for each of 25 independent Markov chain model realizations. In this work, we arrive to similar conclusions. Using sensitivity tests to determine convergence of our results, we maintain the 40 particles, and perform prognosis for 50 independent realizations of the probabilistic characterization of the usage profile. Both the accuracy and precision of the EOD depend strongly on this last parameter, and due to i) the highly variable current demand in each Markov chain, and ii) that the prognosis horizon covers a whole discharge of the battery pack, we double their number of realizations. All the parameters used in our long-term SOC prognosis are listed in Table 6.

Table 6. Example of Long-term SOC prognosis parameters

Parameter	Value
Number of particles N_p	40
Realizations of Markov model	50
Realizations of prognosis	50
Process Noise (ω_1, ω_2)	$[\mathcal{N}(0, 0.17), \mathcal{N}(0, 0.12)]$
Hazard zone (SOC)	4-6%
x_1 initial condition	0.15
x_2 initial condition	100%
k_0 prognosis start	0 s

Table 7. Example of Long-term SOC prognosis parameters

#s	$\mathbb{E}\{EOD\}$	95% c.i.	$\text{JITP}_{5\%}$	$\text{JITP}_{15\%}$
1	6957	[6658, 7286]	6681	6732
2	7058	[6492, 7477]	6540	6691
3	7070	[6585, 7471]	6620	6754
4	7172	[6924, 7474]	6954	7049
5	7291	[6923, 7492]	6952	7082
6	6855	[6525, 7263]	6559	6649
7	7137	[6639, 7478]	6696	6886
8	7242	[6957, 7478]	7041	7114

Table 7 shows (a fraction) of our results for the $\mathbb{E}\{EOD\}$, 95% confidence intervals, $\text{JITP}_{5\%}$, and $\text{JITP}_{15\%}$ for 10 out of the 50 realizations of the prognosis algorithm. It is worth noting that $\mathbb{E}\{EOD\}$ is indeed a random variable, and that only in a couple of realizations the algorithm overestimates the EOD ground truth (7193 s). This is a positive aspect of our implementation as, in the case of EVs, underestimating the time when a recharge is needed results in not being stranded (and contributes to alleviate the effects of range anxiety).

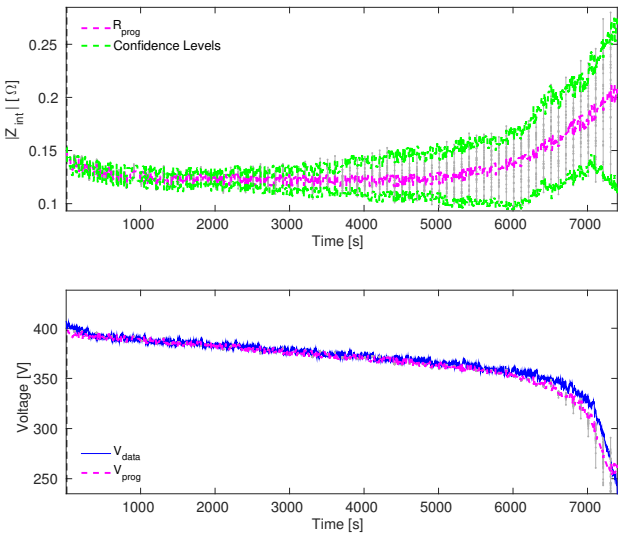


Figure 9. Top: prognosis result for internal resistance (purple). Confidence levels are shown in green, while individual particles are displayed as gray points. Bottom: prognosis result for the voltage (purple). The *measured* voltage is shown in blue, while individual particles are again displayed as gray points.

Figures 9 and 10 illustrate the combined results for all 50 independent and randomized realizations of the algorithm. Figure 9 shows the long-term prognosis for both the internal impedance module and voltage of the battery pack. Here we observe two important features. The first is that the internal resistance follows a very similar trend to the one in Figure 7(a): this effectively means that the results of the more complex SS model (with $\varepsilon_{prognosis}$) include those of the simpler case (with $\varepsilon_{filtering}$). Second, we can see how

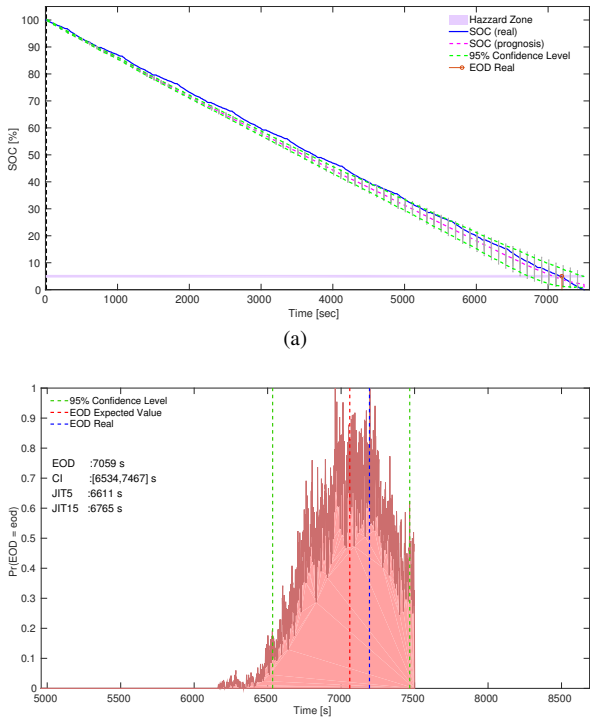


Figure 10. Top: SOC long-term prognosis (purple), while its ground truth is displayed as a blue line. Confidence levels are shown in green, and individual particles as gray points. Bottom: Statistical characterization of the EOD pdf. This is the result of 50 independent realizations of the algorithm. Vertical lines show EOD confidence intervals, expected value, and its ground truth.

the predicted voltage achieves a much better agreement with its ground truth than in previous investigations (Pola et al., 2015), especially at the low-energy end of the curve. The explanation resides in a) the implementation of the metamodel for the internal resistance, which yields to inevitable better results for its long-term prediction, and b) a good parameterization of the OCV curve at the low-SOC interval, with the addition of the ζ parameter).

Figure 10 illustrate the SOC long-term prognosis and EOD characterization. From the latter, we can extract values of $\mathbb{E}\{EOD\} = 7059\ s$, $[6534 - 7467\ s]$ 95% confidence interval, $\text{JITP}_{5\%} = 6611\ s$, and $\text{JITP}_{15\%} = 6765\ s$. The EOD ground truth is $7193\ s$, so again we see how the overall performance of the algorithm stays in the conservative side essential for EVs, and how the JITPs are effective in characterizing the tail of the EOD pdf. The obtained confidence interval is $\sim 13\%$ of the EOD ground truth, which is expected for a prediction horizon that encompasses a full battery pack discharge in which highly variable current demands need to be met.

The obtained statistics and their comparison to ground truth values give us confidence as to using our simulation engine as an effective tool to determine battery autonomy. The problem of *when* an EV will need a re-charge is now one we can answer within an stochastic formulation that helps the decision-making process.

5. BATTERY PACK REMAINING LIFE: PROGNOSIS SCHEMES APPLIED TO THE STATE-OF-HEALTH

Our last objective is to predict battery pack remaining life as a function of i) pack size, and ii) operational conditions through cycles. Given this information, we can establish when a replacement will be required. Again, we work within an stochastic formulation of the problem, looking to provide measures of risk and confidence intervals for the battery end-of-life (EOL). The lithium-ion battery pack is one of the most important components in EVs, and its replacement is expensive. Therefore, the offered warranty must be financially feasible for the company to offer, but attractive enough for drivers to actually buy the EV.

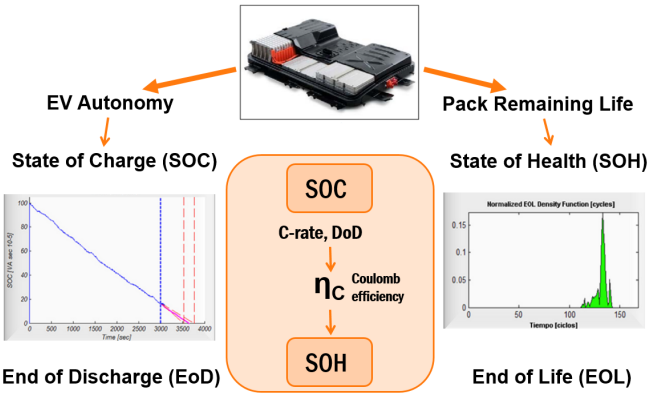


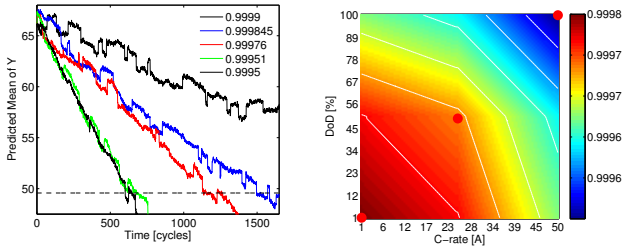
Figure 11. Scheme illustrating how individual cycle conditions, namely DoD and current demand affect the Coulomb efficiency and, in turn, the evolution of the SOH.

The purpose of our simulation engine is to provide a prediction for each set of EV operating conditions. Thus in the n-step SOH prognosis, we consider how individual cycle conditions affect the capacity of the pack. Specifically we build a metamodel which receives depth-of-discharge (DoD) and current demand from the SOC scheme. These inputs are translated into a dependency for the Coulomb efficiency (Section 5.1). In turn, it is this parameter which drives the SOH, as is evident from the simplified model in Section 5.2. The prediction and remaining useful life (RUL) analyses are developed in Section 5.3. Figure 11 illustrate this process, portraying the link between SOC and SOH developed in this paper.

5.1. Metamodel B: Coulomb efficiency

The Coulomb efficiency can be understood from the physics of an electrochemical system. Essentially it describes the ion transfer efficiency between cathode and anode (Bard & Faulkner, 2001). During charging, lithium ions are transferred to the carbon (graphite) anode, forming a layer called solid electrolyte interface. This layer is not removed completely during a discharge, and thickens as cycles go by. In an analogous process, the cathode also develops a restrictive layer known as electrolyte oxidation (Urbain et al., 2008; Miles, 2001).

From an empirical point of view, the Coulomb efficiency η_C is a measure of how much usable energy is expected for the discharge in progress relative to the capacity exhibited during the previous cycle (Olivares et al., 2013). Again, we resort to the concept of metamodeling to describe the Coulomb efficiency. A priori, η_C depends on operational and environmental variables (Bond, Burns, Stevens, Dahn, & Dahn, 2013), such as temperature, pressure, rest times between cycles, depth of discharges, and current intensity for both charge and discharge processes. We select the most important operational variables to enter the η_C metamodel: DoD and current intensity demand during the discharge process.



(a) (b)

Figure 12. (a) Capacity curve simulation for different values of the Coulomb efficiency (b) Metamodel (or 2-D lookup table) for $\eta_C(DoD, I)$.

Since no information about Coulomb efficiency for the Nissan Leaf battery pack is available, we put ourselves in three scenarios, i) one in which η_C strongly depends on DoD and current magnitude, ii) one in which the dependence is weak, and iii) an intermediate case between i) and ii). Figure 12(a) shows what a simulation of a capacity curve for a battery pack ($C_0 = 66.2\ Ah$) might look like for different values of η_C. We choose grids in both DoD:$\{0\%, 50\%, 100\%\}$ and current intensity:$\{1\ A, 25\ A, 50\ A\}$, and by means of piecewise linear interpolation between the three points in each of these grids, we obtain η_C from the simulation as a function of DoD and current, respectively. Combining both dependencies, and assuming separation of variables, we obtain a model (or 2-D lookup table) as shown in Figure 12(b). It is worth noting that this may give rise to capacity curves with variable Coulomb

Table 8. SOH SS model parameters

Parameter	Value
α_2	0.15
β	1
x_1 initial condition	66.2
x_2 initial condition	0.2
Process Noise ω_1	$\mathcal{N}(0, 0.05)$
Process Noise ω_2	$\mathcal{N}(0, 0.05)$
Measurement Noise ν	$\mathcal{N}(0, 0.1)$

efficiencies (i.e. not necessarily a straight line): It all depends on the operational conditions of individual discharge cycles.

5.2. State space model for the SOH

We focus on an empirical SS model. Previous investigations (Olivares et al., 2013) have tested Bayesian filtering and PF-based prognosis that include self-regeneration phenomena. In this work we adopt the simplified approach of (Olivares et al., 2013), where degradation is described by a linear Gaussian dynamic system:

State transition equations

$$x_1(k+1) = \eta_C(DoD, I) \cdot x_1(k) + \omega_1(k) \quad (28)$$

$$x_2(k+1) = \alpha_2 \cdot x_2(k) + \beta \cdot U(k) + \omega_2(k) \quad (29)$$

Measurement equation

$$y(k) = x_1(k) + x_2(k) + \nu(k) \quad (30)$$

where k is the cycle index; x_1 is a state representing the battery SOH; x_2 is a state associated with additional available SOH due to regeneration phenomena; η_C is the Coulomb efficiency, and α_2, β are model parameters. U is an external input, where $U(k) = 1$ if a regeneration phenomenon is detected at cycle k or $U(k) = 0$ if not. Finally, $y(k)$ is the measured SOH. Process (ω_1 and ω_2) and measurement (ν) noises are assumed Gaussian. Parameters employed here are listed in Table 8.

The external input U must be statistically characterized by occurrences of the self-recharge phenomena in a given dataset. Olivares et al. examined 68 of them from accelerated degradation tests at the NASA Ames Prognostic Center of Excellence. We don't count with such dataset and thus have considered[3] $U(k) = 0\ \forall k$ for prognosis-purposes.

5.3. Prognosis scheme applied to the SOH SS model

For long-term predictions, evolution in time is determined by i) the sizing of the lithium-ion battery pack, ii) the usage profile model, and iii) individual cycle conditions as the DoD and current demand. When no measurements are acquired, the uncertainty associated to the state pdf is propagated through a prognosis algorithm. In our case the system is both linear and Gaussian, so we use the prediction equations of the Kalman filter (Kalman, 1960). They are the optimal solution for this problem, and are specified by:

$$\hat{x}(k+1) = A \cdot \hat{x}(k) + B \cdot u(k) \quad (31)$$

$$P(k+1) = A \cdot P(k) \cdot A^T + R_{ww} \quad (32)$$

where $\hat{x}(k)$ represents the state expectation, A and B are matrices that define the state equations and can be built from the values listed in Table 8, $P(k)$ is the state covariance matrix at the $k-th$ cycle, and R_{ww} is the process noise covariance matrix, also in Table 8. We can iterate these equations and compute an n-step prediction for the SOH, defining the hazard zone at a capacity of $49.65\ Ah$ (75% of C_0). This is consistent with Nissan's claim of battery packs retaining $70-80\%$ of their capacity after a decade.

The goal is to statistically characterize the EOL through an expectation value, 95% confidence intervals, and JITPs. All this starting from i) the known initial condition of a new battery pack, and ii) a fixed mission range of $105\ km$, with the usage profile model of Section 4. As aging progresses capacity will diminish, making DoDs much higher, and at some point turning the mission unfeasible. Since DoD increases gradually, a gradual decrease in Coulomb efficiency would be expected as cycles go by. The simulation engine, however, has the capacity to set mission ranges arbitrarily, and also to use different model usage profiles for different cycles. These are input options that the user can choose.

Having chosen the intermediate dependence for $\eta_C(I, DoD)$ (Figure 12), now we can list the steps of the SOH long-term prediction scheme and then analyze its results:

- a) Start from the initial conditions listed in Table 8.
- b) For the discharge cycle, run the prognosis scheme applied to the SOC SS model (see Section 4 for details and the steps of that algorithm).
- c) From all prognosis and Markov chain realizations in b), extract an average for the DoD and current in that discharge.
- d) Use average DoD and I magnitude to evaluate metamodel B, $\eta_C(DoD, I)$.
- e) Using the obtained value for $\eta_C(DoD, I)$, run the prediction stage of the Kalman filter, Eqns. (31) and (32), to evolve the SOH.
- f) Having updated the capacity of the battery pack, go back to b), and iterate.
- g) Iteration finishes when capacity has gone down the hazard zone (75% of C_0), and EOL statistical characterization is achieved.

[3] However, we have kept the equations in their original form in the code, so it is an easy implementation when the external input can be characterized

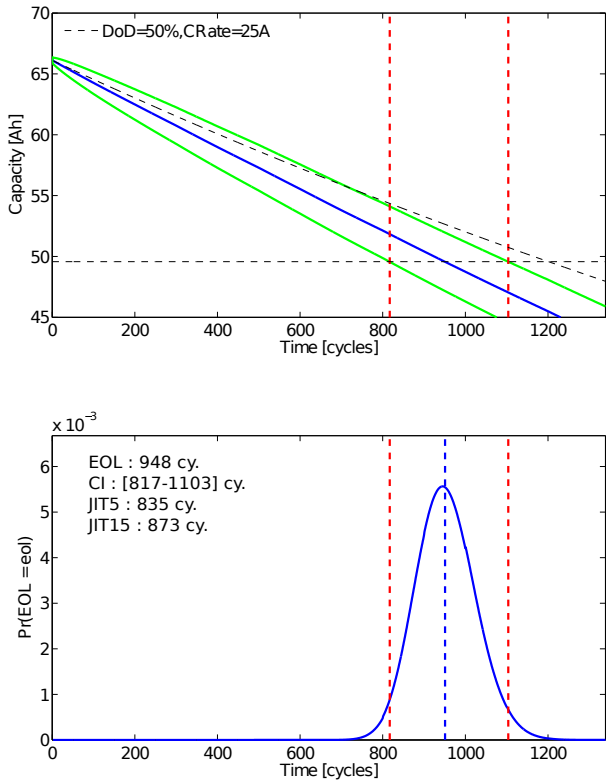


Figure 13. Top: SOH long-term prognosis (blue line) and its 95% confidence level at each cycle (green lines). The dash-black degradation curve represent the evolution of the system if a constant $\eta_C(DoD = 50\%, I = 25\ A)$ had been considered for all cycles. The horizontal dash line represent 75% of C_0. Bottom: EOL characterization, the blue vertical line represents $\mathbb{E}\{EOL\} = 948$ cycles, while the red lines show the 95% confidence interval between cycles $[817 - 1103]$.

Figure 13 (top) illustrate the SOH long-term prognosis. It is important to note how our choice of an intermediate dependence for $\eta_C(DoD, I)$ (Figure 12) governs future behavior. If we were to use a constant $\eta_C(DoD = 50\%, I = 25\ A)$ through cycles, effectively neglecting operational conditions, we would have obtained the dash-black line as the degradation curve of the system. That would result in an overestimation of the RUL of more than 200 cycles! This is the reason why taking into account metamodel B -and thus the interplay between SOC predictions and SOH evolution- is so important. Figure 13 (bottom) shows the EOL characterization. We can extract values of $\mathbb{E}\{EOL\} = 948$ cycles, 95% confidence interval between cycles $[817 - 1103]$. $\text{JITP}_{5\%}$ and $\text{JITP}_{15\%}$ occur at cycles 835 and 873, respectively, characterizing the tail of the EOL pdf. The obtained confidence interval is $\sim 30\%$ of the expected EOL. It is the consequence of uncertainty propagation in a simulation engine that i) performs long-term prognosis for SOH (and the SOC in every individual cycle), and ii) considers the interplay between operational conditions, and how these affect both metamodels for internal resistance and Coulomb efficiency (and, in turn, the SOC and SOH).

Now we present another run of the engine, to illustrate how results change with another prescription for the Coulomb efficiency. First we choose the strong dependency of η_C on DoD and discharged current intensity. We aim to maximize the information found in the SOC prognosis step, so instead of just extracting averages for the DoD and current, we obtain probability mass functions (pmfs) for them. A kernel smoothing function is applied to the pmfs (Bowman & Azzalini, 1997) to derive pdfs. The situation changes from having scalars (in the previous run) to pdfs now.

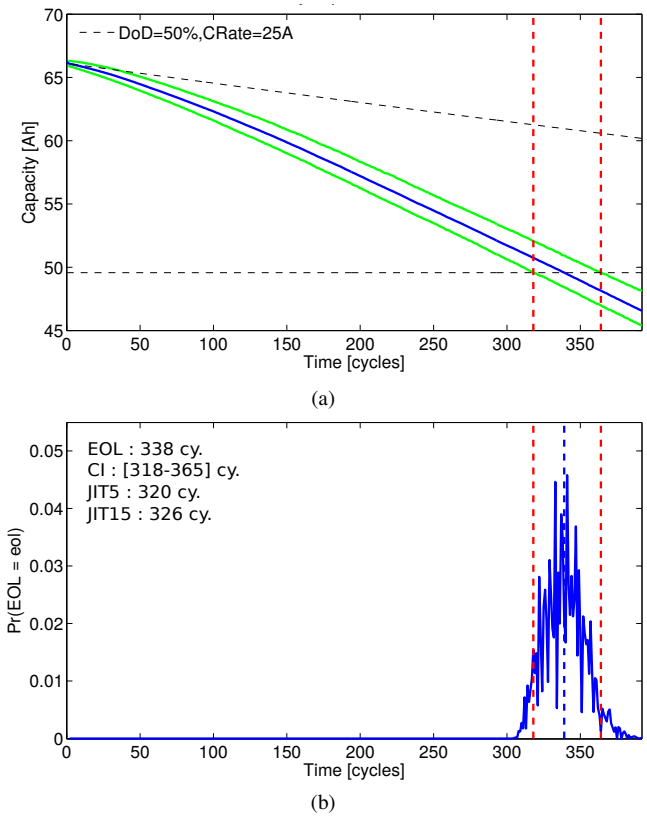


Figure 14. (a) SOH long-term prognosis (blue line) and its 95% confidence level at each cycle (green lines). The horizontal dash line represent 75% of C_0. (b) EOL characterization, the blue vertical line represents $\mathbb{E}\{EOL\} = 338$ cycles, while the red lines show the 95% confidence interval between cycles $[318 - 365]$.

The new results are presented in Figure 14. There is a striking difference with respect to those of Figure 13, due to our choice of an strong dependence for $\eta_C(DoD, I)$. This translates, as expected, into a much reduced remaining useful life for the battery pack (top-left panel). As now we are introducing a pdf instead of an scalar for η_C, the state expectation propagation $\hat{x}(k)$ is not longer Gaussian. Due to this same reason, the EOL pdf is not symmetric as in the previous run.

Figure 14 (bottom) shows the EOL characterization. We can extract values of $\mathbb{E}\{EOL\} = 338$ cycles, 95% confidence interval between cycles $[318 - 365]$. $\text{JITP}_{5\%}$ and $\text{JITP}_{15\%}$ occur at cycles 320 and 326, respectively, characterizing the tail of the EOL pdf. This much reduced expectation for the useful life does not correlate with the Nissan Leaf actual performance, making the case for a strong dependency of metamodel B weak.

5.4. Final considerations

As in the previous section, the obtained statistics and their comparison to real-life values for the Nissan Leaf give us confidence as to using our simulation engine as an effective tool to determine battery pack remaining life. The problem of *when* a replacement will be needed is now one we can provide answers within an stochastic formulation that helps the decision-making process.

6. CONCLUSIONS

When claims on range or remaining useful life for a battery pack are made by auto companies, operational conditions are often not depicted clearly. The public need to take these claims as conservative expectations, as no statistical details are given. What happens if the EV is driven aggressively? Or is charged erratically, or always operated until all energy is drained? How do these conditions affect both the recharge in a cycle-by-cycle basis, and the prospects of an expensive battery pack replacement in the long-term?

We have developed a simulation engine able to answer these questions. Given a series of user-defined inputs, it can determine battery autonomy and remaining life from an stochastic point of view, i.e. it provides expectation values, confidence intervals, and risk measures for the EOD and EOL. This is sufficient to take informed decisions about the state of our car. We have taken into account operational conditions on individual cycles, and how they affect single discharges and, in turn, the overall evolution of the SOC and SOH (driven by the metamodels A and B, on internal ohmic resistance and Coulomb efficiency respectively).

ACKNOWLEDGMENT

This work has been partially supported by FONDECYT Grant Nr. 1170044, and the Advanced Center for Electrical and Electronic Engineering, Basal Project FB0008. Aramis Pérez was supported by the University of Costa Rica (Grant for Doctoral Studies) and CONICYT-PCHA/Doctorado Nacional/2015-21150121.

NOMENCLATURE

AESC	Automotive Energy Supply Corporation
AVT	Advanced Vehicles Testing Activity
DoD	Depth-of-Discharge
EOD	End-of-Discharge
EOL	End-of-Life
ESS	Energy Storage System
EV	Electric Vehicle
HWFET	Highway Fuel Economy Schedule
OCV	Open Circuit Voltage
pdf	Probability Density Function
pmf	Probability Mass Function
PF	Particle Filter
RUL	Remaining Useful Life
SIS	Sequential Importance Sampling
SMC	Sequential Monte Carlo
SOC	State-of-Charge
SOH	State-of-Health
SS	Space-State

REFERENCES

Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, *50*, 174–188.

Bard, A., & Faulkner, L. (2001). *Electrochemical methods fundamentals and applications*. John Wiley and Sons.

Bond, T., Burns, J., Stevens, D., Dahn, H., & Dahn, J. (2013). Improving precision and accuracy in coulombic efficiency measurements of li-ion batteries. *J. Electrochem. Soc*, *160*, 521-527.

Bose, B. (2010). Global warming: Energy, environmental pollution, and the impact of power electronics. *IEEE Industrial Electronics Magazine*, *4*, 6-17.

Bowman, A., & Azzalini, A. (1997). *Applied smoothing techniques for data analysis*. Oxford University Press Inc.

Charkhgard, M., & Farrokhi, M. (2010). State-of-charge estimation for lithium-ion batteries using neural networks and ekf. *IEEE Transactions on Industrial Electronics*, *57*, 4178-4187.

Creal, D. (2012). A Survey of Sequential Monte Carlo Methods for Economics and Finance. *Econometric Reviews*, *31*, 245–296.

Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, *10*, 197–208.

Doucet, A., & Johansen, A. (2008). A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, *12*, 656–704.

Espinoza, P. (2017). *A simulation engine for ion-lithium battery packs in electric vehicles based on energetic au-*

tonomy and remaining useful life criteria (Unpublished master's thesis). University of Chile.

Friedman, L. W. (1996). *The simulation metamodel*. Kluwer Academic Publishers.

Gregory, P. (2005). *Bayesian logical data analysis for the physical sciences*. Cambridge University Press.

Han, H., Xu, H., Yuan, Z., & Shen, Y. (2014). A new soh prediction model for lithium-ion battery for electric vehicles. In *2014 17th international conference on electrical machines and systems* (p. 997-1002).

Ikezoe, M., Hirata, N., Amemiya, C., & Miyamoto, T. (2012). *Development of high capacity lithium-ion battery for nissan leaf* (Tech. Rep.). SAE Technical Paper.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *ASME. J. Basic Eng.*, *82*, 35-45.

Kim, N., Rousseau, A., & Rask, E. (2016). Parameter estimation for a lithium-ion battery from chassis dynamometer tests. *IEEE Transactions on Vehicular Technology*, *65*, 4393-4400.

Liu, J., & West, M. (2001). Sequential monte carlo methods in practice. In (p. 197-223). Springer.

Macharis, C., Lebeau, P., Mierlo, J. V., & Lebeau, K. (2013). Electric versus conventional vehicles for logistics: A total cost of ownership. In *Electric vehicle symposium and exhibition* (p. 1-10).

Miles, M. (2001). Recent advances in lithium battery technology. In *Gallium arsenide integrated circuit symposium* (p. 219-222).

Musso, C., Oudjane, N., & Gland, F. L. (2001). Sequential monte carlo methods in practice. In (p. 247-260). Springer.

Navarrete, H. (2014). *Caracterización estadística del perfil de uso de baterías para el pronóstico del estado-de-carga* (Unpublished master's thesis). University of Chile.

Olivares, B., Cerda, M., Orchard, M., & Silva, J. (2013). Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. *IEEE Transactions on Instrumentation and Measurement*, *62*, 364–376.

Orchard, M., Tang, L., Goebel, K., & Vachtsevanos, G. (2009). A novel rspf approach to prediction of high-risk low-probability failure events. In *Annual conference of the prognostics and health management society.*

Orchard, M., Tang, M., Saha, B., Goebel, K., & Vachtsevanos, G. (2010). Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices. *Studies in Informatics and Control*, *19*, 209-218.

Orchard, M., Tobar, F., & Vachtsevanos, G. (2009). Outer feedback correction loops in particle filtering-based prognostic algorithms: Statistical performance comparison. *Studies in Informatics and Control*, *18*, 295-304.

Orchard, M., & Vachtsevanos, G. (2009). A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control*, *31*, 221–246.

Pattipati, B., Sankavaram, C., & Pattipati, K. (2011). System identification and estimation framework for pivotal automotive battery management system characteristics. *IEEE Transactions on Systems, Man, and Cybernetics Society*, *41*, 869-884.

Penna, J., Nascimento, C., & Rodrigues, L. (2012). Health monitoring and remaining useful life estimation of lithium-ion aeronautical batteries. In *Aerospace conference, 2012 ieee* (p. 1-12).

Pola, D., Navarrete, H., Orchard, M., Rabié, R., Cerda, M., Olivares, B., ... Pérez, A. (2015). Particle-filtering-based discharge time prognosis for lithium-ion batteries with a statistical characterization of use profiles. *IEEE Transactions on Reliability*, *64*, 710-720.

Rahmoun, A., Biechl, H., & Rosin, A. (2012). Soc estimation for li-ion batteries based on equivalent circuit diagrams and the application of a kalman filter. In *Electric power quality and supply reliability conference* (p. 1-4).

Saha, B., Goebel, K., Poll, S., & Christophersen, J. (2009). Prognostics methods for battery health monitoring using a bayesian framework. *IEEE Transactions on Instrumentation and Measurement*, *58*, 291-296.

Salkind, A., Fennie, C., Singh, P., Atwater, T., & Reisner, D. (1999). Determination of state-of-charge and state-of-health of batteries by fuzzy logic methodology. *Journal of Power Sources*, *80*, 293 - 300.

Snihir, I., Rey, W., Verbitskiy, E., Belfadhel-Ayeb, A., & Notten, P. (2006). Battery open-circuit voltage estimation by a method of statistical analysis. *Journal of Power Sources*, *159*, 1484–1487.

Urbain, M., Rael, S., Davat, B., & Desprez, P. (2008). Energetical modelling of lithium-ion battery discharge and relaxation. In *Ieee power electronics specialists conference* (p. 3628-3634).

Wansart, J., & Schneider, E. (2010). Modeling market development of electric vehicles. In *Annual ieee systems conference* (p. 371-376).

Weng, C., Sun, J., & Peng, H. (2013). An Open-Circuit-Voltage Model of Lithium-Ion Batteries for Effective Incremental Capacity Analysis. In *Dynamic systems and control conference.*

Wishart, J., Carlson, R., Chambon, P., & Gray, T. (2013). The electric drive advanced battery project: Development and utilization of an on-road energy storage system testbed. In *Sae world congress* (p. 1-30).

Xiong, R., He, H., Sun, F., & Zhao, K. (2013). Evaluation on state of charge estimation of batteries with adaptive extended kalman filter by experiment approach. *IEEE Transactions on Vehicular Technology*, *62*, 108-117.

Yang, H., Gao, Y., Farley, K., Jerue, M., Perry, J., & Tse, Z. (2015). Ev usage and city planning of charging station installations. In *Wireless power transfer conference* (p. 1-4).

Young, K., Wang, C., Wang, L., & Strunz, K. (2013). *Electric vehicle integration into modern power networks*. Springer Science.

Zeff, S. (2016). My electric journey with a nissan leaf. *IEEE Consumer Electronics Magazine*, *5*, 79-80.

BIOGRAPHIES

Pablo A. Espinoza received his B.S. from Catholic University of Chile in 2007, and a M.S. in Astronomy from the University of Arizona in 2012. In 2017 he obtained a M.S. in Electrical Engineering from University of Chile. His research interests include control systems, prognostics, sensor fusion, computer vision, and electric/self-driving cars.

Aramis Pérez is a Research Assistant at the Lithium Innovation Center (Santiago, Chile) and Professor at the School of Electrical Engineering at the University of Costa Rica. He received his B.Sc. degree (2002) and Licentiate degree (2005) in Electrical Engineering from the University of Costa Rica. He received his M.Sc. degree in Business Administration with a General Management Major (2008) from the same university. Currently he is a doctorate student at the Department of Electrical Engineering at the University of Chile under Dr. Marcos E. Orchard supervision. His research interests include parametric/non-parametric modeling, system identification, data analysis, machine learning and manufacturing processes.

Marcos E. Orchard is Associate Professor with the Department of Electrical Engineering at Universidad de Chile and was part of the Intelligent Control Systems Laboratory at The Georgia Institute of Technology. His current research interest is the design, implementation and testing of real-time frameworks for fault diagnosis and failure prognosis, with applications to battery management systems, mining industry, and finance. His fields of expertise include statistical process monitoring, parametric/non-parametric modeling, and system identification. His research work at the Georgia Institute of Technology was the foundation of novel real-time fault diagnosis and failure prognosis approaches based on particle filtering algorithms. He received his Ph.D. and M.S. degrees from The Georgia Institute of Technology, Atlanta, GA, in 2005 and 2007, respectively. He received his B.S. degree (1999) and a Civil Industrial Engineering degree with Electrical Major (2001) from Catholic University of Chile. Dr. Orchard has published more than 100 papers in his areas of expertise.

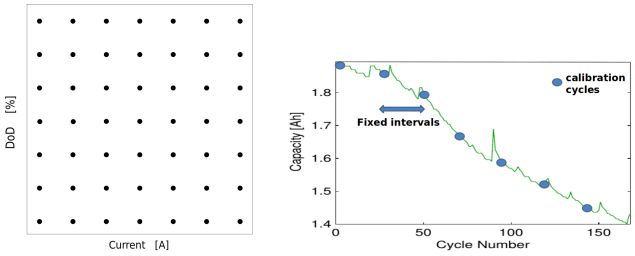
Hugo F. Navarrete received the B.Sc. degree in electrical engineering, and certification in civil electrical engineering from the University of Chile in 2013 and 2014, respectively. His research interests include health management in energy storage devices, and automatic control.

Daniel A. Pola received his B.S. and M.Sc. degrees in electrical engineering from the University of Chile in 2012 and 2014, respectively. His research interests include estimation and prognosis based on Bayesian algorithms, health management for energy storage devices, system identification, and control systems.

APPENDIX

We briefly establish a laboratory methodology for the discharge of lithium-ion cells. More details and results are given in (Espinoza, 2017). The methodology needs to characterize i) the Coulomb efficiency parameter that govern degradation curves, and ii) the internal impedance of a given type of cell. In the absence of *exhaustive* empirical data, these quantities have been assesed by metamodels A and B.

For i), we investigate the Coulomb efficiency dependency on DoD and current. This can be derived from the slope of the capacity curve on degradation tests that extend over the useful life of the battery. For ii) we characterize the internal resistance as a function of SOC, current and aging. This can be derived from special calibration discharge cycles, with spaced current pulses.



(a) (b)

Figure 15. (a) Schematic of a 2-D grid of DoD and current magnitude, where each point represents an experiment of a discharging cell. (b) Typical degradation curve for a lithium-ion cell. In our design, every cycle along the green line would be a discharge at a given DoD *and* current level. Blue circles would represent calibration cycles, spaced at regular intervals.

To understand the proposed experimental design, let us consider Figure 15(a). Schematically, this represents a 2-D grid (DoD and current), with each black point representing a discharge. Lithium-ion cells can have useful lifes of $>$ hundreds of cycles, and we are looking for significant degradation to obtain the slope of the downward capacity curve $\eta_C(DoD, I)$. It is easy to see why testing for dependencies on more *independent* variables, or dimensions (than DoD, I), is not viable for reasonable time-scales.

Figure 15(b) illustrates a typical degradation curve for a lithium-ion cell. In our proposed experimental design, each cycle at the green line would be a discharge at a given DoD *and* current level, i.e. a black point from the previous plot. These would be repeated over time, until interrupted at regular intervals for two calibration cycles. The interval is set at 40 cycles (or 5% of the expected useful life), We know these are intended to provide information about the internal resistance, but what do we mean exactly by a calibration cycle? This type of cycle corresponds to a *full* discharge, and feature a series of pulses as a function of time (or SOC). These need to be sufficiently spaced, such that the battery does not exhibit a capacitor-like behavior. The intensity of the pulses will give us the R_{int} dependency with current magnitude. And if we repeat these calibrations cycles along the capacity curve, over many cycles we will also acquire a dependency on aging, thus obtaining $R_{int}(SOC, I, \tau)$.