



HAL
open science

Robust and adaptive online estimation of Li-ion battery cell capacity

Paul-Henri Michel, Vincent Heiries

► **To cite this version:**

Paul-Henri Michel, Vincent Heiries. Robust and adaptive online estimation of Li-ion battery cell capacity. VPPC 2023 - 2023 IEEE Vehicle Power and Propulsion Conference, Oct 2023, milan, Italy. 10.1109/VPPC60535.2023.10403135 . cea-04520715

HAL Id: cea-04520715

<https://cea.hal.science/cea-04520715>

Submitted on 25 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Robust and adaptive online estimation of Li-ion battery cell capacity

MICHEL Paul-Henri
CEA-LITEN/L2EP

Power Electronics and Energy Management Lab
17 rue des martyrs, 38054 Grenoble, France
e-mail: paul-henri.michel@cea.fr

HEIRIES Vincent
CEA-LETI/L2EP

Power Electronics and Energy Management Lab
17 rue des martyrs, 38054 Grenoble, France
e-mail: vincent.heiries@cea.fr

Abstract—The proposed algorithm aims to estimate the instantaneous capacity of an electrical battery. The battery capacity is a crucial parameter of the battery management system and the knowledge of its value is necessary for the state of charge, state of function and remaining useful life estimators. Moreover, operation of lithium batteries raises safety issues, and the very accurate estimation of the battery capacity makes it possible to always guarantee its use in its Safe Operating Area (SOA). As a starting point, the classical Coulomb counting relation is taken as an observation equation in a Kalman framework based on a state equation of the capacity evolution. The observation equation is a biased linear relation with the variables of interest being affected by errors and corrupted by outlier samples. A random consensus resampling is thus applied to reject these outliers and a pre-estimation of the ordinate at the origin of the affine function is performed. The bias reduction is performed thanks to a modified hough-transform. This hough-transform uses the a-priori prediction of the capacity and a multi-resolution recursive approach is used to optimize the bias reduction while limiting the computational complexity. The algorithm is robust to the presence of a high level of outlier and the Kalman approach makes the estimation adaptive. Moreover a statistical test is realized to determine whether the estimated capacity should be directly injected back in the SOC estimator or not. Our algorithm is tested on a database delivered by the NASA DashLink platform. Statistical simulations confirm the robustness and the adaptive nature of the method. The obtained RMS of the relative errors over the entire ageing cycles are below 3%.

Index Terms—state of health, robust estimation, capacity, ageing model, kalman filtering

I. INTRODUCTION

The increasing use of electrical batteries is a clear attempt to be prepared for the “post-oil” era. But even if the development of the electric mobility is a vertuous trend, the mineral and metal resources involved are critically limited. The case of lithium salt is a flagrant example. In the next few years the part of lithium demand in the battery sector would be up to 50% of the entire demand with an average growth rate of 15% [1]. So, on the one hand the recycling and ecological use of batteries is a necessity, and on the other hand the battery should be operated optimally. Consequently, it is required to estimate the battery indicators accurately. In this paper we focus on the capacity estimation. Indeed, a precise and robust estimation of the battery capacity leads to a more efficient usage and allow to increase its operating margins. Moreover,

as the capacity is a crucial parameter of the cell, with a more precise estimation, a safe operation can be guaranteed, and the probability of hazardous event can be reduced. These two benefits improve the life cycle of the battery. The capacity of electrical batteries cannot be directly measured because its value results in a complex relationships between electrical and chemical parameters of the internal components of the cell. Thus, since the capacity cannot be measured it must be estimated. There exists many laboratory protocols to determine the nominal capacity at given operating conditions but these methods are offline ones and consequently they cannot be embedded online in a Battery Management System (BMS). The aim is to estimate online regularly the battery cell capacity to not only evaluate the State of Health (SOH), but also feed the other algorithms of the BMS that estimate the other indicators : State of Charge (SOC), State of Available Power (SoAP), State of Function (SOF), State of Energy (SOE). Current estimators are well documented in the literature. The total least squares is currently an interesting method to obtain the capacity and a recursive approach is formulated by Plett[2]. Obviously many papers deal with machine learning techniques such as support vector regression or neural networks[3]. However learning algorithms need an exhaustive database of battery cycling to demonstrate their efficiency and ability to generalize. Such databases exist but only for a few range of cell references, and are very expensive to build. The incremental analysis is also a laboratory method but it can be used online coupled with learning technique[4]. The problematic of the database remains here the same. Researchers are currently developping electrical and thermal models of the capacity evolution[5]. These models are crucial but it cannot be used in an embedded system without a signal processing approach as the inputs are biased and corrupted by different errors. The precision of these cited methods is mostly confirmed, but here we focus also on the robustness to outliers. The combination of a state - observation equations aims to increase the average precision. For that we generate a scatter plot of the variations of SOC and current on a thin time grid. The robust estimation of the slope gives us the capacity update. This estimation is performed thanks to a random consensus integrated in the Kalman filter. The state equation of the capacity evolution does not use neither the SOC nor the current. This equation

can be minimalist, at least, we just want a global tendency of the capacity decrease. A pre-processing technique is necessary to reduce the bias mostly caused by the current. Therefore, we perform a bias reduction knowing the prediction of capacity. Once this bias reduction is performed the entire consensus and estimation are computed. At the end, the test hypothesis can be processed in parallel and the new capacity is injected back in the SOC estimator once the benefit of it is confirmed. We start with some definitions and context descriptions. The algorithm protocol is then presented. Finally, preliminary results are summarized with their statistical data and visual plots.

II. PROBLEM STATEMENT

In the following, the $\overline{(\cdot)}$ operator stands for the empirical mean while μ is the mathematical expectation. The capacity evolution is difficult to model because the underlying physical phenomena are dependent of a lot of correlated variables linked together by non-linear functions. Furthermore, except for very specific applications, the diversity of uses makes the capacity estimation and prediction highly complex. In [6], A. Barré gives a statistical analysis of the battery degradations. The capacity depletion is dependent of the time related variables, battery SOC profile, battery temperature and ambient temperature. As the variation of SOC for a given cycle is a time dependent variable then this SOC variation is linked to the capacity evolution. Indeed, the Coulomb counting is defined by the following equation on the sampled data. We denote the sampling period T_e and I^{th} the true current, τ the Coulomb efficiency, \widehat{SOC} the estimated SOC and C_{opt} the capacity that minimizes the observation error. These variables are expressed in the international system of units except for the capacity which is in A.h. The index i designs the i^{th} statistical realisation :

$$\begin{aligned} \Delta SOC_{k_i} &= \sum_{j=0}^{N_i} (SOC_{k_i+j+1} - SOC_{k_i+j}) \\ &= \sum_{j=0}^{N_i} \tau \frac{T_e I_{k_i+j}^{th}}{3600 C_{opt}(k_i+j)}, k_i + N_i < k_{i+1} \end{aligned} \quad (1)$$

We generate at instant l ($l = l + 1$ once $k_i = k_M$) a set Ψ_l of M samples $S_{k_i} = (\widehat{SOC}_{k_i}, \Delta I_{k_i})$ where $\eta_{I_{k_i+j}}$ is the additive current noise and it follows an uncentered and non-stationary gaussian noise. It can be demonstrated using a second order stationary condition on the capacity variations that (1) leads to :

$$\begin{aligned} \Delta I_{k_i} &= \sum_{j=0}^{N_i} \tau \frac{T_e}{3600} (I_{k_i+j}^{th} + \eta_{I_{k_i+j}}) \\ &\approx C_{opt}(l) \Delta SOC_{k_i} + \tau \frac{T_e}{3600} \sum_{j=0}^{N_i} \eta_{I_{k_i+j}} \\ &= \Delta I_{k_i}^{th} + \eta_{\Delta I_{k_i}} \end{aligned}$$

The real SOC is unknown but we are considering here that we have an estimation of it obtained as the output of a recursive and stationary biased or unbiased estimator. The

indicator variable $\xi_k^{(i)}$ is set to 1 when the SOC variation is aberrant (i.e. if follows the outlier process Θ_{k_i}) and to 0 in inlier case :

$$\Delta \widehat{SOC}_{k_i} = (1 - \xi_k^{(i)}) \Delta \widehat{SOC}_{k_i} + \xi_k^{(i)} \Theta_{k_i}$$

The inlier variable $\Delta \widehat{SOC}_{k_i}$ is gaussian and unbiased distributed, and the density of $\Delta \widehat{SOC}_{k_i}$ knowing $\xi_k^{(i)} = 0$ remains centered and tends towards gaussianity. We define the inlier region in space (disjoint of the outlier one) by the distance criteria d which incorporates the prediction of capacity and the sample S_{k_i} :

$$S_{k_i} \in \Psi_l^{in} \Leftrightarrow d(S_{k_i}, \widehat{C}_l) < d_i^{MAX}(C_{opt}(l))$$

This distance can be defined by the frontier of the 3σ region of the joint density of S_{k_i} or even by a maximal distance criteria (innovation or orthogonal distance) relative to the expected precision on the estimated capacity. This distance is denoted d_i^{MAX} and the inlier region is called Ψ_l^{in} .

A. Analysis of the bias term

The current measurement is biased and corrupted by an additive and centered noise $\eta_{\Delta I_{k_i}}^c$ and we define T as the average cell temperature :

$$\Delta I_{k_i} = \Delta I_{k_i}^{th} + N_i \underbrace{\mu_{I,c}}_{offset} + \sum_{j=0}^{N_i} k(T_{k_i+j}, I_{k_i+j}^{th}) \underbrace{I_{k_i+j}^{th}}_{proportional} + \underbrace{\eta_{\Delta I_{k_i}}^c}_{noise}$$

If only translation and rotation are caused by the offset terms then we can define the y-intercept. The scatter shape will be implicitly analyzed in the intercept reduction algorithm. This intercept term is the time-dependent mean of the measurement errors. Its theoretical definition is the statistical mean of ΔI_{k_i} knowing the SOC variation is null :

$$b_l^{th} = E[\Delta I_{k_i} | \Delta SOC_{k_i} = 0], \forall C_{opt}, \forall k_i \leq k_M$$

In other terms this intercept is the y-intercept of the affine function which models the dependency between the SOC variation and the charge of current. The delta of current is the combination of the intercept b_l^{th} , rotational and noised terms :

$$\Delta I_{k_i} = \Delta I_{k_i}^{th} + b_l^{th} + \underbrace{\mathcal{R}(b_l^{th}, N_i)}_{rotation} + \underbrace{\mathcal{D}(N_i, \overline{N_i})}_{distorion} + \eta_{\Delta I_{k_i}}^c$$

The rotation component and distorsion terms are shown in Figure.1 and we understand that these last terms (plus noise variable) are, in mean sense, independent of the y-intercept (b_l^{th}). In other words, it means that the scatter is centered around a line and this line can be the result of a translation and rotation transforms.

As the observation equation has two degrees of freedom composed of the intercept and the line slope which is the capacity, then an unbiased y-intercept estimator must be independent of any capacity related variables if and only if the distorsion term is zero mean and uniformly distributed.

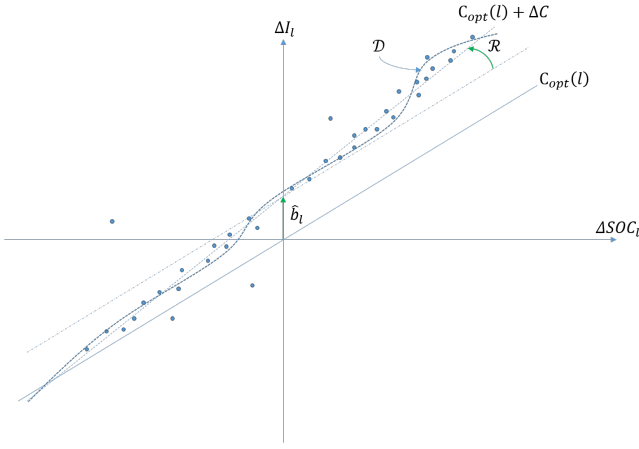


Fig. 1. y-intercept and other terms

B. A state equation of the capacity evolution

The assumption is made here that there exists a state model of the capacity evolution (as an example, obtained through data processing of test campaign). This model can be non-linear and is denoted f , we can thus combine the state and the observation in the following state-space model where the delay between l and $l - 1$ is at least $\sum_{i=1}^M N_i T_e$:

$$S_{k_i} \in \Psi_l \rightsquigarrow \begin{cases} C_l = f(C_{l-1}, T_l, \dots) + w_l & [1 \times 1] \\ \Delta I_l = C_l \Delta \widehat{SOC}_l + b_l^{th} + V_l & [M \times 1] \end{cases} \quad (2)$$

Owing to [7] in III.B the objective function to be minimized in the Kalman filter can also be written as the sum of the two respective Mahalanobis' norms computed on state errors and on the measurement errors. In (2) the measurement equation is a linear and non-stationary regression form with errors in the variables and outlier presence. Under these conditions the Kalman update would be non-robust due to the leverage point effect on the Mahalanobis norm of the measurement errors caused by outlier presence which are non-modeled. Therefore, the inliers and the outliers must be classified in order to only make the update on the basis of the inlier set, i.e. classify the data to define the state-space model (2) on $S_{k_i} \in \Psi_l^{in}$ and define its noise terms. The state noise is a gaussian, white and centered process. Its variance is denoted $E[w_l^2] = Q_l$. This noise term w_l is assumed independent with the state C_m for $m \leq l$, moreover the measurement noises are mutually independent with the state. As a result, the inputs of f must be independent of the noise terms contained in the SOC variation. If we note \vec{V}_l the random vector of the measurement process then this zero mean gaussian and identical distributed process has the following correlation matrix $E[V_l V_l^T] = R_k$ which is not necessarily diagonal. Its components are :

$$\vec{V}_l^{(i)} = \mathcal{R}(b_l^{th}, N_i) + \mathcal{D}_i(N_i, \bar{N}_i) + \eta_{\Delta I_{k_i}}^c - C_l \eta_{\Delta SOC_{k_i}} \quad (3)$$

The conditional density function of \vec{V}_l knowing N_i is the density of the sum of two independent and gaussian random variables which is a gaussian random variable. It can be shown

that the distortion term is relative to the standard empirical deviation of N_i and the empirical mean of theoretical current plus its constant offset. The distortion \mathcal{D}_i is not necessarily stationary here and is not always evenly distributed in the space of statistical realisations.

III. METHODOLOGY

A. Creation of set Ψ_l

The time dynamics of the capacity evolution gives us the minimal rate of estimation updates. Furthermore, increasing the range of ΔSOC for a given M improves the precision of the capacity estimation. But generating samples at high ΔSOC requires a long delay and thus decreases the estimation rate. It can also be noticed that generating samples at very low ΔSOC makes the estimator faster but the signal-to-noise ratio (SNR) is clearly reduced. A trade-off between the SNR and the slowest time dynamic must be respected. We suppose that the upper limit $\epsilon_C(l)$ (the error on capacity approximation) is known and leads to an order of magnitude of the maximal number of cycles between two capacity updates. Then, given M , the sampling distribution of ΔI_{k_i} bounds can be tailored.

An intuitive distribution is the truncated gaussian density function because its range is broad while concentrating the values around a well-chosen mean. The global range may be divided into M evenly distributed segments. If we refer to the noise model proposed in the equation (3) then, if expert knowledge on the current profiles exist, we can minimize both the bias and the deviation terms due to N_i . The additional bias term is not included in the y-intercept estimation and is relative to the uneven distribution of $N_i - \bar{N}_i$ while the deviation terms have a higher amplitude as the standard deviation (if evenly distributed) of N_i is high.

1) *Mean of current predictable*: In such a case we can adequately generate Ψ_l by reducing both the deviation and the bias terms. If there exists a lot of current modes in the profile operation of the battery, we may also suppress the rotational bias.

2) *Mean of current unpredictable*: When no information about the mean of the current is known (i.e. when the battery application is kind of stochastic) then no prediction of the best sampling technique of Ψ_l can be made. This situation can be improved using specific algorithm techniques but this problematics is not discussed in this paper.

B. Intercept reduction with Hough-transform

The proposed algorithm for intercept reduction is inspired by the Adaptive Hough Transform described in [8]. The Hough-Transform (HT) is a method that finds lines in a noisy dataset and potentially corrupted with aberrant samples. The core idea of this method is to express the cartesian parameterization of an affine function in its parameters space. A noisy line is a flatten peak in this dual space. Finding the maxima of this peak gives an estimation of the intercept b_l^{th} . Here, instead of seeking among the entire set of possible solutions with a constant resolution, we recursively update, center and scale this set and refine the resolution of search.

The centering and scaling are performed thanks to a connected label components algorithm in order to work on the peak region that contains the future global maxima. The space of parameters must be initially large, but in our case we can use the capacity prediction and its standard deviation to reduce its size. The consistency of the estimator is demonstrated for unimodal and symmetrical inlier distribution in [9]. The theoretical replacement breakpoint converges to $1/2$ on the sole condition that the resolution of the HT is well calibrated and if an upper bound on the outlier ratio is known and largely below 50% when M is a small value.

Here we use this algorithm to have an estimation of the y-intercept. This algorithm is not sufficient for our capacity estimation because the Hough-Transform is more a detector than an estimator and thus the HT precision is not high enough for the capacity estimation requirements, especially since the HT is not a time adaptive and recursive algorithm contrary to the Kalman filter. However if the y-intercept term is highly above the standard-deviation of the measurement noise then it is necessary to “detect” it. The intercept obtained from the HT method is unbiased, and provides a significant reduction of the Kalman filter observation bias. The resolution and precision values of the estimation are used to get a confidence interval.

At convergence, i.e. at sufficient resolution, the maxima must be equal or up to $(1-\alpha)M \times r_{in}$ at $\frac{\alpha}{2}$ -th 100-quantile of the error variable where r_{in} is the inlier ratio, M the number of samples and α the risk level ($\alpha \ll 1$). If this not the case then the scatter should not be a line and the estimation must be postponed.

C. Capacity estimation by one-point RANSAC and UKF

We insert the state model of capacity evolution in an Unscented Kalman Filter (UKF) after having determined the y-intercept to perform successively the consensus and the capacity update on the inlier set. The Unscented Kalman Filter was proposed in the 2000s by Julier et al. in [10][11] and this filter is an extension of the classical Kalman to non-linear models. Contrary to the Extended Kalman filter, the Unscented Kalman Filter works with points (called sigma points) around the mean state to empirically estimate the mean and covariance propagation through the non-linear functions. These mean and covariance estimations are obtained using the scaled unscented transform (UT). The ability of the UT to approximate the propagation of the two first moments through a non-linear model is actually better than the first-order derivative approximation made in the EKF [12]. Moreover, for small state dimensions, the UKF complexity is about the same than the EKF one.

The set of inliers is obtained from a consensus approach called one-point RANSAC proposed by Civera et al. in the case of visual odometry problematic [13]. This technique is a derivative of the RANSAC algorithm (Random Sample Consensus) previously introduced in 1981 by Fischler and Bolles [14], which is a powerful technique for regression problem in presence of outlier observation samples. In our case, estimating the capacity is equivalent to obtain the slope of the linear and local relationships between the delta of

current and the variations of SOC. In other words, we have to find the affine function that fits the best the samples of Ψ_l^{in} . An affine function is entirely parametrized by its slope and its y-intercept and thus, taking two samples is enough to characterize it. RANSAC algorithm selects these two random points and test (exhaustively or at a determined confidence bound) each possible line. The line that fits the best the scatter is then chosen. The one-point RANSAC realizes the same procedure but, as the capacity prediction is known thanks to the Kalman recursion, then selecting one point is enough to draw the line and perform the consensus. Contrarily to the RANSAC, the one-point RANSAC is recursive and the theoretical number of hypotheses is largely smaller than the one needed in RANSAC algorithm for a given confidence bound of inliers detection. The procedure is intuitive and the Figure.2 shows some of the one-point RANSAC iterations :

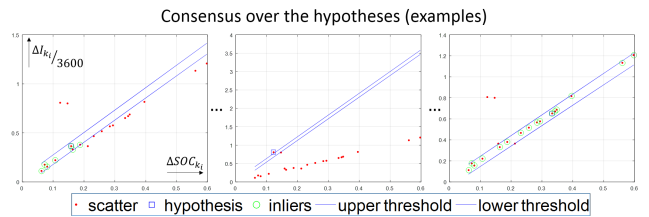


Fig. 2. one-point RANSAC procedure

It should be noticed that the knowledge of b_l improves the observation update resolution. It would have been less precise to estimate the y-intercept with only one of the hypothesis realisation. Estimating this intercept in a first part is primordial. Once a sufficient set of inliers is found the “real” update of Kalman filter over the entire set Ψ_l^{in} is performed and the estimated capacity is then obtained.

D. Retroaction using a hypothesis test

Here the SOC estimator is a recursive filter that attempts to minimize a given criteria, for example if the SOC estimator is a Kalman-based one then the objective function that is minimized is the mean-squared error of the estimated state given all the previous and instantaneous observations. Offline simulations can characterize the upper limit on capacity error (above ϵ_C) up to which the SOC estimation is unbiased or stationary biased. The case “stationary biased” is enough in our case because it leads to a SOC variation unbiased. From electrochemical considerations we can draw up a coarse but sufficient maximal variation of theoretical capacity and consequently we have an order of magnitude about the minimal delay needed between two capacity retroactions in the SOC estimator[7].

Of course, for a given capacity precision, the faster we update the capacity the most adaptive will be the SOC estimator. But this capacity must not be injected back too frequently to avoid instabilities linked to models switching. A classical scheme is to analyse the normalized innovation squared terms. Under gaussian assumptions, these terms are χ^2 -squared distributed and a hypothesis test can be performed

to decide whether the updated capacity feedback would be efficient to improve the accuracy.

IV. RESULTS

Our algorithm has been validated on the DashLink dataset provided by NASA in open access. They realized for this dataset deep discharges at constant rate and at ambient temperature on four different 18650 Li-Ion cells. The Open Circuit Voltage (OCV) is not delivered with the dataset, thus we had to perform an estimation of this using the impedance data and the voltage measurements. It has to be noted that this estimated OCV is not precise. Moreover the capacitance of the first order Warburg impedance of the cells is also not delivered and we estimated it at beginning of life of the cell using an unique measure of the time constant of a step response. The constant rate of current (constant mean) is the most difficult case because, as we have seen in the previous part, in such a configuration we cannot reduce the variance and potential rotation bias around the theoretical line. At time we do not dispose of an electrochemical model of the capacity ageing relative to the temperature. Thus the state model f is a coarse arithmetic depletion of capacity relative to a feature image of the number of battery cycles, and this model is averaged over the four cells of the NASA to keep generalization and avoid learning dependency. All this to say the presented results should be largely improved with more detailed initial battery cell characterizations. The SOC cannot be used as a feature because no dependency between state and measurements is allowed. We have chosen as feature the sum of squared temperatures accumulated between two capacity. Such a choice makes sense using thermics considerations. Indeed, the Joule loss can be linked to the integral of the heat flow from the cell to external environment. This heat transfer is proportional to the sum in discrete time of cell temperature when the ambient temperature is constant. Therefore, as the Joule loss is an image of electrical power exchanged, the sum of temperature between two updates can define a cycle related variable. It should be noticed that the use of squared temperature is a way to deal with both negative and positive temperatures. Absolute function could be used but its non derivability at 0 may be problematic in the general case. So the squared temperature is kept as a feature. We write below the model of capacity depletion where c is a normalization constant, T the average temperature of the cell, ω the slope of our model and L the total number of samples :

$$C_l = C_{l-1} - \omega T_e \frac{\sum_{j=1}^L T_j^2}{c}$$

This model is a first order approximation of capacity depletion relative to an “average” of the squared temperatures. Such a modeling is obviously biased and heavily unprecise. In addition to that no calendar ageing is taken into account in the capacity evolution here. In the future this model should be improved. But for our first validation this is considered sufficient, and it will shows the interest of combining a state model and a Kalman updating stage. To propose a statistical

sensitive-study of our estimator we repeat 50 conditionally independent realisations of the entire ageing set given the load profiles (around 550 hours of real profile repeated in simulation 50 times over the four cells) where the scatter Ψ is a new one at each temporal and statistical realisation. Even if the current profiles are always the same along the ageing, repeating the same profile with a different scatter at each step performs a sensitivity analysis.

TABLE I
STATISTICS OF THE PROPOSED CAPACITY ESTIMATOR

Cell's reference		B0005	B0006	B0007	B0018
$\bar{\mu}(\hat{C} - C)$	A.h	0.01	0.056	-0.0070	0.020
$\bar{\sigma}(\hat{C} - C)$	A.h	0.0068	0.014	0.0075	0.0085
$\bar{\mu}[\frac{\hat{C}-C}{C}]$	%	0.72	3.8	-0.086	1.4
$\overline{RMS}(\hat{C} - C)$	A.h	0.032	0.066	0.029	0.039
$\overline{RMS}[\frac{\hat{C}-C}{C}]$	%	2.0	4.5	1.7	2.7
$\overline{P}[\mu(\hat{C} - C) \leq 3\sigma]$	%	99.2	99.3	99.1	99.1

On Figures. 3-4, have been plotted two results of battery cell capacity estimation over time. The true cell capacity evolution is plotted in circle blue, whereas the estimated capacity is plotted in red solid line (in yellow and purple the 3σ bounds). The cyan blue solid line represents the capacity pure prediction using only the capacity evolution model.

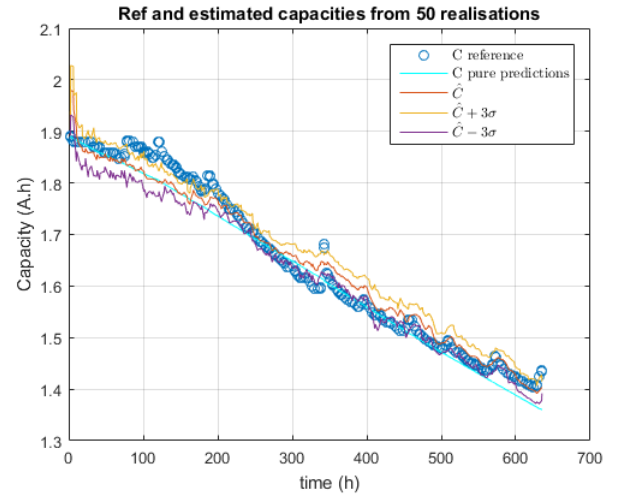


Fig. 3. statistical results on the B0007 cell

We first observe that the pure prediction of capacities (open-loop, with out the Kalman filter update) can be highly biased in long-term. This can be viewed for the cell B0018 in Figure.4. And this example confirms the need for this Kalman filter correction and its recursive scheme. In the case of B0007 cell results are less demonstrative because the predicted capacities obtained in open-loop is fortunately good. Even if the state equation is long-term biased, the estimated capacity follows the reference trajectory. Moreover we can observe the algorithm ability to converge rapidly towards abrupt capacity

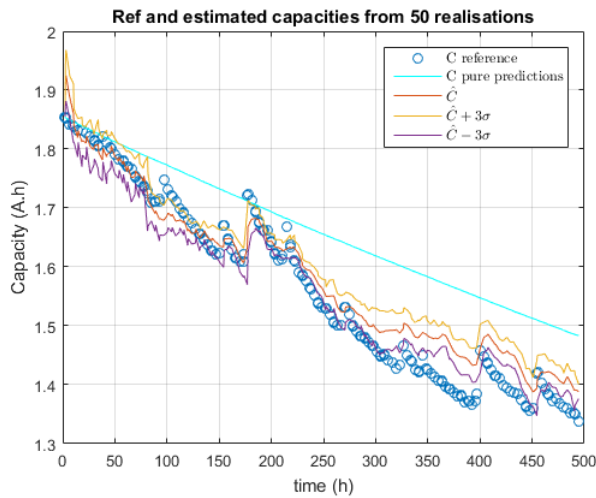


Fig. 4. statistical results on the B0018 cell

change. These two examples (repeated 50 times along the entire ageing cycles) represent statistical confirmations of the adaptability and the high level of accuracy obtained with our algorithm. As it was explained in the introduction the other great advantage of the proposed method is its robustness. Indeed all these tests and its results exhibit this robustness : the probability that the centered error is below 3σ is better than 99%. The histograms are not shown in this paper but the estimator distributions are quite symmetrical and their kurtosis is near the one of gaussian distribution.

V. CONCLUSIONS

The capacity estimator proposed in this paper was developed in order to improve both the robustness and the precision of the estimation during all the life cycle of an electrical battery. The robustness criteria is the most important because the capacity estimation is used in adjacent algorithms of the battery management system, and we know that outliers can corrupt the input data. The combination of bias reduction and outlier rejection method makes the algorithm robust. Moreover the virtuous insertion of these preprocessing algorithms in a Kalman framework make the global algorithm robust and adaptive. The obtained results are a statistical confirmation of this robustness and precision. Most of the RMS of the relative errors are below 3%. Future works should develop the mathematical concepts and study the algorithm complexity. Improving the data precision and working with more precise models of capacity evolution (get from the electrochemical works as an example) should also highly improve the performance of this algorithm.

ACKNOWLEDGMENT

We thank NASA and Dawn McIntosh, contributor of the online collaborative platform DashLink, for providing this essential data set.

REFERENCES

- [1] J. Labbe and G. Daw, "Panorama 2011 du marché du Lithium," BRGM, Rapport public RP-61340-FR, Jul. 2011.
- [2] G. L. Plett, "Recursive approximate weighted total least squares estimation of battery cell total capacity," *Journal of Power Sources*, vol. 196, no. 4, pp. 2319–2331, Feb. 2011.
- [3] A. A. Hussein, "Capacity Fade Estimation in Electric Vehicle Li-Ion Batteries Using Artificial Neural Networks," *IEEE Transactions on Industry Applications*, vol. 51, no. 3, pp. 2321–2330, May 2015.
- [4] C. Weng, Y. Cui, J. Sun, and H. Peng, "On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression," *Journal of Power Sources*, vol. 235, pp. 36–44, Aug. 2013.
- [5] M. Ecker, J. B. Gerschler, J. Vogel, S. Käbitz, F. Hust, P. Dechent, and D. U. Sauer, "Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data," *Journal of Power Sources*, vol. 215, pp. 248–257, Oct. 2012.
- [6] A. Barré, F. Suard, M. Gérard, M. Montaru, and D. Riu, "Statistical analysis for understanding and predicting battery degradations in real-life electric vehicle use," *Journal of Power Sources*, vol. 245, pp. 846–856, Jan. 2014.
- [7] Z. Chen, *Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond*. Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada: Natural Sciences and Engineering Research Council of Canada, 2003.
- [8] J. Illingworth and J. Kittler, "The Adaptive Hough Transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 690–698, Sep. 1987.
- [9] A. Goldenshluger and A. Zeevi, "The Hough transform estimator," *The Annals of Statistics*, vol. 32, no. 5, pp. 1908–1932, Oct. 2004, arXiv: math/0503668.
- [10] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of 1995 American Control Conference - ACC'95*, vol. 3, Jun. 1995, pp. 1628–1632 vol.3.
- [11] E. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*. Lake Louise, Alta., Canada: IEEE, 2000, pp. 153–158.
- [12] J. K. Uhlmann, "Dynamic Map Building and Localization_new Theoretical Foundations," Ph.D. dissertation, Oxford, Oxford, England.
- [13] J. Civera, O. G. Grasa, A. J. Davison, and J. M. M. Montiel, "1-Point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry," *Journal of Field Robotics*, vol. 27, no. 5, pp. 609–631, May 2010.
- [14] M. A. Fischler and R. C. Bolles, "Random Sample Consensus : a paradigm for model fitting with applications to image analysis and automated cartography," *Graphics and Image Processing*, vol. 24, no. 6, 1981.