36th International Electric Vehicle Symposium and Exhibition (EVS36)
Sacramento, California, June 11-14, 2023

Performance Diagnostics in Photovoltaic-Lithium-Battery Installations using Inconsistent Field Data

Alan G. Li[1], Xia Zeng[2], Wouter Parys[2], Md Sazzad Hosen[2,3], Theodoros Kalogiannis[2], Matthias Preindl[1], Joeri van Mierlo[2], Maitane Berecibar[2]

[1]*Columbia University in the City of New York, 500 W. 120th St., New York, NY, 10027, USA*

[2]*Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium*

[3]*Md Sazzad Hosen (corresponding author), md.sazzad.hosen@vub.be*

Executive Summary

Performance diagnostics of batteries in solar-photovoltaic and battery systems are important, especially if using second-life electric vehicle batteries. Currently, the battery pack is often oversized and sub-optimally managed. This is partially due to the lack of high-quality data collection and processing on-site. Practical implementation of advanced diagnostics tools such as machine learning remains out of reach. Here a robust pulse-based diagnostics method is proposed for calculating the equivalent series resistance of the battery pack. It takes advantage of 'natural' pulse instances in the system. Feature selection of the voltage and current measurements is combined with simple regression techniques to obtain an estimate. Data from 5 real-world installations, split into 9 distinct datasets, is used to assess the method. While there was no opportunity to compare the estimates with the true resistance values, the consistency of the results suggests that our method may be generally applicable.
Keywords: battery ageing, battery SoH, lithium battery, photovoltaic, pulse power

1 Introduction

Residential solar photovoltaic (PV) and lithium-ion battery (LIB) installations, exemplified in Fig. 1, are increasingly popular for consumers aiming to reducing reliance on the electrical grid [1]. PV-battery systems may increase energy security for the household, reduce electric utility bills, and reduce carbon emissions [2]. In remote locations, islanded PV-battery systems may provide electrical services critical for social development [3], where lead-acid batteries are often used. Much effort has been focused on optimisation of the battery pack, which not only ensures that the use of solar energy can be shifted away from the instance of availability, but also constitutes a significant portion of the total cost [4].

Many factors affect battery performance. The energy management system (EMS) controls the flow of energy between the PV panels, the battery, and the grid. Different EMS strategies lead to different rates of degradation, thus affecting the useful pack lifetime [5]. Knowing the load dynamics is important for EMS development because they govern the specific degradation trajectory [6]. Knowledge of the load behaviour can also be used to optimally size the battery packs, which decreases the total costs and reduces wasted capacity [7].

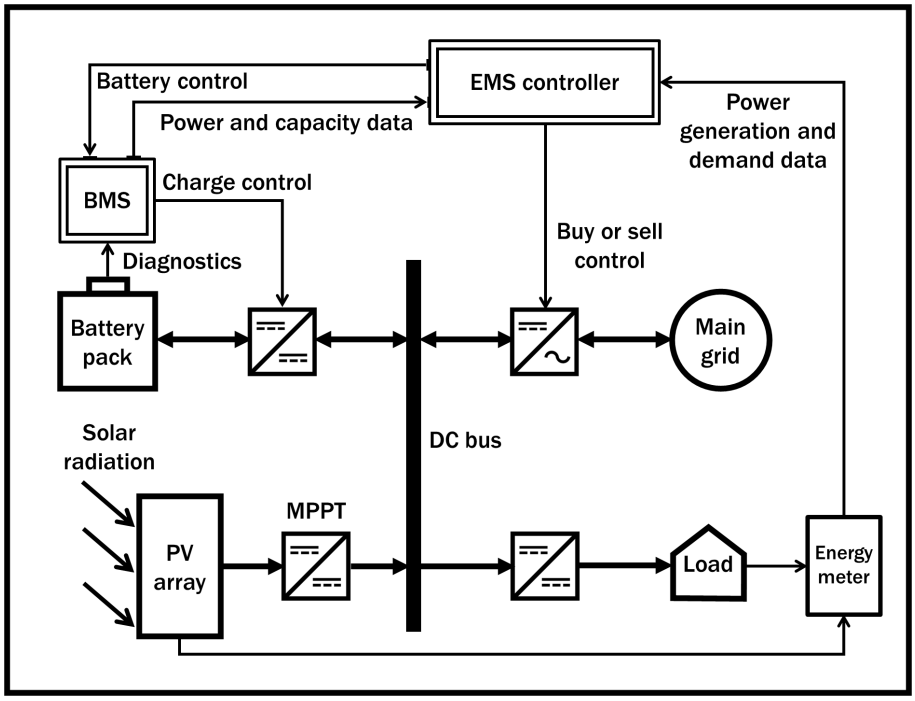


Figure 1: Diagram of PV-battery microgrid, reproduced from [1]

Better state of health (SoH), state of charge (SoC), and state of power (SoP) estimation in the battery management system (BMS) are crucial for optimising pack usage [8]. By reducing the error bounds, there is more confidence to use the pack to its full capabilities [9]. In real-world installations, however, SoC estimation is at best rudimentary, while degradation modelling remains an active research area [10, 11]. Degradation models are often constructed from simplified physical principles [4, 12] or equivalent-cycles [13]. While many models have been proposed, they are not usually implemented in real systems due to highly variable data, limited computational power, or the inability to characterise the batteries before use [14, 15].

While knowledge of the actual degradation rate of the cells would be useful, current installations are limited to methods that are robust to inconsistent data inputs and computationally lean. Most studies use simulations and laboratory tests, which are not available in the real world. Data-driven machine-learning techniques have thus gained popularity for their ability to use real-world voltage, current, and temperature data to assess battery health and performance [3, 16, 17].

Results from [3], in particular, directly address the challenges in estimating battery degradation using real-world data from PV and lead-acid battery systems. It is noted that the incomplete cycling, inconsistent data quality, and lack of controlled tests hinder the ability to calculate the battery SoC or SoH [18]. Thus the equivalent series resistance R_0 was identified as a health metric. Using Gaussian process regression and laboratory open-circuit voltage (OCV) testing, R_0 measurements are obtained. It is shown that R_0 can predict the end-of-life with up to 82% accuracy at the point of failure. The challenges to R_0 estimation are linked with its dependence on the instantaneous operating points of temperature, current, and SoC, as well as slow-changing degradation level [19]. Additionally, there is considerable uncertainty in extending the behaviours of laboratory-characterised cells to a population of real-world systems. Finally, it is unclear if results from lead-acid cells can extend to lithium-ion systems.

Here we examine the challenges and opportunities of using pack-level data from real grid-connected PV-LIB installations to develop a health and performance estimation technique from R_0 measurements. We investigate the challenges and opportunities of using real-world data instead of laboratory techniques and propose a simple method that can capture the effects of resistance change in the system. In contrast to previous studies, our diagnostics methods assume no knowledge of the lithium-ion chemistry or pack configuration.

Our study continues to Section 2 where we discuss the datasets. Methods are described in Section 3. Results are presented and discussed in Section 4. The report is concluded in Section 5.

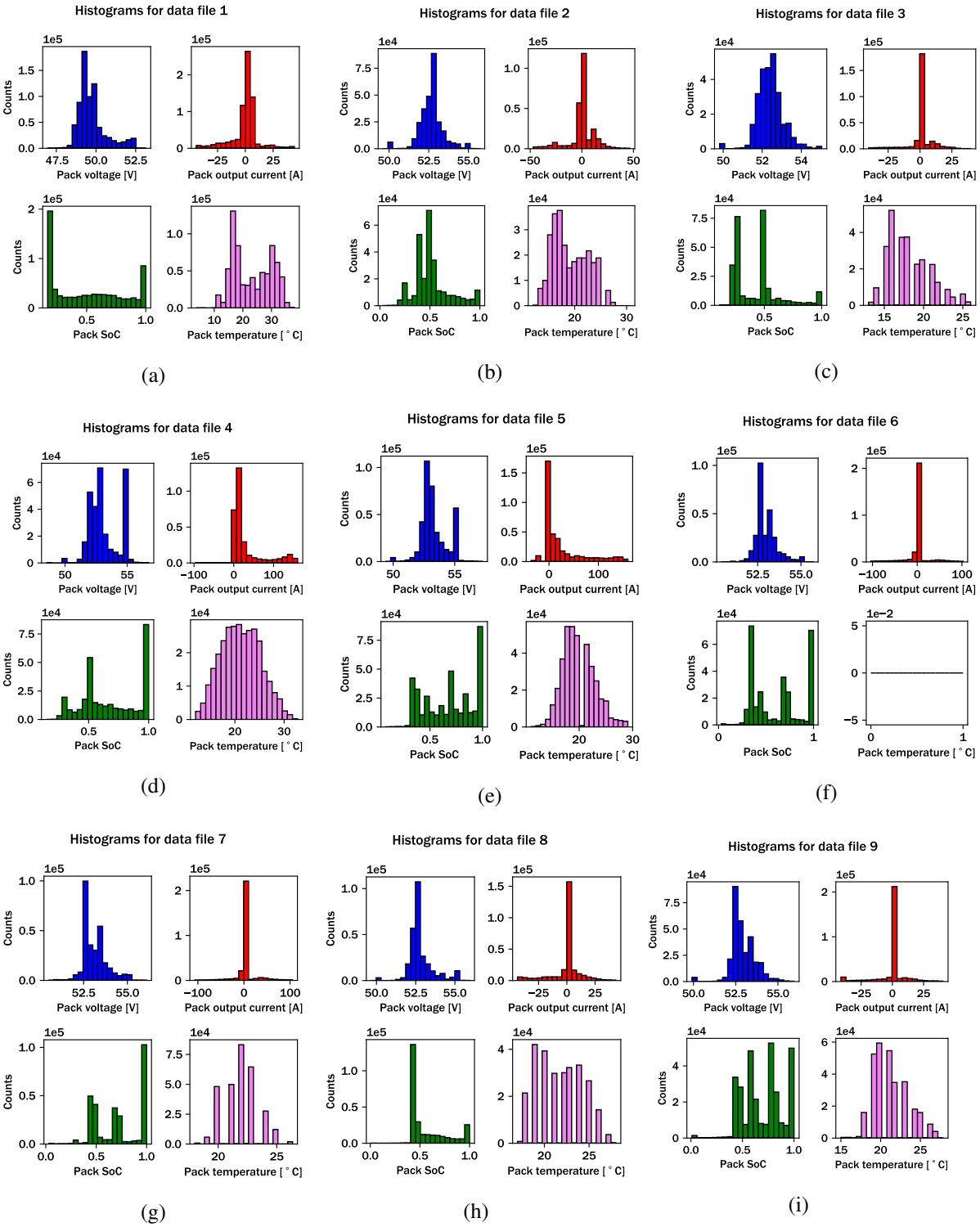


Figure 2: Histograms of the voltage, current, SoC, and temperature for all data files

Table 1: Names and locations of PV-battery installations from Victron Energy

Dataset #	Name	Location	Acquisition period
1	Carmen Pepe	Spain	Apr 2021-22
2,3	Hannes Holm	South Africa	Jun-Dec 2021, Apr-Oct 2022
4,5	House Holm	South Africa	Jun-Dec 2021, Apr-Oct 2022
6,7	Ian Bell	South Africa	Jun-Dec 2021, Apr-Oct 2022
8,9	Ian Hamilton	South Africa	Jun-Dec 2021, Apr-Oct 2022

Table 2: PylonTech US3000C battery pack specifications

Characteristic	Value	Units
Chemistry	LFP	—
Nominal voltage	48	V
Max charge voltage	53.5	V
Min discharge voltage	44.5	V
Nominal capacity	3.552	kWh
Usable capacity	3.374	kWh
String size	16	cells

2 Data Collection

Data are downloaded from the Victron Energy world database [20]. Key data includes the pack voltage, current, temperature, and average pack SoC. Data on individual cells is not available. We list all the installations in Table 1. The number refers to the data file, which is split into two non-contiguous parts for the last four locations. In some cases, the pack serial number is known, such as location 1 (Carmen Pepe, Spain), which used a PylonTech US3000C battery pack with the specifications shown in Table 2. Though laboratory testing of this pack was not available, this would greatly aid in characterisation. For data files 2-9, the datasets are incomplete and are missing 4 months from December to April. They are located in the southern hemisphere, South Africa.

Features of all datasets are shown as histograms in Fig. 2, describing how often certain values of the voltage, current, temperature, and SoC are measured. Outliers are removed, defined as any value exceeding 3 standard deviations from the mean of the distribution. The SoC is assumed to be an approximate pack-averaged value estimated on-site and is defined as

$$z'(t) = \frac{q(t)}{Q_0} \tag{1}$$

where Q_0 [Ah] is the initial nominal maximum pack capacity and $q(t)$ [Ah] is the estimated remaining capacity at time t. This is not the usual definition of SoC because Q_0 does not capture the effects of degradation. Since this aggregated metric treats the pack as a single cell, there is typically unused cell capacity which may even exacerbate degradation [15].

From Fig. 2 it can be seen that the datasets have highly distinct trends. In most cases, the voltage remains close to 50 or 52 V; current is commonly very close to 0; and the temperature is around 25°C. The SoC measurements, however, show large variation and certain datasets have more frequent high-current discharges, such as 4 and 5. We can also see that dataset 6 had no temperature measurements. Interestingly, even for data files from the same location there is considerable variation, particularly in voltage and SoC. Since the current distributions remain approximately unchanged, this could reflect different operating conditions rather than a change in loading.

Specific features of the data are exemplified in Fig. 3 for location 1, with the minimum, average, and maximum daily profiles for voltage, current, temperature, and SoC. Note that current refers to the pack output current and the minimum and maximum are determined by the number of coulombs passed. The daily trends of battery charging around solar noon and the consequent rise in SoC and temperature can be clearly observed.

The number of data samples (directly proportional to the sampling frequency) recorded for each day and location are shown in Fig. 4a. It can be seen that the data resolution is inconsistent and data file 1 even features a dramatic increase in sampling frequency. For consistency in data processing, the data

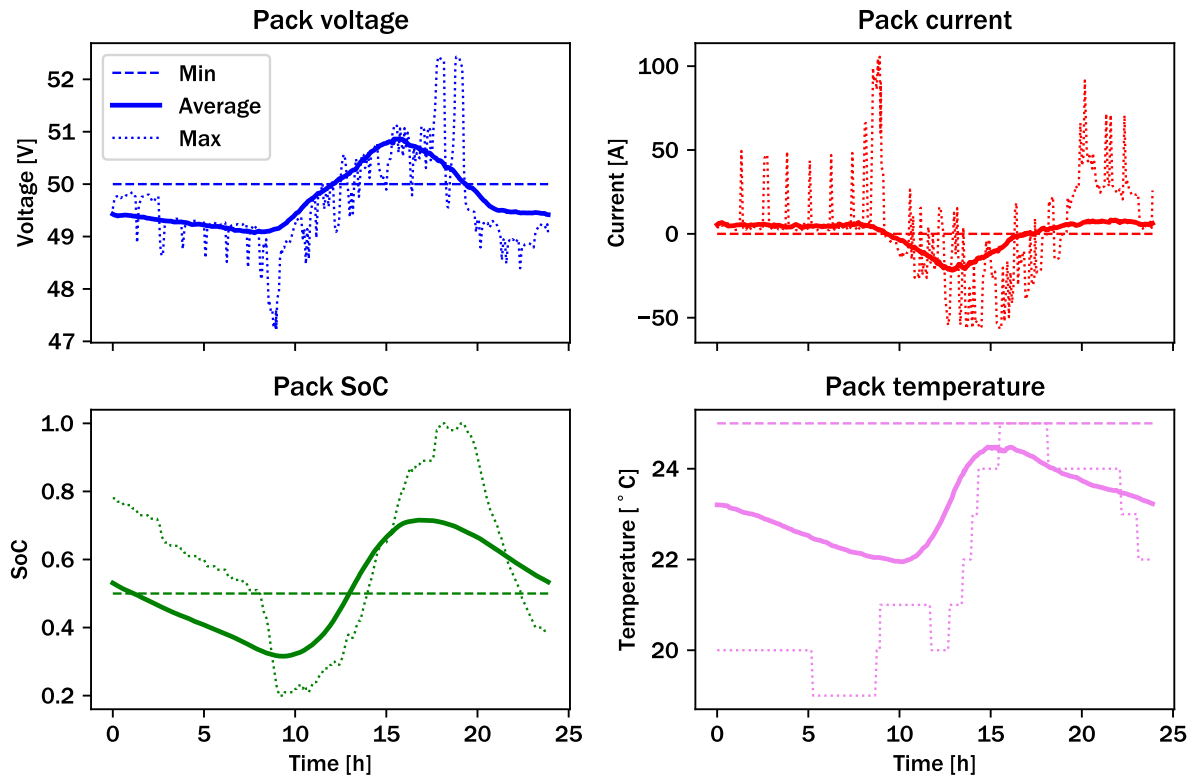
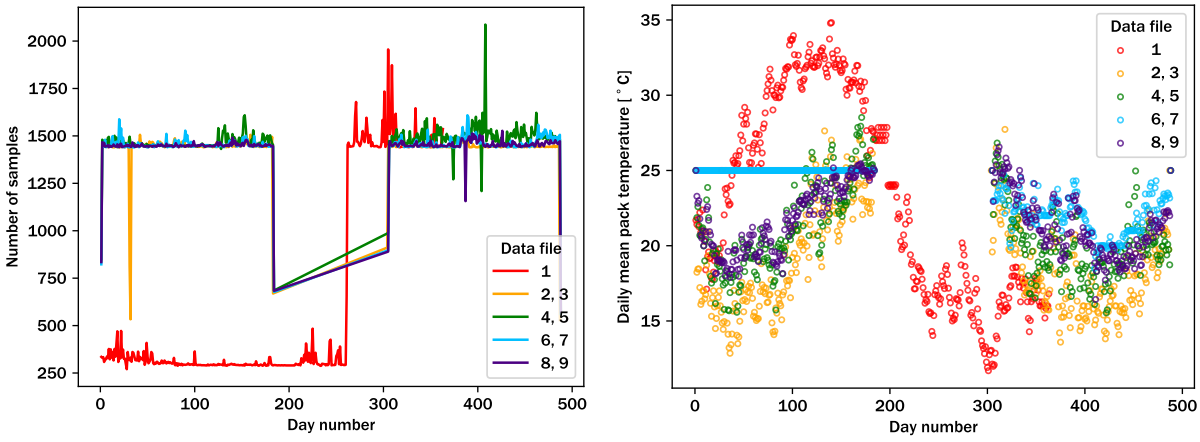


Figure 3: Voltage, current, temperature, and SoC 24-hour profiles for dataset 1, showing the average daily curve and the profiles corresponding to the days with minimum and maximum amounts of coulombs passed



(a) (b)

Figure 4: Plots of day-varying parameters in the data files, showing (a) Number of data points recorded per day and (b) Mean daily pack temperature

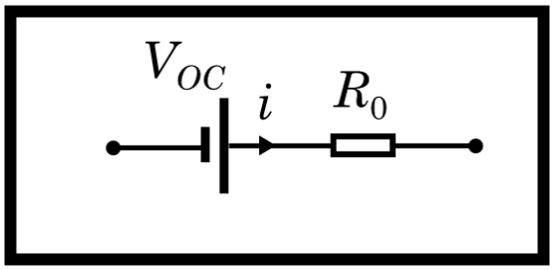


Figure 5: Diagram of 0-order equivalent circuit model

are therefore linearly interpolated to the maximum recorded value such that the number of samples per day is identical across each dataset. Seasonal temperature variation is observed in Fig. 4b. There is considerable variation for each day but the warmer summer temperatures in dataset 1 can be observed from days 50–200 and cooler spring temperatures for the other datasets.

3 Methods

The health and performance metrics most important for operation are the SoH and impedance, often described as SoP. We focus on R_0 estimation, similar to [3].

The equivalent series resistance R_0 is represented in Fig. 5. Since the minimum sampling interval was 44 seconds (sampling frequency of magnitude 10^{-2} Hz) many of the transient overpotentials dynamics are unobservable [21, 22]. Ohmic and charge transfer overpotentials, in particular, reach steady-state within this sampling interval. Even if diffusion may still be observable at 10^{-2} Hz, the Nyquist limit reduces the observable frequencies. Thus the 0-order equivalent circuit model (ECM) may be used. Furthermore, the R_0 measurement represents an aggregated cell impedance. The terminal voltage is therefore modelled as

$$V(t) = V_{OC}(z,t) - R_0 i(t) \tag{2}$$

where i [A] is the output current from the cell, V_{OC} is the OCV, a function of the state of charge z. If we take the derivative with respect to the time-varying current, we can write R_0 as the gradient

$$R_0 = \frac{\partial V_{OC}(z,t)}{\partial i(t)} - \frac{\partial V(t)}{\partial i(t)} \tag{3}$$

Assuming the condition $\frac{\partial V_{OC}(z,t)}{\partial i(t)} \to 0$, i.e., assuming that the OCV is constant relative to the current, we can write

$$R_0 = \left| \frac{\partial V(t)}{\partial i(t)} \right| \tag{4}$$

Since the terminal voltage is also a function of OCV, for this approximation to be valid we require the terminal voltage to have greater dependence on the current than on the SoC. We can see that these conditions do not hold when the OCV or hysteresis effects are the primary cause of voltage change.

To calculate R_0, voltage and current data are fitted to Equation 4. This is performed with careful selection of voltage-current data and least-squares linear regression. The algorithm is enumerated below:

1. Select window length N.
2. For each day number n_{day}, create a pool of potential start times t_k of the regression window with the criteria

$$\begin{cases} i(t_k) = 0 \\ i(t_{k+1}) - i(t_k) > 3 \quad \text{[A]} \end{cases} \tag{5}$$

   This aims to reduce the influence of OCV and ensures that there is sufficient current variation to obtain a good fit. As can be seen from Fig. 2, the current is most commonly at low values, meaning that the 3 A threshold allows for many candidate windows.

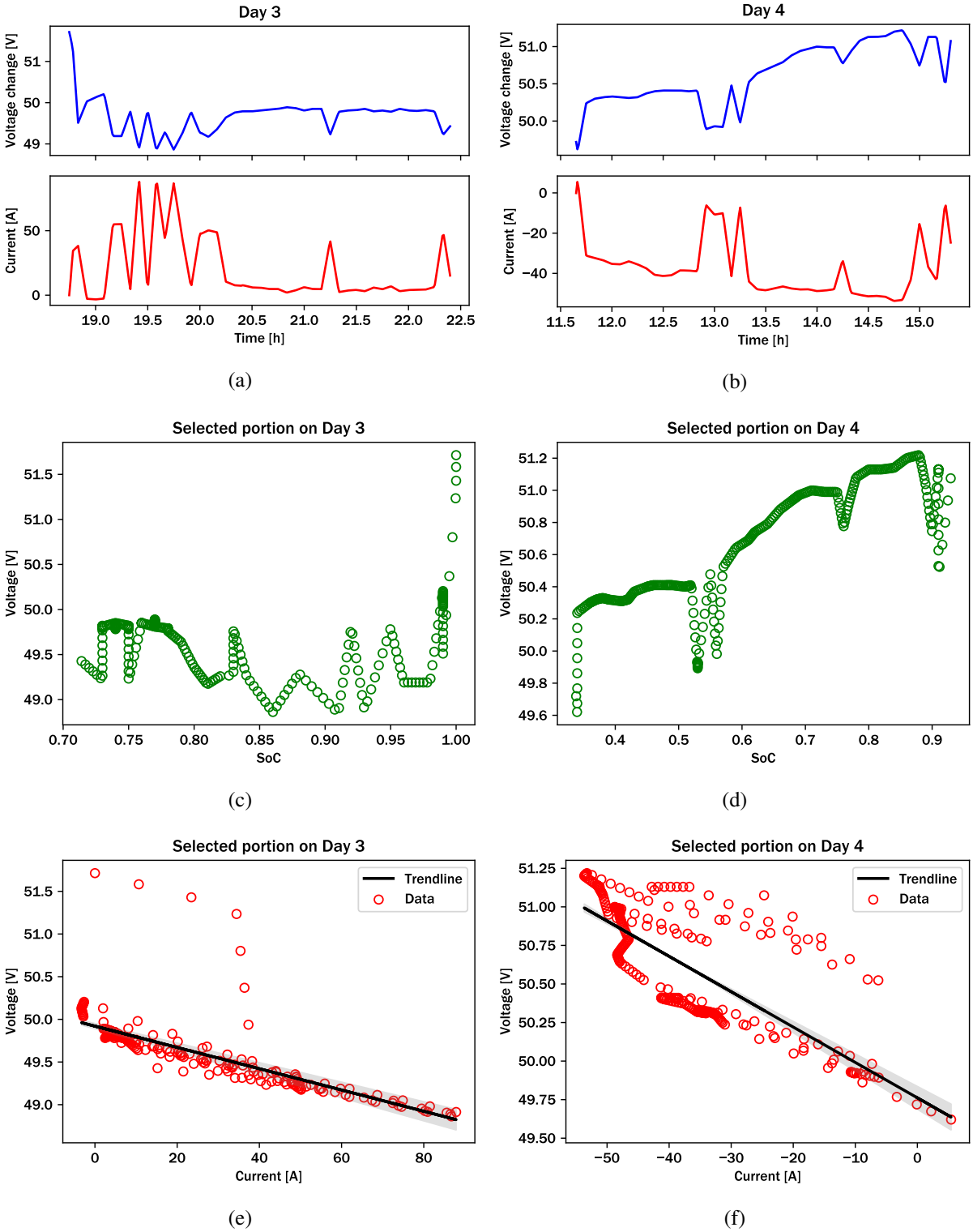


Figure 6: Regression window data from dataset 1, showing plots of (a)-(b) voltage and current over time, (c)-(d) voltage against SoC, and (e)-(f) voltage against current, demonstrating low standard error (left column) and high standard error (right column)

3. For each candidate data window, calculate the least-squares feature vector using

$$w_k = (X_k^T X_k)^{-1} X_k^T y_k \tag{6}$$

where $X_k \in \mathbf{R}^{N\times 2}$ is given by

$$X_k = \begin{pmatrix} 1 & V(t_k) \\ 1 & V(t_{k+1}) \\ \vdots & \vdots \\ 1 & V(t_{k+N-1}) \end{pmatrix} \tag{7}$$

and $y_k \in \mathbf{R}^N$ is given by

$$y_k = \begin{pmatrix} i(t_k) & i(t_{k+1}) & \cdots & i(t_{k+N-1}) \end{pmatrix}^T \tag{8}$$

4. Find the value of k that yields the lowest standard error,

$$k_{\text{min}} = \arg\min_k \|y_k - X_k w_k\|_2 \tag{9}$$

5. The R_0 prediction for the day is given by the second element of $w_{k_{\text{min}}}$

$$\hat{R}_0(n_{\text{day}}) = w_{k_{\text{min}},2} \tag{10}$$

Sample regression windows are shown in Fig. 6. It is evident that the Day 3 portions (Figs. 6a, c, e) result in a much better fit than the Day 4 portions. As suggested by Equation **??**, this is because the average gradient of the voltage-SoC behaviour is close to 0. In contrast, the portions for Day 4 (Figs. 6b, d, f) demonstrate strong voltage-SoC variation leading to higher standard error. This suggests that the ohmic resistance is best diagnosed from instances of high-amplitude, high-frequency current pulses.

4 Results and Discussion

Results of resistance prediction using window length of $N = 300$ are shown in Fig. 7. From Fig. 7a it can be seen that the resistance is approximately constant or increasing for each data file. This agrees with expectations – degradation is known to decrease SoH and increase impedance [23].

Uncertainties in the prediction may be due to variations in SoC, SoH, temperature, current, and pack properties. These relationships are not known but are the result of complex dynamics between the LIB cells, pack configuration, BMS algorithm, and power electronics. As exemplified in Fig. 7b, the measurements are consistent, suggesting that the method is robust against data fluctuation. Further work is needed to use our results for predicting battery end-of-life.

Decreasing trends in the resistance measurement with respect to the day number, as detailed in Fig. 7c, may be explained by the coupling of effects between SoC, SoH, temperature, and current. All four variables are known to affect the battery impedance. This is due to the effects of ion transport and electrode dynamics [23]. Ion kinetics are captured by battery overpotentials. High currents may reduce the effects of diffusion and charge transfer, thus reducing the overall impedance. Higher temperature, meanwhile, can improve ion transport, thus reducing impedance. Impedance also decreases near mid-range SoC values because at the extremes either the positive or negative electrode are nearly full, hindering charge transfer [3]. Lower SoH reflects electrode degradation, which may result from structural change, dendrite growth, and solid-electrolyte-layer formation — all of which increase impedance. Finally, cell-to-cell variation can cause large variation in the overall pack impedance [9]. Since the R_0 measurement aggregates all of these effects, without additional processing it may not be an ideal indicator for battery degradation. Indeed, the reported values of R_0 may be more accurately described as the aggregated impedance Z over a 44-second time interval.

Our primary hyperparameter is the window length, chosen to be 300 based on the results shown in Figs. 7d-e. As can be seen, if the window is too small then the there is high error in the prediction. In contrast, if the window is too big then the number of available predictions drops considerably. This is because the number of unique windows is limited by the number of data points in a day. Thus $N = 300$ strikes

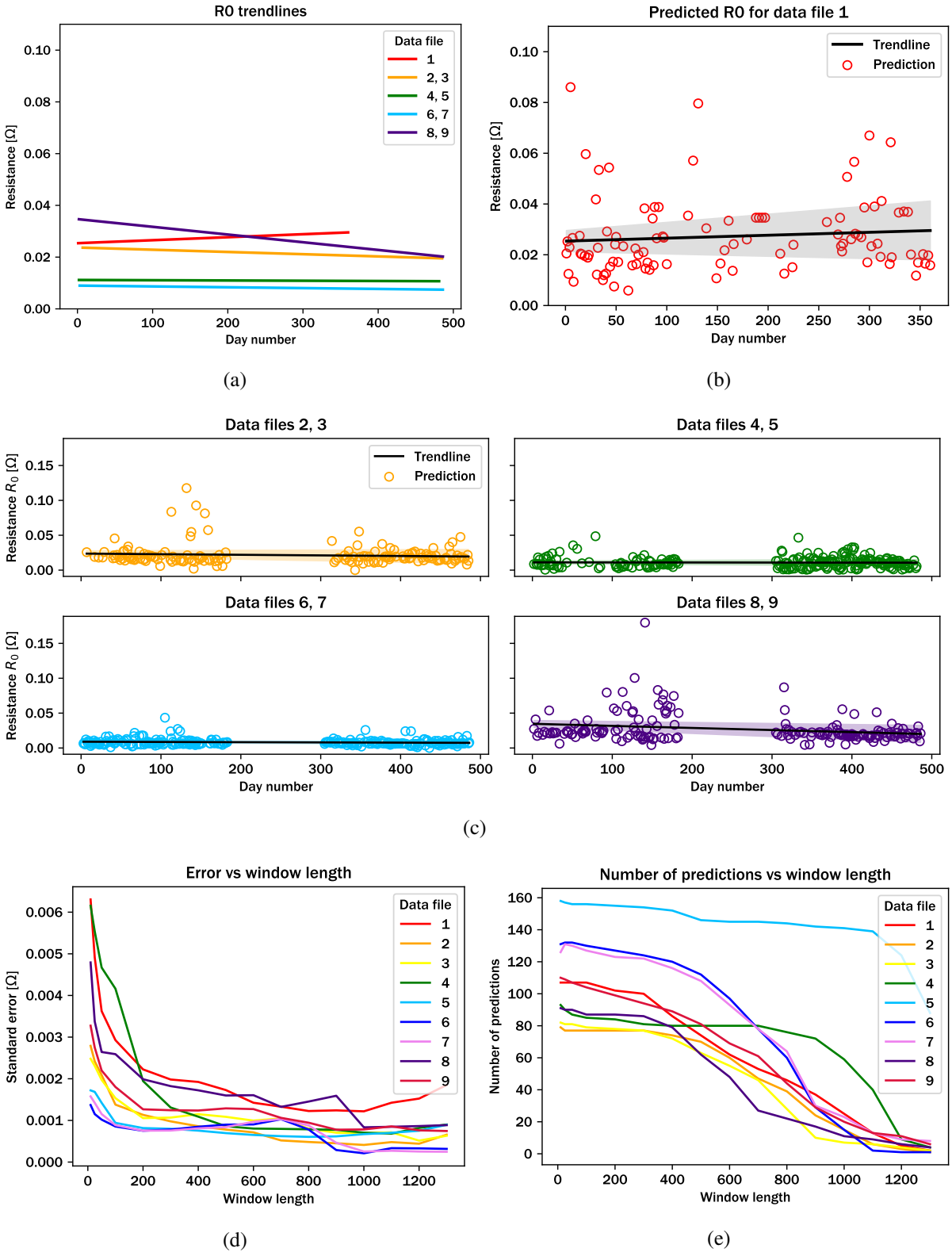


Figure 7: Resistance prediction results for each data file, showing (a) Overall trendlines, (b) Detail of predictions and trendline for dataset 1, (c) Details of predictions and trendlines for data files 2–9, (d) Standard prediction error against window length, and (e) Number of available predictions against window length

a balance between low error and a high number of predictions. This appears to be consistent across all datasets.

As implied above, the datasets present several limitations. We have no ability to perform experimental tests on the battery packs in the installations and thus no way to validate our conclusions using real observations. Obtaining datasets where there is a known time of failure, such as in the PV-lead-acid systems in [3], could provide some form of experimental validation, but any measurements between day 0 and the end-of-life will remain unverified. Gaps in the datasets and issues with data quality reduce the certainty in our conclusions. There are no cell-level measurements, so we cannot easily understand how the pack data translates to the cell level. The pack SoC is likely to have accumulated errors as well. Finally, the R_0 measurement could be affected by physical disturbances in the system, such as an impact to the wiring that would not necessarily affect the LIB cells but still alter the ohmic resistance.

Disaggregating the effects of operational parameters and pack configuration to achieve a better SoH indicator would greatly increase the utility of our work. As demonstrated in [3], if the general degradation trajectory curve is assumed to be known, then the effects of SoC, SoH, temperature, and current can be decoupled from the aggregated impedance. This would allow us to isolate the SoH-driven component and thus obtain a pseudo-SoH metric.

5 Conclusion

Real-world PV-grid systems present many challenges and opportunities for degradation diagnostics. Data inconsistency and inability to verify conclusions have made real-world datasets difficult to research. We have nevertheless shown that valuable trends may still be identified from first principles. In particular, we propose an algorithm for estimating the equivalent series resistance R_0. Our results are obtained from a 0-order ECM and ordinary least-squares estimation of carefully-designed data window selection process. Our method offers a promising way to understand the system degradation with minimal data collection.

Further research may include OCV-SoC reconstruction using instances where the OCV can be deduced with high confidence. This could also lead to incremental-capacity-inspired analysis for pack-level SoH estimation. With on-site access to a PV-battery installation and the ability to interrupt service for diagnostics, high-quality validation measurements can be obtained. This would ultimately lead to lower costs, higher cycling capacity, and greater adoption of PV power for the reduction of carbon emissions.

Acknowledgments

This project was supported by VLAIO with the ICON project MAMUET (grant number HBC.2018.0529). Alan G. Li was supported by a Fellowship of the Belgian American Educational Foundation.

References

[1] A. G. Li and M. Preindl, "Assessing degradation-aware model predictive control for energy management of a grid-connected pv-battery microgrid," *IEEE Transportation Electrification Conference (ITEC)*, 2022.

[2] M. Dubarry, N. C. Cortez, and D. Matthews, "Data-driven direct diagnosis of pv connected batteries," *PREPRINT*, 2022.

[3] A. Attio and D. A. Howey, "Predicting battery end of life from solar offgrid system field data using machine learning," *Joule*, vol. 5, pp. 3204–3220, 2021.

[4] M. F. Zia, E. Elbouchiki, and M. Benbouzid, "Optimal operational planning of scalable dc microgrid with demand response, islanding, and battery degradation cost considerations," *Appl. Energy*, vol. 237, pp. 695–707, 2019.

[5] S. Ouedraogo, G. A. Faggianelli, G. Pigelet, G. Notton, and J. L. Duchaud, "Performances of energy management strategies for a photovoltaic/battery microgrid considering battery degradation," *Solar Energy*, vol. 230, pp. 654–665, 2021.

[6] M. Dubarry, G. Baure, and A. Devie, "Durability and reliability of ev batteries under electric utility grid operations: Path dependence of battery degradation," *J. Electrochem. Soc.*, vol. 165, no. 5, A773–A783, 2018.

[7] Y. Wei, T. Han, S. Wang, Y. Qin, L. Lu, X. Han, and M. Ouyang, "An efficient data-driven optimal sizing framework for photovoltaics-battery-based electric vehicle charging microgrid," *J. Energy Storage*, vol. 55, no. 105670, 2022.

[8] A. G. Li, W. Wang, A. C. West, and M. Preindl, "Health and performance diagnostics in li-ion batteries with pulse-injection-aided machine learning," *Appl. Energy*, vol. 315, no. 119005, 2022.

[9] A. G. Li and M. Preindl, "State of charge imbalance classification of lithium-ion battery strings using pulse-injection-aided machine learning," *IEEE Transportation Electrification Conference (ITEC)*, 2022.

[10] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. V. Mierlo, and P. V. den Bossche, "Critical review of state of health estimation methods of li-ion batteries for real applications," *Renew. Sus. Energy. Rev.*, vol. 56, pp. 572–587, 2016.

[11] S. Wang, D. Guo, X. Han, L. Lu, K. Sun, W. Li, D. U. Sauer, and M. Ouyang, "Impact of battery degradation models on energy management of a grid-connected dc microgrid," *Energy*, vol. 207, no. 118228, 2020.

[12] M. Alramlawi and P. Li, "Design optimization of a residential pv-battery microgrid with a detailed battery lifetime estimation model," *IEEE Trans. Indus. Appl.*, vol. 56, no. 2, 2020.

[13] M. A. I. Martins, L. B. Rhode, and A. B. de Almeida, "A novel battery wear model for energy management in microgrids," *IEEE Access*, vol. 10, no. 118228, 2022.

[14] J. Reniers, G. Mulder, S. Ober-Blobaum, and D. A. Howey, "Improving optimal control of grid-connected lithium-ion batteries through more accurate battery and degradation modelling," *J. Power Sources*, vol. 379, pp. 91–102, 2018.

[15] J. M. Reniers, G. Mulder, and D. A. Howey, "Unlocking extra value from grid batteries using advanced models," *J. Power Sources*, vol. 487, no. 229355, 2021.

[16] M. S. Hosen, J. Jaguemot, J. V. Mierlo, and M. Berecibar, "Battery lifetime prediction and performance assessment of different modeling approaches," *iScience*, vol. 24, no. 2, 2021.

[17] S. Khaleghi, M. S. Hosen, D. Karimi, H. Behi, S. H. Beheshti, J. V. Mierlo, and M. Berecibar, "Developing an online data-driven approach for prognostics and health management of lithium-ion batteries," *Appl. Energy*, vol. 308, no. 118348, 2022.

[18] S. Zhang, M. S. Hosen, T. Kalogiannis, J. V. Mierlo, and M. Berecibar, "State of health estimation of lithium-ion batteries based on electrochemical impedance spectroscopy and backpropagation neural network," *World Electr. Veh. J.*, vol. 12, no. 3, 2021.

[19] J. de Hoog, J. Jaguemont, A. Nikolian, J. V. Mierlo, P. V. den Bossche, and N. Omar, "A combined thermo-electric resistance degradation model for nickel manganese cobalt oxide based lithium-ion cells," *Appl. Thermal Eng.*, vol. 135, pp. 54–65, 2018.

[20] Victron Energy. (2022). "World installations," [Online]. Available: `https://vrm.victronenergy.com/world/#`.

[21] A. G. Li, M. M. Wu, Y. A. Fahmy, and M. Preindl, "Fast time-domain impedance spectroscopy of lithium-ion batteries using pulse perturbation," *IEEE Transportation Electrification Conference (ITEC)*, 2022.

[22] A. G. Li, K. Mayilvahanan, A. C. West, and M. Preindl, "Discrete-time modeling of li-ion batteries with electrochemical overpotentials including diffusion," *J. Power Sources*, vol. 500, no. 229991, 2021.

[23] A. G. Li, A. C. West, and M. Preindl, "Towards unified machine learning characterization of lithium-ion battery degradation across multiple levels: A critical review," *Appl. Energy*, vol. 316, no. 119030, 2022.

Presenter Biography

Xia Zeng received her Masters degree in Vehicle Engineering from Chongqing University, China in 2019. She was a visiting research student with the Department of Electrical Engineering, Chalmers University of Technology, Sweden, from 2017 to 2018. In 2021, she started her Ph.D. in the Department of Electrical Engineering and Energy Technology (ETEC) at Vrije Universiteit Brussel, Belgium. Her research focuses on battery monitoring and lifetime estimation through advanced sensing techniques.