The Journal of Open Source Software

DiffCapAnalyzer: A Python Package for Quantitative Analysis of Total Differential Capacity Data

Nicole L. Thompson[1], **Theodore A. Cohen**[2, 3, 4], **Sarah Alamdari**[1], **Chih-Wei Hsu**[1], **Grant A. Williamson**[2], **David A. C. Beck**[1, 5], **and Vincent C. Holmberg**[1, 2, 6]

1 Dept. of Chemical Engineering, University of Washington **2** Molecular Engineering and Sciences Institute, University of Washington **3** Dept. of Materials Science and Engineering, University of Washington **4** Dept. of Chemistry, University of Washington **5** eScience Institute, University of Washington **6** Clean Energy Institute, University of Washington

DOI: 10.21105/joss.02624

Software

- Review
- Repository
- Archive

Editor: Jeff Gostick

Reviewers:

- @yangbai90
- @WardLT

Submitted: 11 July 2020
Published: 18 October 2020

Summary

In order to study long-term degradation and charge storage mechanisms in batteries, researchers often cycle these electrochemical cells for hundreds or even thousands of charge and discharge cycles. The raw data produced during cycling can be interpreted via a variety of techniques that each highlight specific aspects of how the battery is functioning. Differential capacity (dQ/dV) analysis, one such technique, results in plots of the differential capacity – the charge introduced into the battery during a small change in voltage – *vs.* the voltage. Electrochemical reactions result in significant charge introduced into the cell across a small voltage window. In the differential capacity plot, this behavior results in a peak for each electrochemical reaction. Therefore, differential capacity plots are particularly useful for highlighting the various electrochemical events occurring within the cell, specific to each cycle (Aihara et al., 2016; Christophersen et al., 2006; Christophersen & Shaw, 2010; Marzocca & Atwater, n.d.; Torai, Nakagomi, Yoshitake, Yamaguchi, & Oyama, 2016; Weng, Cui, Sun, & Peng, 2013). In turn, these peaks carry important characteristics of the electrochemical reaction. For example, the location of the peak indicates at what voltage the reaction occurs, and the area of the peak is linked to the amount of charge exchanged in the reaction.

We present DiffCapAnalyzer, a Python package for extracting and tracking differential capacity curve features through multiple charge and discharge cycles. DiffCapAnalyzer provides cleaned dQ/dV curves, peak locations, peak heights, peak areas, and other characteristics specific to each cycle from raw battery cycling data.

Statement of Need

Traditionally, when using differential capacity plots, researchers have drawn conclusions based on an arbitrarily chosen subset of cycles and reported mainly qualitative claims on how peaks shift during cycling, due to the difficulties in analyzing the full amount of data produced in the differential capacity plots. Additionally, although it is known that peak shapes and areas correlate to important electrochemical events, only a few papers report using peak deconvolution as a method to interpret dQ/dV plots (Aihara et al., 2016; Bian, Liu, & Yan, 2019; He, Bian, Liu, Wei, & Yan, 2020; Huang, 2019; Torai et al., 2016). Further, there does not exist any standardized method for peak deconvolution of differential capacity plots. These issues can largely be attributed to the lack of software designed for investigating sets of dQ/dV curves. Prior to DiffCapAnalyzer, no open source software has been available to researchers for the analysis of differential capacity curves through peak deconvolution.

Description

The software described herein, DiffCapAnalyzer, has been developed to address the drawbacks associated with differential capacity analysis by processing cycling data in a chemistry-agnostic manner. This is done by calculating differential capacity from the given raw cycling data using Equation 1, cleaning and smoothing the dQ/dV plots, and performing automatic peak locating and deconvolution for every cycle within the dataset.

$$(dQ/dV)_i = (Q_i - Q_{(i-1)})/(V_i - V_{(i-1)}) \tag{1}$$

In differential capacity curves without any cleaning or smoothing, there is significant noise and large step-wise changes present. This is a common problem when the denominator of Equation 1 approaches zero (Bloom et al., 2005; Huang, 2019). Therefore, in order to accurately identify peaks, the data is cleaned by removing points such that the voltage difference between datapoints is at least 0.001 V. Subsequently, the curve is smoothed using a Savitzky-Golay filter, which is a moving polynomial of specified order fit over a specified number of data points. At the current state of the software, the polynomial order of the Savitzky-Golay filter is set at 3 with a window length of 9 data points, as these seemed the best parameters on the data tested to preserve important features while removing noise. This cleaning process is summarized in Figure 1:

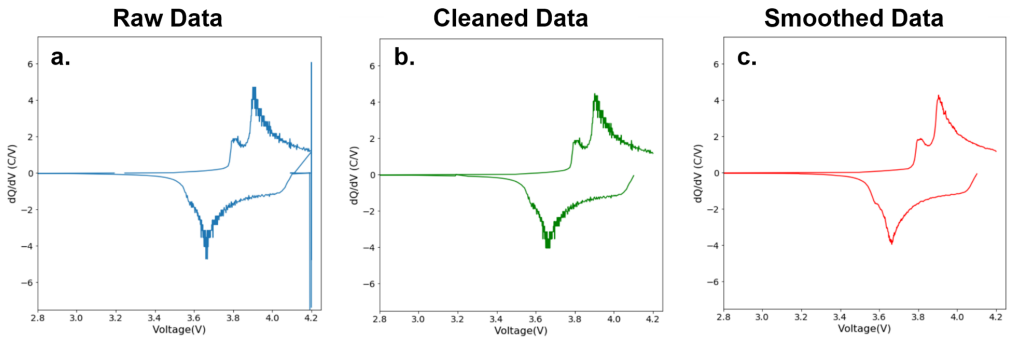

Figure 1: Cleaning process on an example differential capacity curve.

Once the data is clean, the software automatically finds peaks in the dQ/dV curves utilizing the PeakUtils Python package ("PeakUtils," n.d.), and returns the peak heights and the peak locations, as shown by an example cycle in Figure 2a. These peak heights and locations are then used to inform the model build, which is individualized to each cycle contained in the dataset. The model consists of Pseudo-Voigt distributions positioned at the identified peak locations and a baseline Gaussian distribution that captures the area that is not part of the Pseudo-Voigt distributions. The Pseudo-Voigt distribution described by Equations 2 and 3 is simply the linear combination of a Gaussian and a Lorentzian. This distribution is often used in fitting experimental spectral data due to it being a highly generalizable function able to fit a large variety of peak shapes (Wertheim, Butler, West, & Buchanan, 1974).

$$f_v(x, A, \mu, \sigma, \alpha) = \frac{(1-\alpha)A}{\sigma_g\sqrt{2\pi}} \exp\left[-(x-\mu)^2/2{\sigma_g}^2\right] + \frac{\alpha A}{\pi}\left[\frac{\sigma}{(x-\mu)^2 + \sigma^2}\right] \tag{2}$$

$$\sigma_g = \sigma/\sqrt{2\ln 2} \tag{3}$$

Once the model is generated, an optimized fit is found by allowing all parameters to vary except the center position of the Pseudo-Voigt peaks, which are assigned via the previously

identified peak locations. Figure 2b presents an example of an initial model fit and the model fit once optimized specifically for that charge cycle.

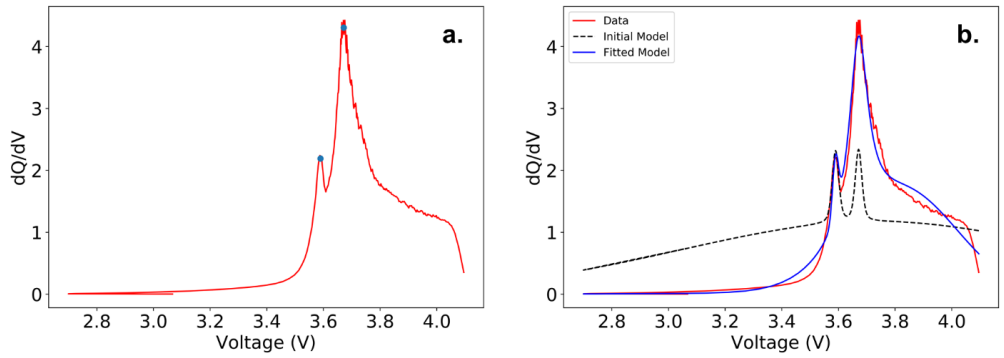

Figure 2: Fitting process on an example differential capacity curve.

Further example model fits can be found on GitHub. From this model, peak areas, widths, and shapes can be extracted and examined to give further insight into the electrochemical processes occurring. The software also utilizes an SQLite database backend to store raw data, cleaned data, model parameters, and peak descriptors for each cycle. In addition to the data processing abilities of the software, a Dash-based web application has been developed where users can upload their own raw data to be processed and visualize the resulting dQ/dV plots and peak descriptors. Users can also evaluate the model fit, alter the threshold for peak detection, and update the model and descriptors in the database. From this application users can also download the cycle descriptors data as a CSV file for their own uses. Further instructions and descriptions of the software functionality can be found in the DiffCapAnalyzer Github repo.

In summary, DiffCapAnalyzer provides the ability to quantitatively analyze battery cycling data. The peak descriptors obtained from DiffCapAnalyzer could be used to classify electrochemical events, visualize battery degradation over time, or as features in state of health analyses. This package also lays the groundwork for a standardized method for cleaning and analyzing this type of data. We hope that this Python package advances the field of electrochemistry and enables researchers to better analyze, interpret, and present their battery cycling data.

Acknowledgments

This project was supported by: Data Intensive Research Enabling Clean Technology (DIRECT) National Science Foundation (NSF) National Research Traineeship (DGE-1633216), the State of Washington through the University of Washington (UW) Clean Energy Institute and the UW eScience Institute, in part upon the work of S.A. and G.A.W. supported by the NSF Graduate Research Fellowship under Grants No. DGE-1762114 and DGE-1256082, respectively, and via funding from the Washington Research Foundation.

References

Aihara, Y., Ito, S., Omoda, R., Yamada, T., Fujiki, S., Watanabe, T., Park, Y., et al. (2016). The electrochemical characteristics and applicability of an amorphous sulfide-based solid ion conductor for the next-generation solid-state lithium secondary batteries. *Frontiers in Energy Research*, *4*. doi:10.3389/fenrg.2016.00018

Bian, X., Liu, L., & Yan, J. (2019). A model for state-of-health estimation of lithium ion batteries based on charging profiles. *Energy*, *177*, 57–65. doi:10.1016/j.energy.2019.04.070

Bloom, I., Jansen, A. N., Abraham, D. P., Knuth, J., Jones, S. A., Battaglia, V. S., & Henriksen, G. L. (2005). Differential voltage analyses of high-power, lithium-ion cells: 1. Technique and application. *Journal of Power Sources*, *139*(1-2), 295–303. doi:10.1016/j.jpowsour.2004.07.021

Christophersen, J. P., Bloom, I., Thomas, E. V., Gering, K. L., Henriksen, G. L., Battaglia, V. S., & Howell, D. (2006). Advanced technology development program for lithium-ion batteries: Gen 2 performance evaluation final report. doi:10.2172/911596

Christophersen, J. P., & Shaw, S. R. (2010). Using radial basis functions to approximate battery differential capacity and differential voltage. *Journal of Power Sources*, *195*(4), 1225–1234. doi:10.1016/j.jpowsour.2009.08.094

He, J., Bian, X., Liu, L., Wei, Z., & Yan, F. (2020). Comparative study of curve determination methods for incremental capacity analysis and state of health estimation of lithium-ion battery. *Journal of Energy Storage*, *29*, 101400. doi:10.1016/j.est.2020.101400

Huang, M. (2019). Incremental capacity analysis-based impact study of diverse usage patterns on lithium-ion battery aging in electrified vehicles. *Batteries*, *5*(3), 59. doi:10.3390/batteries5030059

Marzocca, L. M., & Atwater, T. B. (n.d.). Differential capacity-based modeling for in-use battery diagnostics, prognostics, and quality assurance. *U.S. Army RDECOM Communications, Electronics, Research, Development and Engineering Center Aberdeen Proving Ground, MD 21005*.

PeakUtils. (n.d.).*Bitbucket*. Retrieved from https://bitbucket.org/lucashnegri/peakutils/src/master/

Torai, S., Nakagomi, M., Yoshitake, S., Yamaguchi, S., & Oyama, N. (2016). State-of-health estimation of $LiFePO_4$/graphite batteries based on a model using differential capacity. *Journal of Power Sources*, *306*, 62–69. doi:10.1016/j.jpowsour.2015.11.070

Weng, C., Cui, Y., Sun, J., & Peng, H. (2013). On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *Journal of Power Sources*, *235*, 36–44. doi:10.1016/j.jpowsour.2013.02.012

Wertheim, G. K., Butler, M. A., West, K. W., & Buchanan, D. N. E. (1974). Determination of the Gaussian and Lorentzian content of experimental line shapes. *Review of Scientific Instruments*, *45*(11), 1369–1371. doi:10.1063/1.1686503