Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

Benchmark Synthetic Training Data for Artificial Intelligence-based Li-ion Diagnosis and Prognosis

Matthieu Dubarry [1],*, and David Beck [2]

[1] University of Hawaii, Hawaii Natural Energy Institute; matthieu.dubarry@gmail.com
[2] University of Hawaii, Hawaii Natural Energy Institute; dmbeck@hawaii.edu
* Correspondence: matthieu.dubarry@gmail.com

Abstract: Accurate lithium battery diagnosis and prognosis is critical to increase penetration of electric vehicles and grid-tied storage systems. They are both complex due to the intricate, nonlinear, and path-dependent nature of battery degradation. Data-driven models are anticipated to play a significant role in the behavioral prediction of dynamical systems such as batteries. However, they are often limited by the amount of training data available. In this work, we generated the first big data comprehensive synthetic datasets to train diagnosis and prognosis algorithms. The proof-of-concept datasets are over three orders of magnitude larger than what is currently available in the literature. With benchmark datasets, results from different studies could be easily equated, and the performance of different algorithms can be compared, enhanced, and analyzed extensively. This will expend critical capabilities of current AI algorithms, tools, and techniques to predict scientific data.

Keywords: Machine Learning, Training Data, alawa, AI

In recent years, artificial intelligence (AI) has attracted a lot of attention for energy applications [1-4]. For the diagnosis and prognosis of lithium-ion battery (LiB), one bottleneck is the training data that is not populated enough and not representative of the projected sporadic usage [3,5]. This can hinder performance as models can only be as good as the data they were trained with. Most studies had training datasets below 20 samples. Among the studies with more, [5-12], the study by Severson *et al.* [5,11] stands out with 124 different conditions tested, although only charging conditions were varied. Even though online databases [5,13-18] are a step in the right direction, this is vastly insufficient because LiB degradation is path-dependent and small changes in conditions were shown to lead to drastic differences in durability [19]. This path dependence is an essential aspect to consider for the validation of online diagnosis and prognosis tools [20,21]. With a limited set of training data, the universality of the diagnosis and prognosis tool cannot be proven.

The need for high throughput computational data generation was highlighted in recent reviews [1,2], although no effort towards computational cycling training data has been reported. In this work, we will report the first synthetic benchmark training datasets that englobe the entire degradation spectrum. This will be done using computer-assisted voltage curve generation to remove the need for lengthy and costly experimental campaigns. The datasets will be computed using the mechanistic modeling approach we pioneered, along with other groups, in the mid-2000s [22-25]. The approach has been well validated [26-28] and has been intensively used in recent years with the rise in popularity of the electrochemical voltage spectroscopies (EVS) [29,30]. The use of the mechanistic approach will enable the creation of training datasets several orders of magnitude larger than the current ones encompassing all possible degradation scenarios. This is especially important for prognosis tools as battery capacity loss is known to accelerate at some point upon aging. This is the essential feature than any good prognosis needs to capture. This acceleration usually originates from degradations that have an “incubation” period that does not affect capacity [21]. A complete training dataset should contain examples of paths with an acceleration induced by different components of the degradation, whether it is the loss of lithium, losses in active materials, or resistance.

Compared to traditional battery modeling [31,32], the mechanistic framework is a backward modeling approach where the input is the degradation and the output the cell's voltage and capacity. This makes the method perfect for generating training data and more efficient than electrochemical models because there is no need to find actual physical phenomena that could lead to a given degradation. The initial set up is also simpler without a need for complex parameterization and the only prerequisite being half-cell data versus a reference electrode for both electrodes. In 2017, we proposed to use this approach to perform sensibility analyses by simulating a wide range of hypothetical degradation scenarios [20]. This work is building on our previous effort to enable the creation of comprehensive training data.

In this work, we will focus on diagnosis and prognosis, two applications that require different sets of training data. To illustrate the difference between the two requirements, one can think of two time-dependent processes, one linear and one exponential, intersecting. At the intersection, the diagnostic will be the same, but the prognosis needs to be different. In addition, we will also address another hurdle for AI algorithms, the quest for meaningful learnable parameters. Although some studies seem to fit the full constant current voltage data [33], most studies favor the use of features of interest (FOI) and only focus on a specific part of the electrochemical response. This could be capacity and resistance evolution [10,34-37], curvature [38,39], sections of the voltage response [8,40-42], electrochemical impedance spectroscopy [43,44], variance [5,11], or EVS [33,45-48]. The latter has attracted a lot of attention in recent years since the early work on the technique [22,49,50]. Still, the correlation of the variations with degradation is not trivial and requires significant sensibility analysis to derive universal parameters [20]. Some FOIs proposed in the literature were proven not applicable outside of the tested data [21], which illustrates the need for more extensive training sets and proper sensitivity analysis of the learnable parameters. With controlled and complete training, the theoretical framework to establish relationships between data and models can be explored in much more detail, and results from different studies could be easily equated.

The purpose of this publication is to introduce a big data methodology and showcase its possibilities. The optimization of the technique to reduce computation time and its integration with AI algorithms is out of the scope of this work.

Methodology

The mechanistic approach conceptually follows the one described by Christensen and Newman to simulate cell degradation in the early 2000s [31,32]. Instead of computing-intensive electrochemical models to simulate half-cell electrode behavior, the mechanistic approach [22-25] uses experimental half-cell data for each electrode. Figure 1(a) presents the mechanistic representation of a cell. The matching consists of two parameters, the loading ratio LR between the capacities of both electrodes and their offset OFS that is representative of the amount of lithium lost during the solid electrolyte interphase formation. The circles in Figure 1(a) showcase how the experimental data overlap the emulated cell perfectly. To enhance the differences, the same data can be derived and plotted as incremental capacity, dQ/dV = f(V), Figure 1(b-d). For this EVS, each peak represents a phase transformation in one of the electrodes. A perfect match between the emulated and the experimental data proves that the emulated cell is perfectly replicating the experimental cell voltage variations. Figure 1(b-d) presents examples of the quality of the obtainable fits for cells based on different Li-ion chemistries from our previous work [51-53].

Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

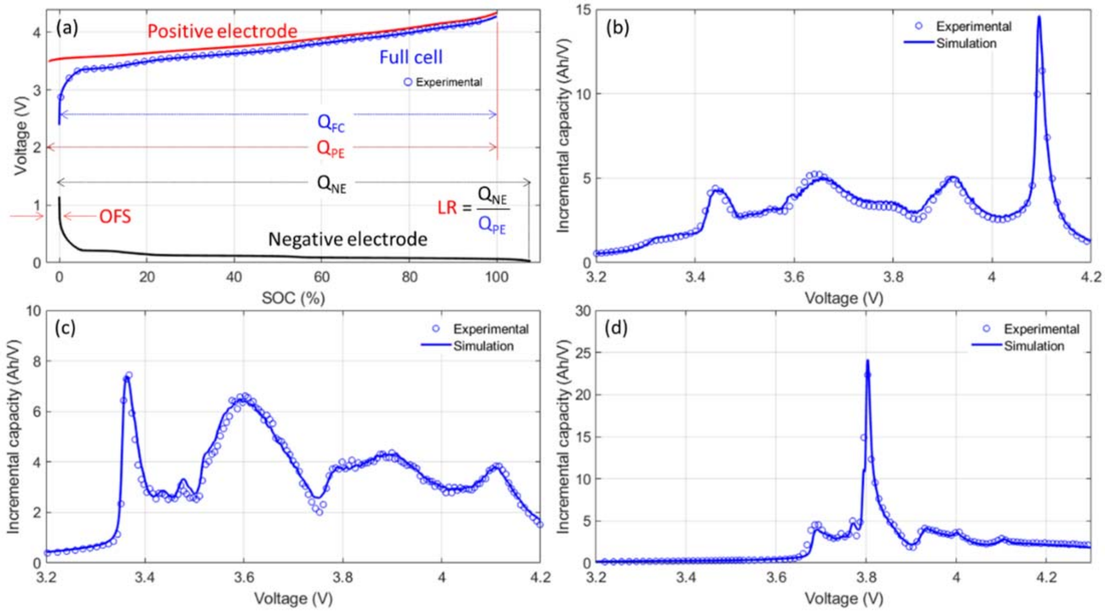


Figure 1: (a) Mechanistic model principles. The full cell response, in blue, corresponds to the response of the positive electrode, in red, minus the one of the negative electrode, in black, after matching. Fit for cells based on (b) a blended Graphite (Gr)-Silicon (Si) NE and $Ni_{.8}Mn_{.1}Co_{.1}O_2$ (NMC_{811}) PE [53], (Gr) negative electrodes (NE) and (c) $Ni_xAl_yCo_{1-x-y}O_2$ (NCA) [51] as well as a (d) blend of $LiCoO_2$ (LCO) and NCA as (PE) [52].

While the active materials are considered stable with aging, their quantity, as well as the amount of Lithium reacting, will change upon degradation. Degradation will then not affect the electrode OCV curves, but it will impact their matching. If less active material is available, the loading ratio between the electrodes will change. If some reactant is lost, the synchronicity of the electrodes will change. These matching changes can be rendered in the mechanistic approach via some scaling of the electrode curves and a translation of one electrode compared to the other, in other words changing LR and OFS in Figure 1(a). Figure 2(a) presents the impact of the variation of LR and OFS on cell capacity. In can be seen that there is an infinity of combination of LR and OFS that can lead to any capacity loss; this is the path dependence of the degradation. This also explains why measuring capacity will never be a good enough diagnosis of true degradation. Changes in the amount of active material are referred to as loss of active material (LAM). LAM can occur at the PE and NE. Change in the amount of lithium reacting is referred to as loss of lithium inventory (LLI). The thermodynamic degradation of Li-ion batteries can be decomposed into these three main degradation modes. LLI is inducing translation and the LAMs scaling, Figure 2(b-d). More details on the technique and the associated equations can be found in [25].

doi:10.20944/preprints202005.0106.v1

Peer-reviewed version available at Journal of Power Sources 2020, 479, 228806; doi:10.1016/j.jpowsour.2020.228806

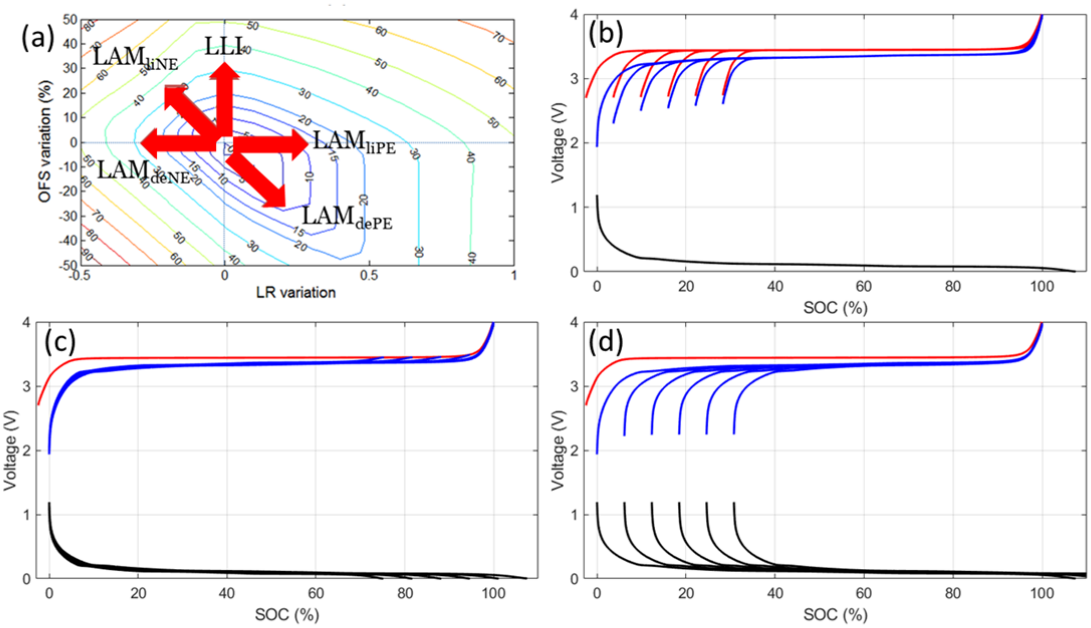


Figure 2: (a) Impact of changes of LR and OFS on the capacity loss. Arrows represent the impact of LLI and LAMs on LR and OFS. (b) Impact of LAM$_{PE}$ and (c) impact of LAM$_{NE}$ with the scaling of the affected electrode compared to the other. (d) Impact of LLI with the translation of the NE compared to the PE. In all cases, the full cell response, in blue, corresponds to the response of the positive electrode, in red, minus the one of the negative electrode, in black, after matching.

The validity of the predicted impact of LAM and LLI on the electrochemical behavior of experimental cells was verified by independent studies [26-28]. Most notably, Kassem *et al.* [26] confirmed that LLI does induce electrode slippage between the NE and Schmidt *et al.* [54] verified that LAM does influence the LR by assembling cells with different amounts of PE versus a constant NE. Moreover, over the past decade, numerous studies [29,30] successfully used this approach to diagnose the degradation of commercial cells. Figure 3(a,b) presents two examples of the fit of the voltage response of aged commercial cells from previous works [53,55]. Figure 3(c,d) presents the comparison of an experimental (from [53]) and a synthetic dataset constructed solely from half-cell data and simple equations that modified the matching of the electrodes based on the quantification of LLI and the LAMs. This highlights how the mechanistic approach is efficient to replicate the voltage variations associated with cell degradation.

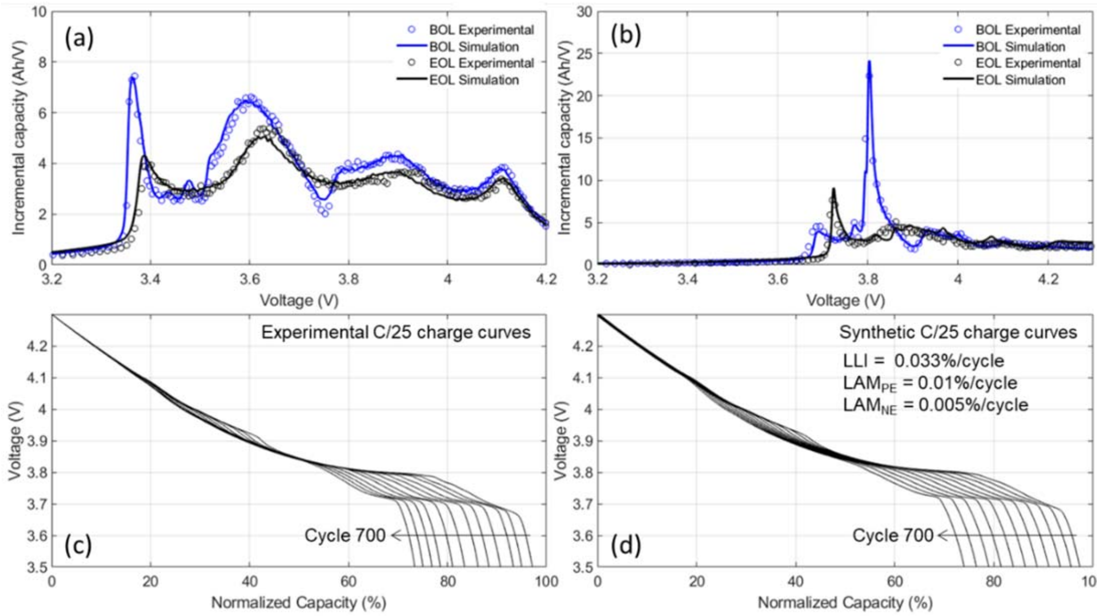


Figure 3: Example of voltage emulation after aging for (a) a Gr//NCA discharge [55] after 1,500 equivalent full cycles based on electric vehicle driving and frequency regulation. (b) A Gr//LCO+NCA charges [53] after 1000 full cycles at 1.5C. (c) Experimental [53] and (d) synthetic voltage variations upon aging calculated from simple linear equations extracted from the diagnosis.

From the mechanistic approach point of view, each thermodynamic degradation can be described by a unique combination of the LLI, LAMPE, and LAMNE triplet. Conversely, scanning every possible combination of the triplet yield any potential degradation from which the voltage response can be reconstructed. This is the unique feature that is used in this work to build synthetic training datasets.

For datasets aimed at diagnosis, the approach used in previous work on an automated diagnosis tool for BMS [20] can be expanded. The [LLI, LAMPE, LAMNE] triplets can be normalized so that their sum is equal to one and they can then be represented in a ternary diagram, Figure 4. By scanning every point of this diagram (red dots) between 0 to 100% degradation for each mode (vertical lines), each possible degradation can be represented, and the associated voltage response calculated to be used in a training dataset. E.g., for the [0.5,0.2.,0.3] triplet, LLI would be varied from 0 to 100% with 1% increments, LAMPE from 0 to 40% with 0.4% increments, and LAMNE from 0 to 60% with 0.6% increments so that their ratio is always conserved.

Peer-reviewed version available at Journal of Power Sources 2020, 479, 228806; doi:10.1016/j.jpowsour.2020.228806

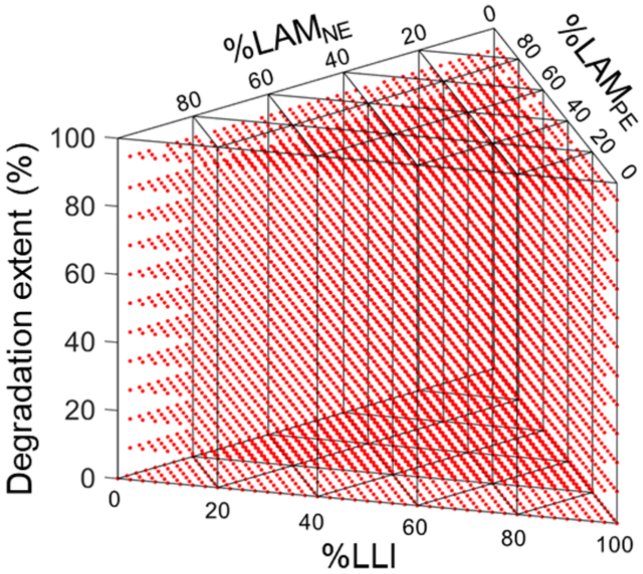


Figure 4: (a) Mechanistic model principles to emulate every possible degradation. Only partial degradation calculations were represented for illustration purposes.

For datasets aimed at prognosis, the evolution of each degradation mode needs to be calculated independently as the ratio between the normalized values of the triplet might not be constant. The triplet values could be calculated, per example, for different linear degradation paces, with paces linear combined with exponential variations, or with a delayed exponential increase, more representative of the real degradation of commercial cells [19,56,57]. Since the voltages associated with all the triplet values were already calculated for diagnosis, no additional simulations are actually necessary and the different degradation paths are all inscribed in Figure 4. They can be reconstructed by following the evolution of the triplet values in the ternary design space and the corresponding voltage variations can be accessed. Figure 5 presents an example of the process for two degradation paths for a Gr//$LiFePO_4$ (LFP) cell. Figure 5(a) plots the evolution of the triplets for each path. For each cycle, as exemplified by the black box in Figure 5(a), a triplet value can be determined and the quantification can be used to find the corresponding point in the ternary design space, Figure 5(b). In this example, path 2 (green) was simulated with linear variations for the three components of the triplet, its path is, therefore, vertical as the ratio between the degradation modes is constant. Path 1 (blue) had some exponential component for LLI and LAM_{PE} and, therefore, the path is more complicated. Nonetheless, based on the triplets and the diagnosis dataset, the capacity loss, Figure 5(c), and the voltage variations, Figure 5(d), can be deciphered. This approach for calculating a prognosis dataset will be faster than simulating each degradation path one by one as it will avoid repeating the same simulations over and over.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 May 2020 doi:10.20944/preprints202005.0106.v1

Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

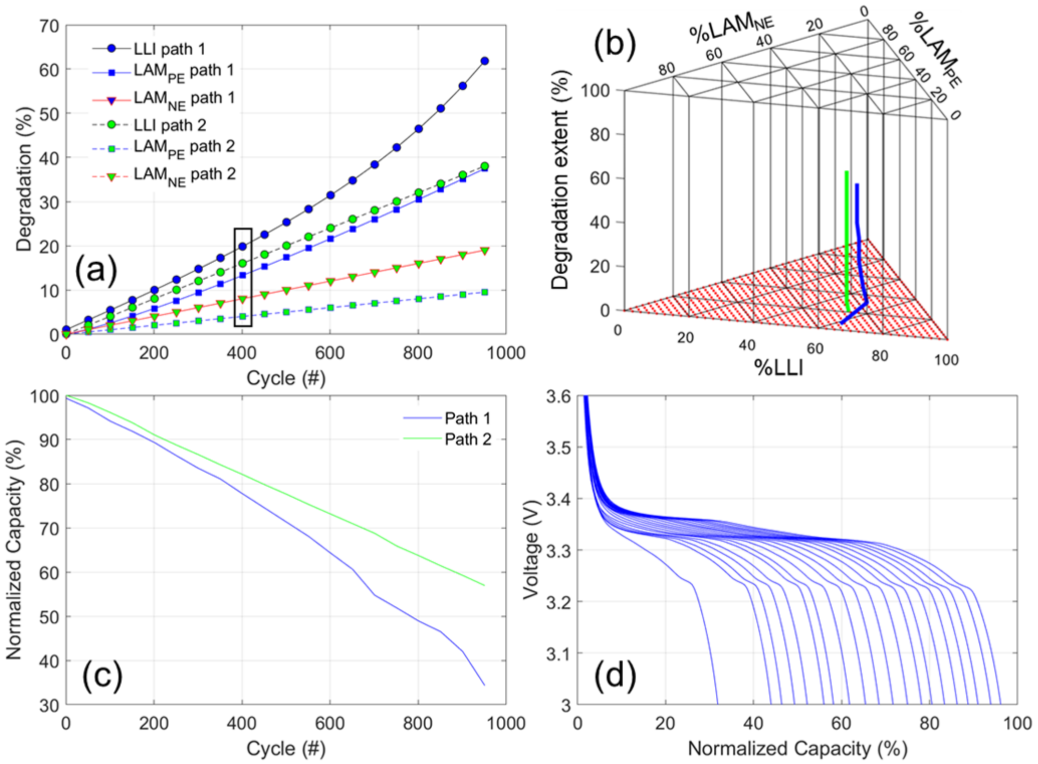


Figure 5: Example of the prognosis dataset calculation for two different paths for a Gr//LFP cell with (a) the triplet variations, (b) their path in the ternary design space, (c) the reconstructed capacity loss, and (d) the voltage variations for path 1.

It must be noted that the description above was simplified for narrative purposes and was centered on the thermodynamic degradation observable at low currents and constant temperature for ease of representation. The full mechanistic approach is much more complex [25]. To be able to simulate different levels of usage, an equivalent circuit model (ECM) is included to handle different rates for each electrode [22-26]. For electrode blends, or changes in electroactive phases upon aging, each component of the electrode must be considered separately (i.e., LAM_{PE} will be splat in LAM_{LCO} and LAM_{NCA} for an LCO-NCA blended electrode). Moreover, for datasets at higher currents or different temperatures, some kinetic parameters such as the ohmic and faradic resistance increase must also be scanned as they will influence the voltage response of the cells. Finally, in some cases, lithium plating, both thermodynamic and kinetic, must be taken into consideration [25]. The validity of the mechanistic approach in all of those cases has been proven in the literature [19,58,59]. As explained in [25,58], resistance changes can be accommodated by shifting the voltage curves up or down depending on the electrode and the regime. Kinetic limitations can be handled by modifying the simulated rate while counterbalancing the polarization changes. Simulating higher rates allows emulating slower kinetics. Simulating lower rates allows emulating faster kinetics. This allows to model different rates and temperatures [58]. Figure 6(a) presents an example from [58] where different temperatures were simulated from 25°C data at different rates. For lithium plating, the solution is to simulate the NE as a blend between Gr and metallic lithium with the metallic lithium initially outside of the potential window, Figure 6(b). If $LAM_{Graphite}$ becomes prominent, or if its resistance pushes its potential below 0V, the metallic lithium will then be dragged into the potential window and start interacting with the PE [19,21,25]. As discussed in the literature [5,19,25,57], thermodynamic lithium plating begins when there is enough LAM_{NE} so that the NE is fully lithiated before the PE is delithiated. The amount of LAM_{NE} necessary for that to happen depends on the initial cell configuration (LR_{ini} and OFS_{ini}) and also on the evolution of LLI and LAM_{PE} [60]. Figure 6(c) plots an example of the evolution of the plating threshold (PT) for a randomly chosen

Peer-reviewed version available at Journal of Power Sources 2020, 479, 228806; doi:10.1016/j.jpowsour.2020.228806

degradation path. Lithium plating will start when LAMNE becomes larger than PT. This does not by itself induce any capacity loss as metallic lithium can be reduced and oxidized. This lithium, however, is usually quickly passivating, which results in additional LLI. To take this effect into account, a correction for LLI can be introduced and the additional LLI is defined as the difference between LAMNE and the PT, when positive, multiplied by a reversibility factor. Figure 6(c) showcases an example of the acceleration in LLI based on different reversibilities. Figure 6(d) presents the associated capacity loss showcasing the acceleration.

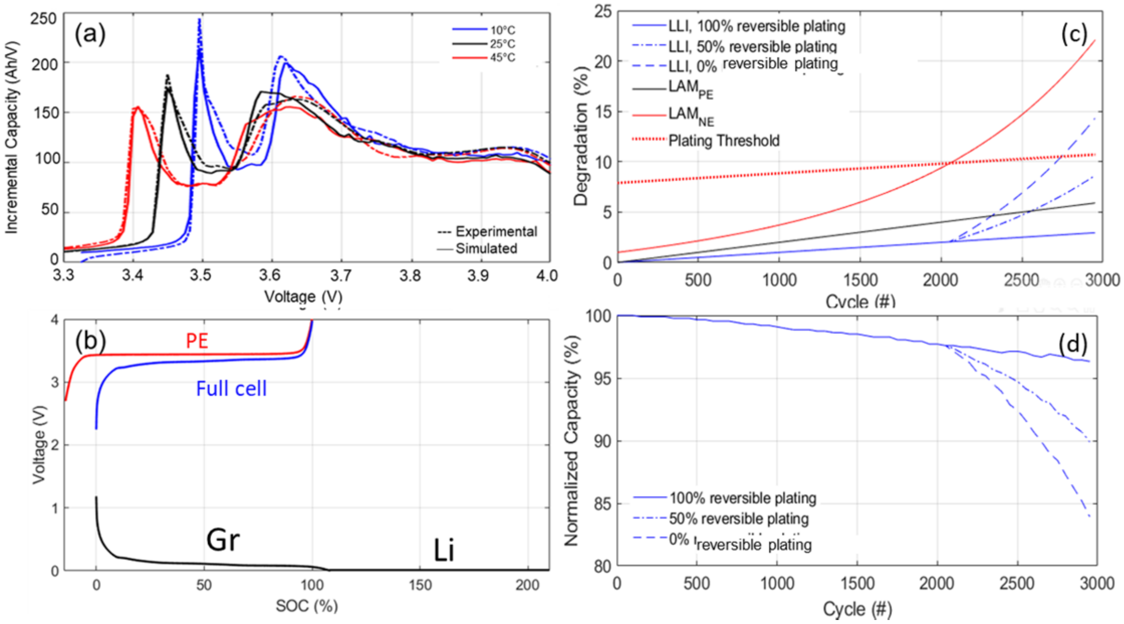


Figure 6: (a) Emulation of the impact of temperature at C/5 for a Gr//NCA cell from [58], and (b) Electrode matching with a blended Gr-Li electrode with the Li outside of the potential window as a reserve. Evolution of (c) LLI and LAMNE depending on the reversibility of lithium plating (0%, 50% or 100%) and (d) the associated capacity loss.

Proof-of-concept datasets

For this proof-of-concept work, a Gr//LFP cell similar to the one used in [20] was chosen because the sensibility analysis of different FOI was already reported and some non-ML diagnosis tools were validated. The data could also be compared to the one from Severson *et al.* [5].

The diagnosis dataset contains more than 500,000 voltage vs. capacity curves. It can be used to train machine learning algorithms, but more importantly, it can also be used to perform sensibility analysis [20] to find out which FOIs are representative of different degradations or if they can be used as universal proxies for capacity loss. An example is presented in Supplementary note 1.

The prognosis dataset contains more than 130,000 different degradation paths with upwards of 3,000,000 individual voltage vs. capacity curves. It can be used to test the validity of varying FOIs proposed in the literature for early prognosis. An example is presented in Supplementary note 2.

Both datasets are available to download (see Data availability section for access information).

Limitations and outlook

In this work, we presented a proof-of-concept methodology to generate big data training datasets for intercalation batteries. The mechanistic approach combines both modeling and experimental techniques to provide a universal tool for the creation of synthetic voltage curves practically indistinguishable from real data. This approach offers the benefits of the broad applicability of the model to various cell chemistries, designs, and operating modes, as well as the high fidelity as inhered from the detailed extraction of the experimental data. The methodology could be applied to the generation of training datasets encompassing the entire degradation spectrum as well as different operating conditions such as rate and temperature. Scenarios leading to lithium plating with adjustable reversibility can be added and the addition of random noise on the data could be considered to make the datasets even more realistic and consider the impact of real-life testing machines' precision and accuracy.

As a proof-of-concept, we generated a diagnosis dataset containing more than individual 500,000 voltage vs. capacity curves and a prognosis dataset with more than 130,000 individual degradation paths for a commercial graphite//LFP battery. These datasets are more than three orders of magnitude larger than the one currently available in the literature and they could be extended at will. We showcased the use of the proposed datasets for FOI sensibility analysis and a diagnosis technique based on the joint evolution of three FOIs since variations of single FOI were proven inefficient. We also used them to propose an early prognosis method based on the extrapolation of the diagnosis of the degradation modes.

Although there is still a lot of work to be done to optimize the technique, results are extremely promising and should accelerate the development of accurate algorithms. The proof-of-concept dataset used in this work were limited to thermodynamic degradations at constant temperature for single-phase electrodes. Taking everything into account (LAMs, LLI, polarization, kinetics, phases, temperatures…) for all the possible evolutions (linear, exponential, delays…) will take enormous calculation time and generate overwhelming datasets. Some optimization and intelligence could be applied to limit the number of degradation paths to test. A solution could be large designs of experiments first to screen the impact of the different parameters then to focus on their representativity in the paths to simulate. Here we provide an open framework to the community to apply advanced optimization techniques to reduce the calculation burden.

Finally, we would like to point out that our intent is no way to remove the need for experimental testing all-together. On the contrary, experimental testing is still essential, and the only way, to decipher which conditions cater to specific degradation paths and its input is, therefore, fundamental to decipher the applicability of batteries for a given application. Our methodology can be used alongside to develop the needed algorithms once a cell was selected.

Methods

Half-cell data. Data from four different types of commercial cells were used in this work. The Gr//NCA cell was a Panasonic 3350 mAh NCR 18650B [51,58,60-62]. The Si-Gr//NMC811 was a Samsung-SDI 3350 mAh INR18650-35E [52], the Gr//LCO+NCA was an LG-Chem 2800 mAh ICR18650 C2” [53], and the Gr//LFP was an A123 2300 mAh ANR26650M [19,63]. All were extensively studied in previous work. Interested readers are referred to the original publications for more details. The cells were slowly discharged to 0% SOC before being carefully disassembled in an Argon-filled glove box where the electrodes, separator, and casing will be separated [64]. Electrode discs (18 mm in diameter) were punched from the harvested electrodes and rinsed in a dimethyl carbonate (DMC) solution. The backside of two-sided electrodes was then wiped clean using cotton swabs soaked in nN-methyl-2-pyrrolidone (NMP) solvent before being the electrode was rinsed again in a fresh DMC solution. The electrodes were tested using a lithium counter electrode. Metallic lithium was applied onto a stainless-steel disk. An electrolyte solution of 1.0 M LiPF6 in ethylene carbonate (EC) + DMC (1:1 by weight) + 2% wt. vinylene carbonate (VC) was used to soak the separator, which consisted of one layer of Whatman GF-D fiberglass discs (12.7 mm in diameter, Whatman), Kent, United Kingdom). Once sealed, the half-cells were taken out of the glove box and connected to a multi-channel Bio-Logic VMP3 potentiostat (Bio-Logic, Claix, France) for testing at different rates from low to high. The complete testing protocol can be found in [64].

Simulations. All simulations were performed under a MATLAB© environment using the ‘alawa toolbox [65]. As in [20], the LFP and graphite electrodes were matched with an LR of 0.95, OFS of 12.5%, and a resistance adjustment of -0.07 to match the A123 ANR26650M experimental data.

The diagnosis training dataset was compiled with a resolution of 0.01 for the triplets and C/25 charges. This accounts for more than 5,000 different paths at the base of the triangle in Figure 4. Each path was simulated with 0.85% increases for each degradation up to 85%. This accounts for 100 simulations per path. The training dataset, therefore, contains more than 500,000 voltage vs. capacity curves and took around 12h to compute on a standard laptop. The 500MB dataset is available to download (see Data availability section for access information).

The prognosis dataset was harder to define as there are no limits on how the three degradation modes can evolve. For this proof of concept work, we considered eight parameters to scan. For each degradation mode, degradation was chosen to follow equation (1).

$$\%degradation = a \times cycle + (exp^{b \times cycle} - 1) \quad (1)$$

Considering the three degradation modes, this accounts for six parameters to scan. In addition, two other parameters were added, a delay for the exponential factor for LLI, and a parameter for the reversibility of lithium plating. The delay was introduced to reflect degradation paths where plating cannot be explained by an increase of LAMs or resistance [55]. The chosen parameters and their values are summarized in Table S1. Figure S1(a,b) presents the evolution of parameters p_1 to p_7. At the worst, the cells endured 100% of one of the degradation modes in around 1,500 cycles. Minimal LLI was chosen to be 20% after 3,000 cycles. This is to guarantee at least 20% capacity loss for all the simulations. For the LAMs, conditions were less restrictive, and, after 3,000 cycles, the lowest degradation is of 3%. The reversibility factor p_8 was calculated with equation (2) when $LAM_{NE} > PT$.

$$\%LLI = \%\mathrm{LLI} + p_8(\%LAM_{PE} - PT) \quad (2)$$

Where PT was calculated with equation (3) from [60].

$$PT = 100 - \left[\left(\frac{100 - \%LAM_{PE}}{100 \times LR_{ini} - \%LAM_{PE}}\right) \times (100 - OFS_{ini} - \%LLI)\right] \quad (3)$$

Varying all those parameters accounted for more than 130,000 individual duty cycles. With one voltage curve for every 100 cycles, it took around 12h to compute on a standard laptop to simulate the more than 3,000,000 voltage vs. capacity curves. The 3GB dataset is available to download (see Data availability section for access information).

Data availability. The datasets used in this study are available at http://dx.doi.org/10.17632/bs2j56pn7y.1 and http://dx.doi.org/10.17632/6s6ph9n8zg.1 for the diagnosis and prognosis dataset, respectively.

Author Contributions: *Conceptualization, M.D.; methodology, M.D.; software, M.D.; validation, M.D. and D.B.; formal analysis, M.D. and D.B.; resources, M.D.; data curation, M.D.; writing—original draft preparation, M.D.; writing—review and editing, M.D. and D.B.; visualization, M.D.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.*

Funding: This work was funded by ONR Asia Pacific Research Initiative for Sustainable Energy Systems (APRISES), award number N00014-18-1-2127. M.D. is also supported by the State of Hawaii.

Acknowledgments: M.D. is thankful to all the past staff that helped define and affine the concepts used in this work, most notably Bor Yann Liaw, Cyril Truchot, Arnaud Devie, George Baure, and David Anseán.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S.P. A Critical Review of Machine Learning of Energy Materials. *Advanced Energy Materials* **2020**, 10.1002/aenm.201903242, doi:10.1002/aenm.201903242.

2. Ng, M.-F.; Zhao, J.; Yan, Q.; Conduit, G.J.; Seh, Z.W. Predicting the state of charge and health of batteries using data-driven machine learning. *Nature Machine Intelligence* **2020**, 10.1038/s42256-020-0156-7, doi:10.1038/s42256-020-0156-7.
3. Vidal, C.; Malysz, P.; Kollmeyer, P.; Emadi, A. Machine Learning Applied to Electrified Vehicle Battery State of Charge and State of Health Estimation: State-of-the-Art. *IEEE Access* **2020**, *8*, 52796-52814, doi:10.1109/access.2020.2980961.
4. How, D.N.T.; Hannan, M.A.; Hossain Lipu, M.S.; Ker, P.J. State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods: A Review. *IEEE Access* **2019**, *7*, 136116-136136, doi:10.1109/access.2019.2942213.
5. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D., et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy* **2019**, *4*, 383-391, doi:10.1038/s41560-019-0356-8.
6. Klass, V.; Behm, M.; Lindbergh, G. A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. *J. Power Sources* **2014**, *270*, 262-272, doi:10.1016/j.jpowsour.2014.07.116.
7. Klass, V.; Behm, M.; Lindbergh, G. Evaluating Real-Life Performance of Lithium-Ion Battery Packs in Electric Vehicles. *J. Electrochem. Soc.* **2012**, *159*, A1856-A1860, doi:10.1149/2.047211jes.
8. Hu, C.; Jain, G.; Schmidt, C.; Strief, C.; Sullivan, M. Online estimation of lithium-ion battery capacity using sparse Bayesian learning. *J. Power Sources* **2015**, *289*, 105-113, doi:10.1016/j.jpowsour.2015.04.166.
9. Richardson, R.R.; Birkl, C.R.; Osborne, M.A.; Howey, D.A. Gaussian Process Regression for In-situ Capacity Estimation of Lithium-ion Batteries.
10. Pan, H.; Lü, Z.; Wang, H.; Wei, H.; Chen, L. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. *Energy* **2018**, *160*, 466-477, doi:10.1016/j.energy.2018.06.220.
11. Attia, P.M.; Grover, A.; Jin, N.; Severson, K.A.; Markov, T.M.; Liao, Y.H.; Chen, M.H.; Cheong, B.; Perkins, N.; Yang, Z., et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **2020**, *578*, 397-402, doi:10.1038/s41586-020-1994-5.
12. Cripps, E.; Pecht, M. A Bayesian nonlinear random effects model for identification of defective batteries from lot samples. *J. Power Sources* **2017**, *342*, 342-350, doi:10.1016/j.jpowsour.2016.12.067.
13. Kollmeyer, P.; Vidal, C.; Naguib, M.; Skells, M. LG 18650HG2 Li-ion Battery Data and Example Deep Neural Network xEV SOC Estimator Script. Available online: https://data.mendeley.com/datasets/cp3473x7xv/3 (accessed on 04/12).
14. Kollmeyer, P. Panasonic 18650PF Li-ion Battery Data", Mendeley Data, v1. Available online: https://data.mendeley.com/datasets/wykht8y7tg/1#folder96f196a8-a04d-4e6a-827d-0dc4d61ca97b (accessed on 04/12).
15. Saha, B.; Goebel, K. Battery Data Set. NASA Ames Prognostics Data Repository. NASA Ames Research Center. Moffett Field, CA, USA. Availabe online: https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/ (accessed on 04/12).
16. Barkholtz, H.M.; Fresquez, A.; Chalamala, B.R.; Ferreira, S.R. A Database for Comparative Electrochemical Performance of Commercial 18650-Format Lithium-Ion Cells. *J. Electrochem. Soc.* **2017**, *164*, A2697-A2706, doi:10.1149/2.1701712jes.
17. Pecht, M. Battery Research Data. Availabe online: https://calce.umd.edu/data (accessed on 04/14).

18. Birkl, C.R.; Offer, G.J. Oxford battery degradation dataset from the Howey Research Group. Availabe online: https://ora.ox.ac.uk/objects/uuid%3a03ba4b01-cfed-46d3-9b1a-7d4a7bdf6fac (accessed on 04/14).
19. Anseán, D.; Dubarry, M.; Devie, A.; Liaw, B.Y.; García, V.M.; Viera, J.C.; González, M. Operando lithium plating quantification and early detection of a commercial LiFePO 4 cell cycled under dynamic driving schedule. *J. Power Sources* **2017**, *356*, 36-46, doi:10.1016/j.jpowsour.2017.04.072.
20. Dubarry, M.; Berecibar, M.; Devie, A.; Anseán, D.; Omar, N.; Villarreal, I. State of health battery estimator enabling degradation diagnosis: Model and algorithm description. *J. Power Sources* **2017**, *360*, 59-69, doi:10.1016/j.jpowsour.2017.05.121.
21. Dubarry, M.; Baure, G.; Anseán, D. Perspective on State-of-Health Determination in Lithium-Ion Batteries. *Journal of Electrochemical Energy Conversion and Storage* **2020**, *17*, 1-25, doi:10.1115/1.4045008.
22. Bloom, I.; Jansen, A.N.; Abraham, D.P.; Knuth, J.; Jones, S.A.; Battaglia, V.S.; Henriksen, G.L. Differential voltage analyses of high-power, lithium-ion cells. 1. Technique and Applications. *Journal of Power Sources* **2005**, *139*, 295-303, doi:10.1016/j.jpowsour.2004.07.021.
23. Honkura, K.; Honbo, H.; Koishikawa, Y.; Horiba, T. State Analysis of Lithium-Ion Batteries Using Discharge Curves. *ECS Transactions* **2008**, *13*, 61-73.
24. Dahn, H.M.; Smith, A.J.; Burns, J.C.; Stevens, D.A.; Dahn, J.R. User-Friendly Differential Voltage Analysis Freeware for the Analysis of Degradation Mechanisms in Li-Ion Batteries. *Journal of the Electrochemical Society* **2012**, *159*, A1405-A1409, doi:10.1149/2.013209jes.
25. Dubarry, M.; Truchot, C.; Liaw, B.Y. Synthesize battery degradation modes via a diagnostic and prognostic model. *J. Power Sources* **2012**, *219*, 204-216, doi:10.1016/j.jpowsour.2012.07.016.
26. Kassem, M.; Delacourt, C. Postmortem analysis of calendar-aged graphite/LiFePO4 cells. *J. Power Sources* **2013**, *235*, 159-171, doi:10.1016/j.jpowsour.2013.01.147.
27. Schmidt, J.P.; Tran, H.Y.; Richter, J.; Ivers-Tiffee, E.; Wohlfahrt-Mehrens, M. Analysis and prediction of the open circuit potential of lithium-ion cells. *Journal of Power Sources* **2013**, *239*, 696-704, doi:10.1016/j.jpowsour.2012.11.101.
28. Birkl, C.R.; Roberts, M.R.; McTurk, E.; Bruce, P.G.; Howey, D.A. Degradation diagnostics for lithium ion cells. *Journal of Power Sources* **2017**, *341*, 373-386, doi:10.1016/j.jpowsour.2016.12.011.
29. Barai, A.; Uddin, K.; Dubarry, M.; Somerville, L.; McGordon, A.; Jennings, P.; Bloom, I. A comparison of methodologies for the non-invasive characterisation of commercial Li-ion cells. *Progr. Energy Combust. Sci.* **2019**, *72*, 1-31, doi:10.1016/j.pecs.2019.01.001.
30. Pastor-Fernández, C.; Yu, T.F.; Widanage, W.D.; Marco, J. Critical review of non-invasive diagnosis techniques for quantification of degradation modes in lithium-ion batteries. *Renewable and Sustainable Energy Reviews* **2019**, *109*, 138-159, doi:10.1016/j.rser.2019.03.060.
31. Christensen, J.; Newman, J. Effect of Anode Film Resistance on the Charge/Discharge Capacity of a Lithium-Ion Battery. *J. Electrochem. Soc.* **2003**, *150*, A1416, doi:10.1149/1.1612501.
32. Christensen, J.; Newman, J. Cyclable Lithium and Capacity Loss in Li-Ion Cells. *J. Electrochem. Soc.* **2005**, *152*, A818, doi:10.1149/1.1870752.
33. Fath, J.P.; Dragicevic, D.; Bittel, L.; Nuhic, A.; Sieg, J.; Hahn, S.; Alsheimer, L.; Spier, B.; Wetzel, T. Quantification of aging mechanisms and inhomogeneity in cycled lithium-ion cells by differential voltage analysis. *Journal of Energy Storage* **2019**, *25*, doi:10.1016/j.est.2019.100813.

Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

34. Nuhic, A.; Terzimehic, T.; Soczka-Guth, T.; Buchholz, M.; Dietmayer, K. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *J. Power Sources* **2013**, *239*, 680-688, doi:10.1016/j.jpowsour.2012.11.146.
35. Liu, D.; Luo, Y.; Liu, J.; Peng, Y.; Guo, L.; Pecht, M. Lithium-ion battery remaining useful life estimation based on fusion nonlinear degradation AR model and RPF algorithm. *Neural Computing and Applications* **2013**, *25*, 557-572, doi:10.1007/s00521-013-1520-x.
36. Liu, D.; Pang, J.; Zhou, J.; Peng, Y.; Pecht, M. Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression. *Microelectronics Reliability* **2013**, *53*, 832-839, doi:10.1016/j.microrel.2013.03.010.
37. Lee, C.; Jo, S.; Kwon, D.; Pecht, M. Capacity-fading Behavior Analysis for Early Detection of Unhealthy Li-ion Batteries. *IEEE Transactions on Industrial Electronics* **2020**, 10.1109/tie.2020.2972468, 1-1, doi:10.1109/tie.2020.2972468.
38. Lee, J.; Kwon, D.; Pecht, M.G. Reduction of Li-ion Battery Qualification Time Based on Prognostics and Health Management. *IEEE Transactions on Industrial Electronics* **2019**, *66*, 7310-7315, doi:10.1109/tie.2018.2880701.
39. Fermín, P.; McTurk, E.; Allerhand, M.; Medina-Lopez, E.; Anjos, M.F.; Sylvester, J.; dos Reis, G. Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells. *Energy and AI* **2020**, 10.1016/j.egyai.2020.100006, doi:10.1016/j.egyai.2020.100006.
40. Yang, D.; Zhang, X.; Pan, R.; Wang, Y.; Chen, Z. A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *J. Power Sources* **2018**, *384*, 387-395, doi:10.1016/j.jpowsour.2018.03.015.
41. Goh, T.; Park, M.; Seo, M.; Kim, J.G.; Kim, S.W. Successive-approximation algorithm for estimating capacity of Li-ion batteries. *Energy* **2018**, *159*, 61-73, doi:10.1016/j.energy.2018.06.116.
42. Goh, T.; Park, M.; Seo, M.; Kim, J.G.; Kim, S.W. Capacity estimation algorithm with a second-order differential voltage curve for Li-ion batteries with NMC cathodes. *Energy* **2017**, *135*, 257-268, doi:10.1016/j.energy.2017.06.141.
43. Eddahech, A.; Briat, O.; Bertrand, N.; Delétage, J.-Y.; Vinassa, J.-M. Behavior and state-of-health monitoring of Li-ion batteries using impedance spectroscopy and recurrent neural networks. *International Journal of Electrical Power & Energy Systems* **2012**, *42*, 487-494, doi:10.1016/j.ijepes.2012.04.050.
44. Saha, B.; Poll, S.; Goebel, K.; Christophersen, J.P. An Integrated Approach to Battery Health Monitoring Using Bayesian Regression and State Estimation. In Proceedings of International Automatic Testing Conference, Baltimore, USA.
45. Weng, C.; Cui, Y.; Sun, J.; Peng, H. On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *J. Power Sources* **2013**, *235*, 36-44, doi:10.1016/j.jpowsour.2013.02.012.
46. Li, Y.; Liu, K.; Foley, A.M.; Zülke, A.; Berecibar, M.; Nanini-Maury, E.; Van Mierlo, J.; Hoster, H.E. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews* **2019**, *113*, doi:10.1016/j.rser.2019.109254.
47. Berecibar, M.; Devriendt, F.; Dubarry, M.; Villarreal, I.; Omar, N.; Verbeke, W.; Van Mierlo, J. Online state of health estimation on NMC cells based on predictive analytics. *J. Power Sources* **2016**, *320*, 239-250, doi:10.1016/j.jpowsour.2016.04.109.

Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

48. He, J.; Bian, X.; Liu, L.; Wei, Z.; Yan, F. Comparative study of curve determination methods for incremental capacity analysis and state of health estimation of lithium-ion battery. *Journal of Energy Storage* **2020**, *29*, doi:10.1016/j.est.2020.101400.
49. Dubarry, M.; Svoboda, V.; Hwu, R.; Liaw, B.Y. Incremental capacity analysis and close-to-equilibrium OCV measurements to quantify capacity fade in commercial rechargeable lithium batteries. *Electrochem. Solid-State Lett.* **2006**, *9*, A454-A457, doi:10.1149/1.2221767.
50. Berecibar, M.; Garmendia, M.; Gandiaga, I.; Crego, J.; Villarreal, I. State of health estimation algorithm of LiFePO4 battery packs based on differential voltage curves for battery management system application. *Energy* **2016**, *103*, 784-796, doi:10.1016/j.energy.2016.02.163.
51. Devie, A.; Dubarry, M. Durability and Reliability of Electric Vehicle Batteries under Electric Utility Grid Operations. Part 1: Cell-to-Cell Variations and Preliminary Testing. *Batteries* **2016**, *2*, 28, doi:10.3390/batteries2030028.
52. Anseán, D.; Baure, G.; González, M.; Cameán, I.; García, A.B.; Dubarry, M. Mechanistic investigation of silicon-graphite/LiNi0.8Mn0.1Co0.1O2 commercial cells for non-intrusive diagnosis and prognosis. *J. Power Sources* **2020**, *459*, doi:10.1016/j.jpowsour.2020.227882.
53. Devie, A.; Baure, G.; Dubarry, M. Intrinsic Variability in the Degradation of a Batch of Commercial 18650 Lithium-Ion Cells. *Energies* **2018**, *11*, 1031, doi:10.3390/en11051031.
54. Schmidt, J.P.; Tran, H.Y.; Richter, J.; Ivers-Tiffée, E.; Wohlfahrt-Mehrens, M. Analysis and prediction of the open circuit potential of lithium-ion cells. *J. Power Sources* **2013**, *239*, 696-704, doi:10.1016/j.jpowsour.2012.11.101.
55. Baure, G.; Dubarry, M. Durability and Reliability of EV Batteries under Electric Utility Grid Operations: Impact of Frequency Regulation Usage on Cell Degradation. *Energies* **2020**, *Submitted*.
56. Waldmann, T.; Hogg, B.-I.; Wohlfahrt-Mehrens, M. Li plating as unwanted side reaction in commercial Li-ion cells – A review. *J. Power Sources* **2018**, *384*, 107-124, doi:10.1016/j.jpowsour.2018.02.063.
57. Yang, X.-G.; Leng, Y.; Zhang, G.; Ge, S.; Wang, C.-Y. Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging. *J. Power Sources* **2017**, *360*, 28-40, doi:10.1016/j.jpowsour.2017.05.110.
58. Schindler, S.; Baure, G.; Danzer, M.A.; Dubarry, M. Kinetics accommodation in Li-ion mechanistic modeling. *J. Power Sources* **2019**, *440*, 227117, doi:10.1016/j.jpowsour.2019.227117.
59. Baure, G.; Devie, A.; Dubarry, M. Battery Durability and Reliability under Electric Utility Grid Operations: Path Dependence of Battery Degradation. *J. Electrochem. Soc.* **2019**, *166*, A1991-A2001, doi:10.1149/2.0971910jes.
60. Baure, G.; Dubarry, M. Synthetic vs. Real Driving Cycles: A Comparison of Electric Vehicle Battery Degradation. *Batteries* **2019**, *5*, doi:10.3390/batteries5020042.
61. Dubarry, M.; Baure, G.; Devie, A. Durability and Reliability of EV Batteries under Electric Utility Grid Operations: Path Dependence of Battery Degradation. *J. Electrochem. Soc.* **2018**, *165*, A773-A783, doi:10.1149/2.0421805jes.
62. Dubarry, M.; Devie, A.; McKenzie, K. Durability and reliability of electric vehicle batteries under electric utility grid operations: Bidirectional charging impact analysis. *J. Power Sources* **2017**, *358*, 39-49, doi:10.1016/j.jpowsour.2017.05.015.
63. Anseán, D.; Dubarry, M.; Devie, A.; Liaw, B.Y.; García, V.M.; Viera, J.C.; González, M. Fast charging technique for high power LiFePO4 batteries: A mechanistic analysis of aging. *J. Power Sources* **2016**, *321*, 201-209, doi:10.1016/j.jpowsour.2016.04.140.

Peer-reviewed version available at *Journal of Power Sources* **2020**, *479*, 228806; doi:10.1016/j.jpowsour.2020.228806

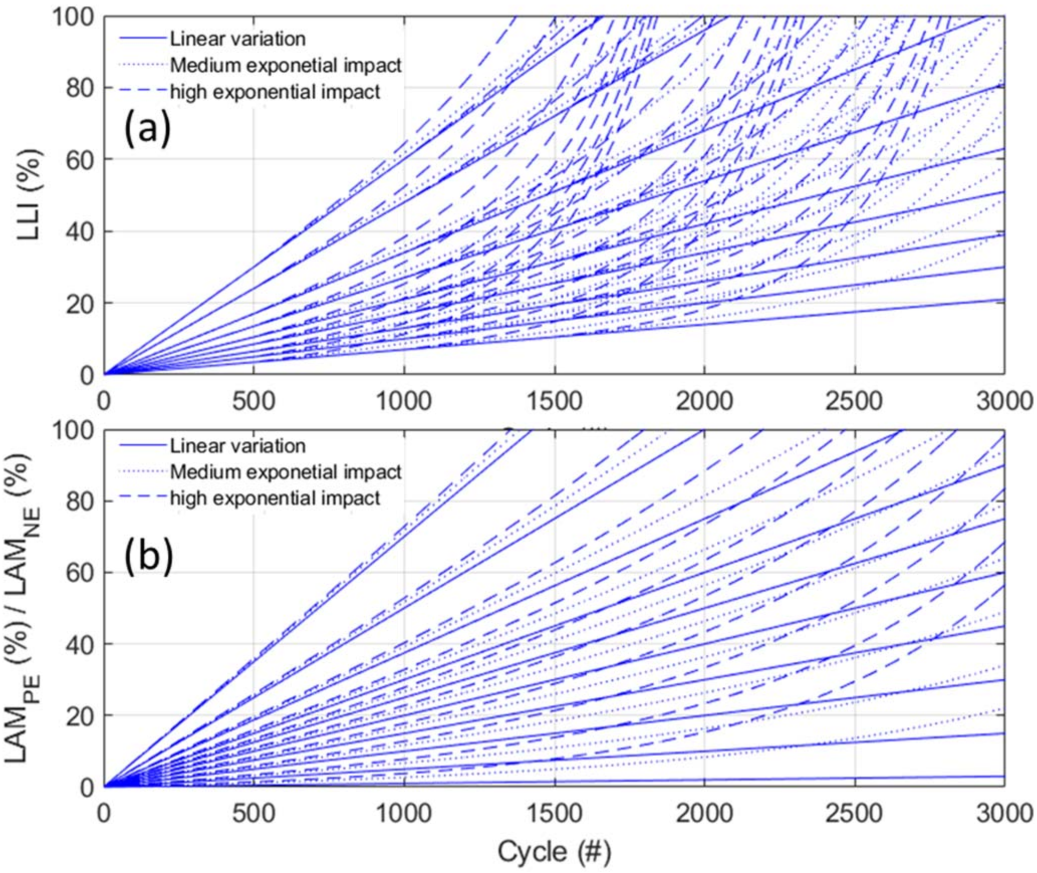
64. Dubarry, M.; Baure, G. Perspective on Commercial Li-ion Battery Testing, Best Practices for Simple and Effective Protocols. *Electronics* **2020**, *9*, 152, doi:10.3390/electronics9010152.
65. HNEI. Alawa central. Availabe online: https://www.soest.hawaii.edu/HNEI/alawa/ (accessed on December 2019).
66. Baure, G.; Dubarry, M. Battery durability and reliability under electric utility grid operations: 20-year forecast under different grid applications. *Journal of Energy Storage* **2020**, *29*, doi:10.1016/j.est.2020.101391.

Supplementary Tables and Figures

Table S1: Summary of scanned parameters for prognosis dataset

Parameter	Description	Values (% per cycle)
p_1	Linear coeff. LLI	0.007, 0.010, 0.013, 0.017, 0.021, 0.027, 0.034, 0.048, 0.06
p_2	Exp. Coeff. LLI	0.000001, 0.002, 0.0033
p_3	Delay exp. LLI	500, 1000, 1500
p_4	Linear coeff. LAM$_{PE}$	0.001, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.0375, 0.05, 0.07
p_5	Exp. Coeff. LAM$_{PE}$	0.000001, 0.001, 0.0013
p_6	Linear coeff. LAM$_{NE}$	0.001, 0.005, 0.0,1 0.015, 0.02, 0.025, 0.03, 0.0375, 0.05, 0.07
p_7	Exp. Coeff. LAM$_{NE}$	0.000001, 0.001, 0.0013
p_8	Plating reversibility	0, 50, 100

Figure S1: Different duty cycles tested for each degradation mode with (a) LLI and (b) the LAMs..



Supplementary Note 1: Diagnosis Dataset Analysis

Figure S2 presents an example of the sensibility of 3 FOIs for capacity loss, LLI, LAM$_{PE}$, and LAM$_{NE}$. The three selected FOIs were the capacity of the first plateau (area under peak ① on Figure S3), the capacity of everything but the first plateau (area under peaks ②-⑤), and the capacity of the last plateau (area under peak ⑤) [20]. As reported in our previous work, [20], the sensibility analysis showcase that no single FOI can track capacity loss nor one of the degradation modes alone since there is always a spread of values. Nonetheless, as reported in [20], their combined variations might be used. This will be discussed in supplementary note 2.

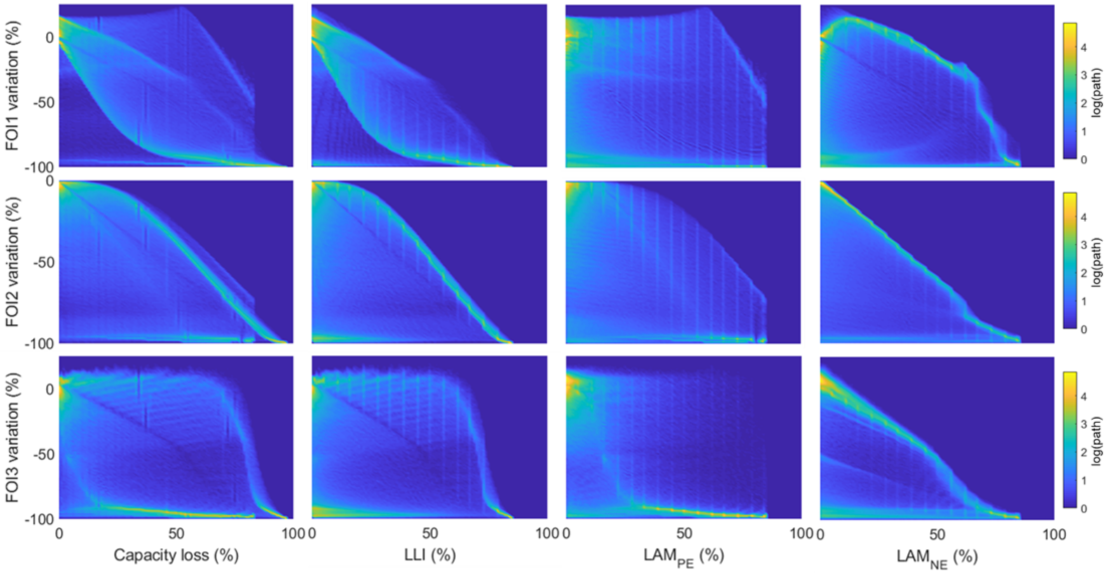


Figure S2: Sensibility analysis for 3 FOI on the Gr//LFP C/25 charges.

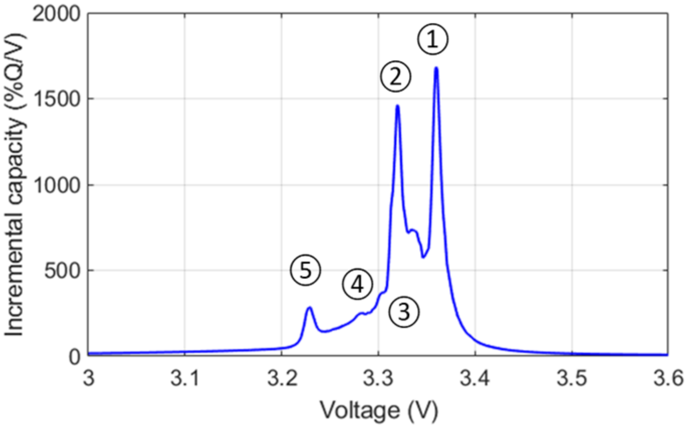


Figure S3: IC signature of a Gr//LFP C/25 charge with the typical peak indexed 1-5.

Supplementary Note 2: Prognosis Dataset Analysis

The prognosis dataset can be used to test the validity of different FOIs proposed in the literature for early prognosis. Figure S4 (a,b) presents the evolution of the variance between cycle one and cycle 100 and 500 respectively as a function of cycle life, defined as the cycle at which 20% capacity loss is reached. When considering a multitude of degradation paths, the variance approach seems less effective than proposed in the literature [5,11] with correlation coefficients of -0.56 and -0.43, respectively, when calculated after 100 and 500 cycles. Looking at the capacity loss after 100 cycles, Figure S4 (c), the correlation is better ($\varrho = 0.63$). Still, the error could be huge for low capacity loss as paths with less than 1% capacity loss after 100 cycles were found to induce 20% capacity loss after just 200 cycles. Finally, as predicted [20,21], the evolution of the first LFP plateau capacity (i.e., the area under peak ① in the Gr//LFP IC curves) is not a good indicator with a Pearson coefficient of 0.51, Figure S4(d). The dataset could be analyzed further by tracking if specific conditions exist where the FOI is accurate, e.g., the first LFP plateau capacity is likely working well when there is little LAM_{NE} [20,21]. This is out of the scope of this publication and will be investigated in future work.

The prognosis dataset can also be used to validate diagnosis techniques. For example, the method proposed in [20] can be tested. As shown in Figure S2 individual FOI information is not enough to reach a proper diagnosis, but their combined variations might be enough. Figure S5 presents the diagnosis results based on the combined variations of 3 FOIs for diagnosis after 100, 200, 500, and 1,000 cycles for the 130,000 computed duty cycles. Out of the more than 500,000 points tested (4 x 130,000), the average errors for estimation of LLI, LAM_{PE}, $LAM_{NE,}$ and capacity loss were of -0.32%, -2.73%, -0,46%, and -0.33% respectively with standard deviations below 1% for all but LAM_{PE} (6%) for diagnosis up to the 500th cycle. The distribution of the results is presented in Figure S5 and the different statistics summarized in Table S2. For all, but LAM_{PE}, most of the errors are comprised between -1 and 1%, with the minimum average error recorded for the diagnosis after 500 cycles and the maximum for the diagnosis after 1,000 cycles. The spread for LAM_{PE} is larger, between -2 and 10%, which was expected since LAM_{PE} is notoriously hard to quantify for LFP cells because of the single voltage plateau [25]. The average LAM_{PE} error decreased with cycle number to 0.74% after 1,000 cycles. Looking more into details, the maximum and minimum errors were always recorded for the diagnosis at 1,000 cycles with maximum and minimum errors mostly below 10% for all the other cycles (except for LAM_{PE}). Future work will investigate the specific combinations that led to large errors, although our previous work [20] showed that it was for degradations unlikely to happen under real-life conditions where LLI usually predominates.

Peer-reviewed version available at Journal of Power Sources 2020, 479, 228806; doi:10.1016/j.jpowsour.2020.228806

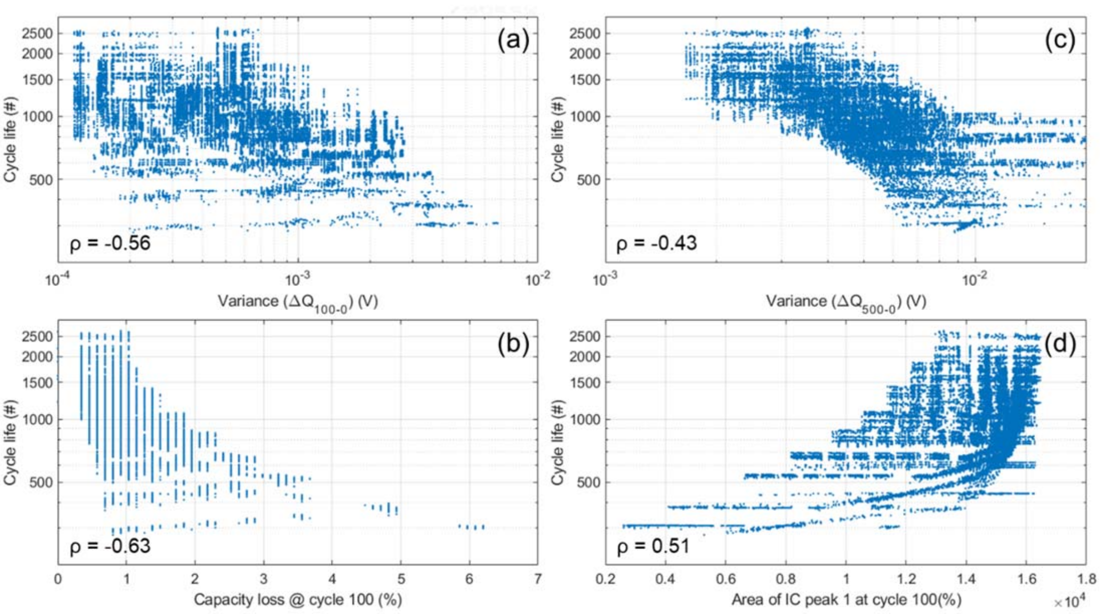


Figure S4: Evolution of the (a) the variance between cycle 100 and 1, (b) the variance between cycle 500 and 1, (c) the capacity loss at cycle 100 and (d) the area of the high voltage IC peak at cycle 100 as a function of cycle life (i.e., cycle at which 20% capacity loss is reached).

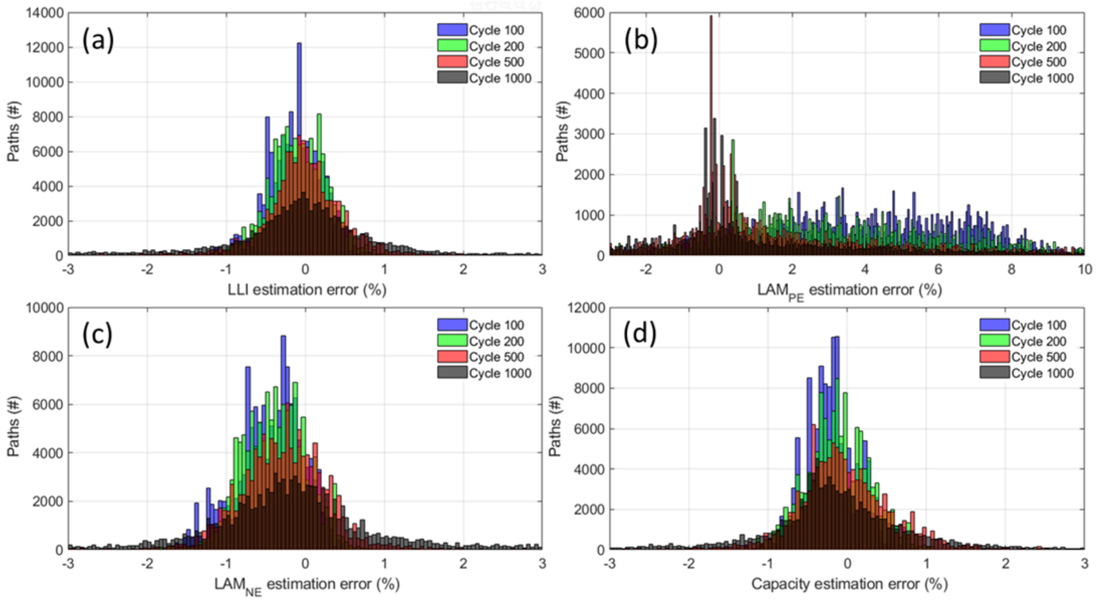


Figure S5 Distribution of the estimation errors for (a) LLI, (b) LAMPE, (c) LAMNE, and (d) the capacity loss based on the evolution of 3 FOIs at cycle 100, 200, 400, and 1000.

Table S2: Main error with standard deviation as well as minimum and maximum error for the three degradation modes and the capacity loss.

	Mean (%)	**Std (%)**	**Minimum (%)**	**Maximum (%)**
LLI	**-0.32**	**2.39**	**-71.86**	**39.31**
cycle 100	-0.12	0.34	-1.53	0.87
cycle 200	-0.08	0.36	-1.42	1.45
cycle 500	-0.02	0.79	-8.73	9.39
cycle 1000	-1.13	4.75	-71.86	39.31
LAMPE	**2.73**	**5.82**	**-61.35**	**63.60**
cycle 100	4.17	2.71	-3.66	13.07
cycle 200	3.47	3.47	-8.82	21.05
cycle 500	2.31	5.84	-34.10	41.15
cycle 1000	0.74	8.98	-61.35	63.60
LAMNE	**-0.46**	**2.36**	**-49.91**	**46.52**
cycle 100	-0.45	0.42	-1.65	0.85
cycle 200	-0.39	0.40	-1.97	1.00
cycle 500	-0.28	0.67	-11.49	4.28
cycle 1000	-0.75	4.78	-49.91	46.52
Capacity loss	**-0.33**	**2.37**	**-60.43**	**43.29**
cycle 100	-0.18	0.35	-1.32	1.16
cycle 200	-0.10	0.41	-1.51	1.86
cycle 500	-0.03	0.92	-10.20	10.68
cycle 1000	-1.04	4.69	-60.43	43.29

The results from Figure S4 raise the important question of whether a single point early prognosis is possible. With the diagnosis approach validated, a possible solution could be to use different diagnoses and extrapolate the evolution of the different parameters to reconstruct the voltage curves using the mechanistic approach, Figure S6. This technique is already used to forecast the evolution of capacity loss for large experimental studies [59,61,66]. Further work is in progress to affine the technique further and to understand under which conditions the estimation is the most accurate.

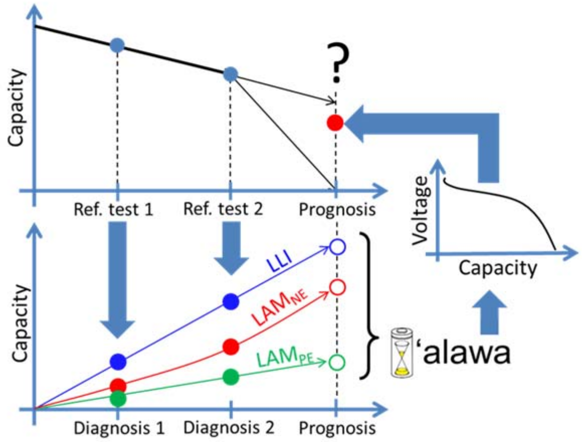


Figure S6: (a) Schematic representation of possible early prognosis solution using the mechanistic approach.