

The challenge and opportunity of battery lifetime prediction from field data

Valentin Sulzer¹, Peyman Mohtat¹, Antti Aitio², Suhak Lee¹, Yen T. Yeh³, Frank Steinbacher⁴, Muhammad Umer Khan⁵, Jang Woo Lee⁶, Jason B. Siegel¹, Anna G. Stefanopoulou¹, and David A. Howey^{2,7,†}

¹University of Michigan, Ann Arbor, MI, USA

²University of Oxford, Oxford, UK

³Voltaiq Inc., Berkeley, CA, USA

⁴Siemens AG, Erlangen, Germany

⁵Continental Automotive GmbH, HE[A]T Artificial Intelligence, Regensburg, Germany

⁶Samsung SDI Company, Ltd., Yongin-si 17084, South Korea

⁷The Faraday Institution, Harwell, UK

† david.howey@eng.ox.ac.uk

June 8, 2021

Abstract

Accurate battery life prediction is a critical part of the business case for electric vehicles, stationary energy storage, and nascent applications such as electric aircraft. Existing methods are based on relatively small but well-designed lab datasets and controlled test conditions, but incorporating field data is crucial to build a complete picture of how cells age in real-world situations. This comes with additional challenges because end-use applications have uncontrolled operating conditions, less accurate sensors, data collection and storage concerns and infrequent access to validation checks. We explore a range of techniques for estimating lifetime from lab and field data and suggest that combining machine learning approaches with physical models is a promising method, enabling inference of battery life from noisy data, assessment of second life condition, and extrapolation to future usage conditions. This work highlights the opportunity for insights gained from field data to reduce battery costs and improve designs.

Keywords: lithium-ion, battery, lifetime, prognostics, knee point, data, applications

Context and scale

To enable the transition to a clean economy and ensure confidence in energy storage technologies, advances are required in reliability, safety, and extended usage of batteries. While headline-grabbing improvements have been made in battery materials, significant advances may also be achieved in managing behavior via enhanced modeling and real-time sensing. These are often chemistry-agnostic and hence can be coupled with future materials and next-generation chemistries such as lithium metal.

Here we explore how physics-based and data-driven modeling informed by measurements from end-use devices enables new battery lifetime models. Although challenging, this will lead to reduced costs by reducing the battery size needed to satisfy warranties and guarantee performance. It will also elucidate degradation mechanisms, improving safety and reducing downtime by enabling appropriate intervention. Finally, it will inform decisions on second life, enabling a circular economy for batteries.

1 Introduction

Batteries are used in a wide variety of applications, from consumer electronics to electric cars, rail, marine, and grid storage systems. A critical need for consumer acceptance in electric vehicles is to achieve longer range and lower cost via pack size reduction [1, 2]. All of these objectives depend on accurate state of health (SOH) estimation and predictions of lifetime under various operating conditions. More accurate lifetime prediction improves battery technology at all stages of a battery’s life. First, it can shorten the product development cycle, for example by elucidating failure mechanisms, in particular if models can be incorporated in a closed loop with experiments [3]. Second, it can be used to optimize manufacturing protocols. Third, improved lifetime prediction can lead to lower warranty and insurance costs, timely preventative maintenance, lower up-front capital cost by reduced over-engineering, and better control of charging and discharging that could prolong life [4]. Finally, it leads to improved prospects for second life applications—supporting the creation of a circular economy around battery manufacturing, re-use and recycling will also be critical as demand is forecast to outpace raw material supply and refinement over the coming decade [5].

The criteria for determining end of life may vary by application, but generally this occurs when the battery can no longer meet the requirements of range, operating time, or maximum power capability under typical usage profiles. The key parameters that affect end of life are capacity (available energy) and internal resistance (available power) [6]. Battery aging depends on intrinsic factors such as manufacturing variability and pack design, and extrinsic factors such as temperature and intensity of usage [7, 8], and is therefore difficult to predict, particularly outside of the laboratory [2]. Existing reviews and perspectives in the litera-

ture consider methods for state of health prediction [9–15], lifetime prediction [9, 14, 16, 17], and/or fusion of physics-based and data-driven models [17–19], but are historically restricted to fairly small datasets under very controlled conditions. Complementary to existing literature, this perspective examines the unique challenges of battery lifetime prediction with field data and for second life applications, and reviews which approaches are most promising for addressing these challenges. This first requires a review of methods using lab data (Section 2). We then lay out the challenges and assess the promising methods for field data analysis (Section 3), where we discuss the additional value that field data offers for lifetime prognostics and the difficulties in obtaining and processing this data. Finally we address high-throughput testing for second life asset evaluation (Section 4).

In the lab setting, which is the best understood and most studied in the literature, the cycling pattern of the batteries can be closely controlled and regular reference performance tests (RPTs) can be performed to quantify health. However, field data from real-world applications exhibits irregular cycling patterns, varying operating conditions, and path-dependent degradation mechanisms, making reliable predictions difficult. This setting is extremely relevant for industrial needs, such as prediction of the remaining useful life of a customer’s electric vehicle or compliance with warranty conditions for grid storage systems, but prognostics using real-world data remains an open research challenge.

Industry sectors such as automotive manufacturers have narrow profit margins and comprehensive certification requirements that necessitate extensive laboratory testing, therefore gathering fleet data may be an additional cost and effort that is hard to justify. Since test data are often not available across the wide range of cells used in packs commercially, a basic set of lab degradation measurements is beneficial as a starting point for understanding the impact of operating parameters on degradation. Additionally, given that automotive cells are considered degraded when they reach 80% SOH (i.e. of their initial capacity), there is a need for capacity and resistance estimation accuracy of at least 5% and ideally 2% to underpin lifetime prediction. Defining the accuracy and confidence levels necessary for a health-conscious battery management system is still an urgent research goal [20].

To build an accurate, general model of battery behavior that covers many usage conditions, a large amount of aggregated data from a population of users is required—it is insufficient only to do this on an individual end user basis. The information that is gathered from intelligently tracking degradation at the fleet level in the field could be used to improve user experience for individual battery end users via over-the-air software updates although there may be regulatory barriers inhibiting this. Finally, at the end of first life (e.g. in an electric vehicle), a battery may be assessed for possible second life application (e.g. grid storage). Estimating health at this point comes with additional difficulties, such as potential lack of historical data and a change

in the future aging mechanism of the battery due to different operating conditions in second life. Controlled RPTs are possible and can be designed around screening and techno-economic analysis of batteries for second life, but are time consuming and require equipment and space that translates to cost, disadvantaging the economics of re-purposed batteries.

To address the challenge of lifetime prediction, three general approaches exist—empirical aging maps, data-driven models, and physics-based models—and later in this paper we also introduce a fourth technique that combines physics-based and data-driven models. Empirical aging maps model capacity fade explicitly as a function of time or charge-throughput, parametrized by operating conditions such as temperature, C-rate, and depth of discharge [7, 21–25]. In the purely data-driven approach, very few assumptions are made about the underlying principles governing the behavior of the battery, and machine learning models are trained using raw input signals (current, voltage, temperature) [26–36]. Another type of data-driven method uses preprocessed features from the voltage, current, temperature, impedance, or power curves as inputs to machine learning models [36–45, 45–54]. Finally, in the physics-based approach, models are constructed from first-principles, with tuning parameters found using a relatively small number of experiments. Such models include differential voltage analysis models [55–60], equivalent circuit models (ECMs) [7, 61–63], and first-principles degradation models based on porous electrode theory [8, 64–70].

Unfortunately, neither the data-driven nor physics-based methods alone can solve the challenge of battery lifetime prognostics from field data. The challenges facing physics-based modeling have been well documented, such as a large number of coupled and nonlinear degradation mechanisms [8], which evolve nearly unobservably from the electrical measurements and so are difficult to parametrize. As a result, not all researchers agree on which degradation mechanisms to model and how to implement them, since many different mechanisms and formulations can explain similar observed degradation behavior [8]. We note that mechanical [71] or acoustic [72] measurements have been proposed as a means to address the lack of observability, but such measurements are not yet widely used.

Meanwhile, data-driven approaches suffer from the ‘curse of dimensionality’, where the amount of data needed to capture all combinations of operating conditions grows quickly with the number of conditions being investigated. This is compounded by the relatively slow rate at which battery lifetime data can be acquired, taking several months or years of experiments for each change in chemistry (e.g. electrolyte additive), form-factor, or manufacturing process. Furthermore, trade secrecy limits how much data is available to individual researchers. The inherently non-linear, path-dependent nature of battery degradation further exacerbates this problem [59, 73–76].

Here, we start by introducing different approaches for life prediction in lab settings (Section 2). The

advantages and disadvantages of each method in terms of computational complexity, data requirements, and accuracy are also discussed. Then, the challenges and opportunities of applying life prediction methods to field data are presented (Section 3), and we suggest that hybrid methods that combine physics-based and data-driven approaches show promise for this because they combine accuracy, robustness to limited or low-quality data, and generality (Section 3). Finally, in Section 4, we discuss the challenges of lifetime estimation for second life applications. These include assessing viability of batteries without having historical data available, in particular determining whether or not the battery has degraded beyond the knee point (i.e. the point beyond which degradation may accelerate towards end of life), and understanding the cost-benefit between testing re-purposed batteries versus the increased revenue from more accurately knowing their SOH.

2 Life prediction from lab data

In a laboratory, battery cycling can be repeated consistently, and conditions such as current and temperature can be closely controlled (Fig. 1). Measurements of ‘ground truth’ battery health—capacity and resistance—can easily be taken with regular RPTs as required, and batteries can be cycled continuously until end of life, which is often specified as the point where measured capacity reaches 80% of the pristine cell capacity [77].

This means that laboratory tests are useful to build models that explore how different operating conditions affect cycle life, and hence determine warranties, maintenance schedules, and system sizing. The operating conditions used in laboratory tests include driving patterns and schedules tailored to very specific applications, such as a specific electric vehicle model. Laboratory tests can also be used to guide the development of new battery chemistries, optimize battery design, and improve manufacturing processes, for example by investigating the effect of parameters such as humidity or formation cycles on performance.

Battery degradation testing is a lengthy process, therefore extreme operating conditions, such as high C-rates or elevated temperatures, are often used to accelerate aging. Even with accelerated aging, it can be slow to assess the degradation impact of individual manufacturing parameters such as materials and processing choices, design factors such as cell size, number of layers and electrode thicknesses, and formation protocols. Additionally, since different manufacturing parameters interact non-linearly, varying each parameter individually may not tell the full story. Therefore domain-knowledge is required to ensure that testing efforts are as effective as possible.

Significant research has been undertaken on modeling lab battery test data, and it can broadly be divided into four categories, summarized in Fig. 2: empirical aging models, pure data-driven methods, feature-based

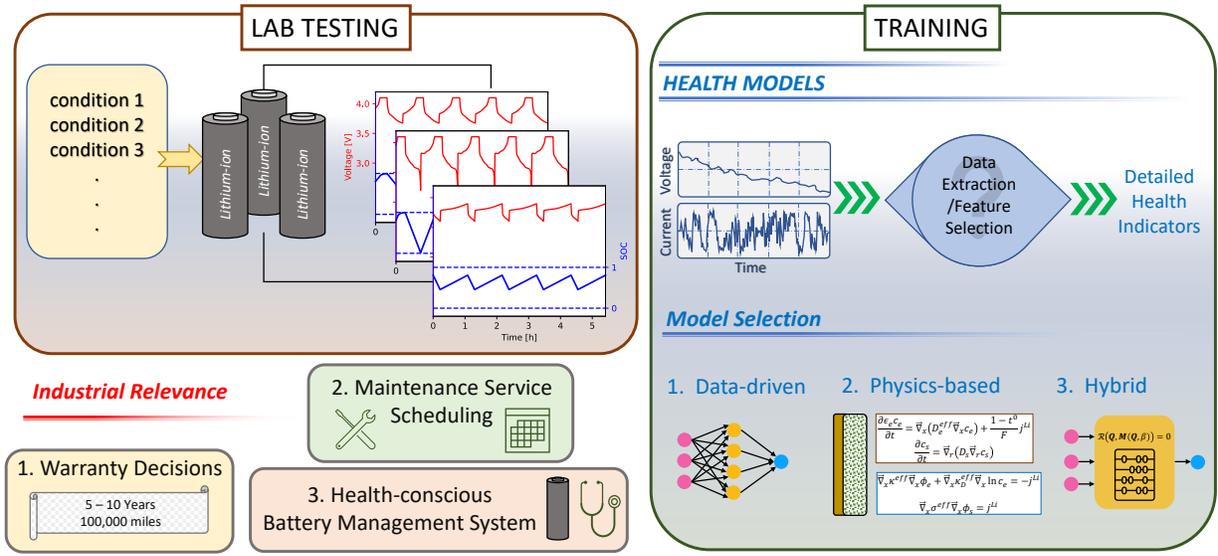


Figure 1: Summary of approaches for prognostics using lab data. A variety of use cases can be tested and health models trained using data-driven, physics-based, or hybrid paradigms. These models are typically developed to underpin warranties, but are not validated beyond the warranty period or in other usage scenarios.

data-driven methods, and physics-based methods. In this section, we review state-of-the-art approaches within each of these categories.

Empirical aging models. A typical measure of battery degradation is the capacity fade curve, which describes how the capacity changes as a function of charge throughput, equivalent cycle number, or time. Therefore the simplest approach for lifetime prediction is to build an empirical model of the capacity fade parametrized by operating condition. This may depend on a number of factors including time and charge throughput. As first proposed by Bloom et al. [21], such empirical models usually have a square-root-of-time dependence, due to diffusion-limited solid electrolyte interphase (SEI) formation, and Arrhenius kinetics for the temperature dependence [7, 21–25]. Further refinements include accounting for C-rate [22–24], average state of charge (SOC) [23, 24], depth-of-discharge range [7, 24], and voltage [7, 25]. Furthermore, recent models separate capacity fade into calendar aging, which depends on time, and cyclic aging, which depends on charge throughput [7, 23, 24].

This approach is simple and easy to implement, making it common in industry for developing maps of lifetime from lab data. However, it has a few limitations. First, very large amounts of data are required to interpolate over all operating conditions. Second, this approach cannot easily account for cell-to-cell

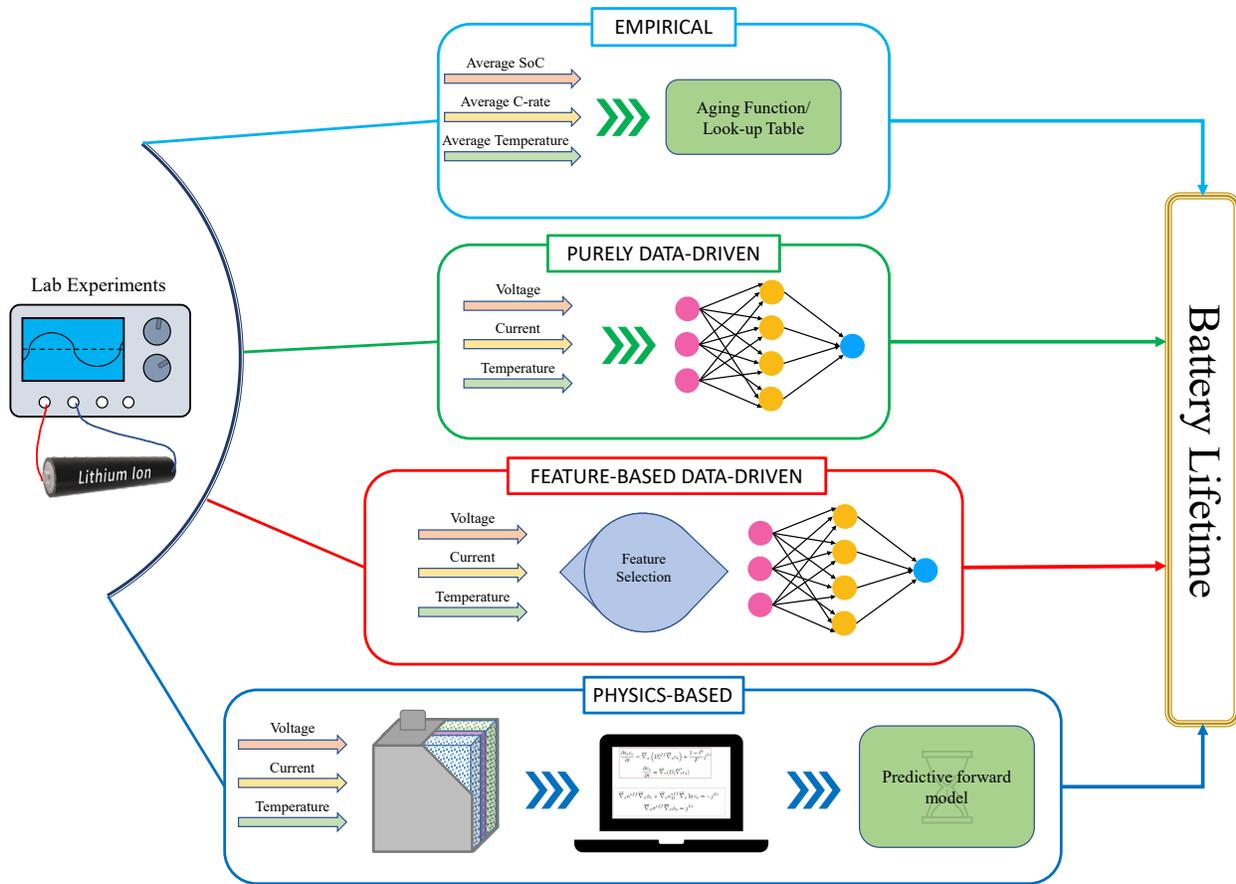


Figure 2: Empirical, purely data-driven, feature-based data-driven, and physics-based approaches to battery lifetime prediction.

variations due to manufacturing or heterogeneous current and temperature distributions within a pack. Third, separate maps are usually developed for cycle and calendar aging, which misses the interactions between these [73], whereas the methods that we will review below can incorporate these interactions as long as appropriate data is included in the training sets. Finally, empirical aging maps may fail to capture knee points in the capacity fade curve [7, 24]. Knee points are a change in the degradation rate likely caused by a change in the underlying mechanism [78] (e.g. SEI growth later leading to lithium plating [66]), and they are a particular challenge for prediction, especially with simple empirical models. To overcome these limitations, prognostics can be improved by also taking into account the individual raw time-series data (such as voltage) of the cells during cycling, instead of just the capacity obtained during the RPTs.

Purely data-driven models. This approach consists of using measurements such as the current and voltage directly as inputs to a machine learning model in order to learn the remaining useful life as the output.

While several studies have shown promise in estimating the present state of health of the battery [26–32, 35, 79], to date there is limited application of these methods to prognostics, with one example being Zhang et al. [33], who use raw electrochemical impedance spectroscopy data to predict remaining useful life (RUL). This may be due to the lack of available data: while each cell has hundreds of cycles to use as training data for SOH estimation, its cycle life is just a single data point per cell.

Another possible approach is to train a machine learning model for ‘forward simulation’, where the model learns how much capacity fade occurs during short intervals based on the existing capacity, current or temperature during the interval [34–36]. Then remaining useful life can be predicted by adding together the capacity fade from all the intervals under typical usage conditions and seeing where the resulting trajectory crosses 80%.

Feature-based data-driven models. In this approach, there is a preprocessing step in which features are extracted from the voltage and current, informed by physical understanding of the cell’s behavior. These features are then used as the inputs to machine learning algorithms. Similarly to pure data-driven methods, much of the literature focuses on present SOH estimation rather than prediction of future SOH [36–44, 46]. However, since the models that take features as inputs are less complex than those that take raw data as input, they require less training data than pure data-driven approaches, and so are more readily applicable to lifetime estimation. For example, Chinomona et al. [47] use various statistical features from the voltage, current, and temperature, while Yun et al. [48] and Greenbank and Howey [49] use as features the time spent within certain voltage, current, power, and temperature ranges.

While those studies use features from the current or voltage within a single cycle, Severson et al. [52] showed that using features generated from changes between different cycles can give very accurate predictions of lifetime, even using a simple regularized linear regression model. Fermín-Cueto et al. [53] improved the accuracy of this method (lower prediction error with fewer cycles) by using more features and a more advanced machine learning algorithm. In a previous paper [54], we applied this method to a different dataset for NMC/graphite cells, achieving accurate prediction of battery lifetime, and showed that these features are closely correlated with loss of lithium inventory in the cells, demonstrating the importance of a physical basis for the features.

Accurate data-driven prediction of the knee point in the capacity fade curve using feature-based data-driven models has recently been demonstrated [49, 53, 80], although data-driven methods may fail to extrapolate correctly under changing usage conditions due to path-dependence [59], and may not even correlate with aging unless a certain depth of discharge is reached during usage [54].

Finally, some feature-based ‘forward simulation’ approaches have also been proposed, using physics-informed features as inputs to Gaussian processes to predict incremental capacity fade and hence remaining useful life under typical operating conditions [45, 50, 51].

Physics-based models. This broad category includes atomistic models, continuum approaches based on porous-electrode theory, through to equivalent circuit models. We include equivalent circuit models here as ‘lumped’ physics models using electrical engineering components to capture the electrochemical behaviour [63, 81]. In terms of degradation, circuit models can be used to estimate and track parameters empirically as cells age [7]. Examples of lifetime prediction using circuit models are Zhang et al. [61] and Chu et al. [62] who employ observers to identify the internal states of an ECM, and then fit an empirical aging map to find future changes in capacity. A challenge with these approaches is that careful tuning is required to achieve robust performance. To address this, Aitio and Howey [82] show that applying Gaussian process regression to identify functional dependencies of model parameters can give smoother and more dependable results when using drive cycle data. Equivalent circuit models are popular since they are easy to implement and parametrize, although recent work has shown that similar computational efficiency can be achieved with models based on porous-electrode theory [83–85].

The other relevant category of physics-based models is those derived from first principles—typically continuum approaches using porous-electrode theory. These may be extended to capture the underlying degradation mechanisms that cause capacity fade, such as SEI layer growth [8, 65–70, 86–88], lithium plating [8, 65–67, 89, 90], and particle swelling and cracking [8, 91]. These models can be used to directly simulate the entire life of the cell under certain operating conditions [8, 64–70]. In theory, such models could be used for prognostics by parametrizing them with early-life data for a particular cell and then simulating the remaining life of the cell and seeing where 80% capacity is reached. An advantage of this approach is that it also provides the predicted cause of failure, enabling remediation strategies before end-of-life such as tightening of safety limits in the battery management system.

The main weakness of these models is their lack of flexibility and parametrization difficulty. While excellent agreement with the available data can be achieved, it is not possible for a mechanistic model to account for every single eventuality. For example even when a large number of degradation sub-models are included, some aspects of the experimental dataset still cannot be fitted accurately [8]. In the following section, we suggest ways in which physics-based models can be made more flexible (and hence model a wider range of degradation of mechanisms) by augmenting them with methods from machine learning.

Accurate parametrization of continuum physics-based models for a fresh cell is challenging, requiring

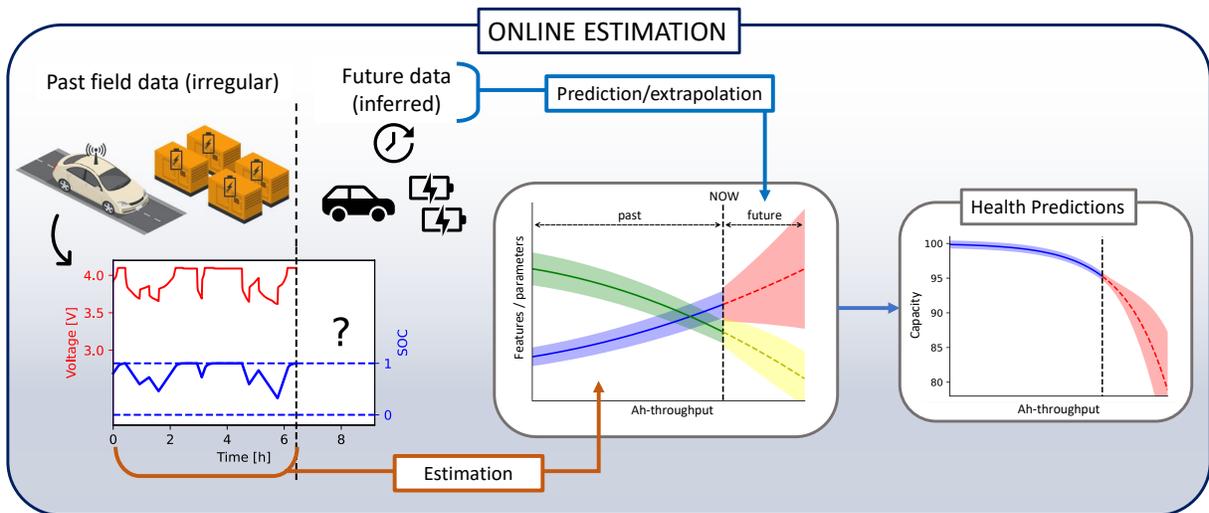
cell teardown and specialized testing equipment [92, 93] to determine a wide range of parameters such as reaction rate constants, conductivities, diffusivities, particle sizes etc. Determining the degradation parameters, such as SEI kinetic parameters, by inverse modeling is even more challenging due to the computational time required to simulate the entire lifetime (e.g. 15 minutes with Ouyang et al.’s relatively simple model [64], compared to microseconds for an empirical aging map). For physics-based models to become more widely adopted for lifetime estimation, this computational time needs to be reduced, either through the development of reduced-order models [83] or numerical methods. Advancing efficient physics-based models is one of the goals of the open-source battery modeling framework PyBaMM [94].

Physical understanding can also be used to define more specific degradation metrics than capacity or resistance. For example ‘degradation modes’, such as loss of active material and loss of lithium inventory, may be estimated from data or models and linked to different underlying mechanisms, such as SEI growth or lithium plating [60]. These can be identified using differential voltage analysis [55–59], and used for prediction, e.g. Hui et al. [95] use differential voltage analysis to identify degradation modes, then predict future change in these degradation modes—achieving better accuracy than empirical aging maps.

3 Life prediction from field data

Lab battery testing is limited in the number of test channels available and the time available for tests. Ultimately what matters is battery performance in real applications. If field data from batteries in end-use applications could supplement lab performance and lifetime tests, this would significantly increase the amount of data available, accelerating our understanding and closing the gap between lab and end-use and ensuring that lifetime prediction algorithms are relevant to industry applications. Fig. 3 summarizes the aims and impacts of lifetime prediction from field data.

Field data versus representative drive cycle testing in the lab. One existing approach to bridge the gap between standard lab tests and field data is to test batteries using representative loading patterns in the lab. To this end, many researchers have characterized ‘typical’ user driving patterns [2, 96, 97]. The advantage of this approach is that it can be performed in the lab with high-accuracy equipment, controlled conditions, and frequent characterizations. However, complementing this with field data is still very valuable for a number of reasons. First, there are never enough testing channels in the lab, and they are never available for long enough to test all combinations of conditions necessary—whose number grows factorially with the number of aging parameters. In contrast, field data may be available at relatively low cost since the cells are already



Impacts on Total Cost of Ownership

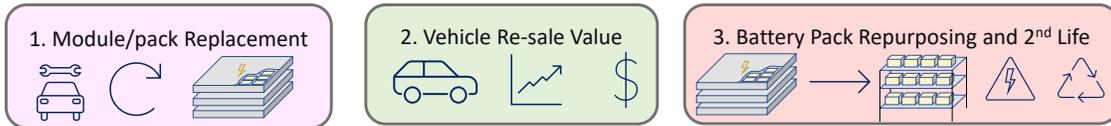


Figure 3: Summary of the aims, process, and impacts of state of health prediction from field data. Challenges include noisy and missing data, uncontrolled partial cycling conditions, lack of regular validation tests, and uncertainty over future usage.

being cycled, and (by definition) the entire range of operating conditions is covered.

Second, lab testing is contrived and different to real world usage. Due to constraints of both time and testing equipment availability, lab tests usually impose extreme conditions and short rest periods between cycles. This can lead to under-prediction of battery lifetime, and hence over-engineering of pack design. Accelerated aging experiments often stress the battery more intensively than standard cycling in the field. This is accounted for by experimentally-developed maps between accelerated experiments and real usage scenarios, but these maps take a long time to obtain, because they require replicating real usage scenarios in the lab, and even then it is difficult to perfectly relate lab tests to field data. For example, it has been shown that rest periods between cycles can be beneficial for the overall life of the battery [73], but lab experiments cannot always incorporate long resting periods due to testing time constraints. For extension of model validity from lab accelerated aging tests to real-world usage, it is important to standardize methods to interpret not just accelerated cycle aging data, but also accelerated calendar aging and mixed cycle plus calendar aging scenarios. For both lab accelerated aging and field data aging research, it is important to rigorously understand the impact of cycles-per-day and rest-time conditions on lifetime.

Third, there could be exogenous factors in the field which are not accounted for in the lab, such as seasonal temperature variations, or mechanical vibrations causing failure. Fourth, from the point of view of understanding battery lifetime, which depends on both intrinsic (manufacturing variability) and extrinsic (usage) factors, more data is always valuable because it increases the statistical confidence with which we can build lifetime and performance models.

Challenges and opportunities of field data. Three main issues arise in this context: the wide variety of uncontrolled usage scenarios, data quality issues, and lack of clear validation data. First, unlike lab-based prognostics where operating conditions are held constant or at least controllable, in the real world, operating conditions can vary throughout a cell's life. Examples for electric vehicle batteries include temperature differences due to seasonality and cell-to-cell variations within a battery pack, or varying SOC windows associated with trip distances. Depending on driving conditions, charging and discharging current/power levels can affect battery degradation to a greater or lesser extent. Recent studies have shown that battery degradation may also depend on the order of the operating conditions, a phenomenon known as 'path-dependence' in the literature [59, 73–76]. Another source of uncertainty in lifetime predictions arises because future operating conditions themselves are uncertain.

Second, field datasets are difficult to work with, for several reasons. Cell-to-cell variations within a pack or modules, less accurate sensors, noise, extreme C-rates, smaller coverage of the available range of

conditions (e.g. microcycling of SOC), missing or false data, and rapidly varying currents make it very challenging to extract features, or fit parameters [54, 98]. There may also be long periods of missing data, for example temperature data while the battery is at rest (and hence not recording). These factors also restrict the identifiability of both mechanistic and data-driven models. For physics-based models, the input signal is not rich enough for parametrization, and therefore simplifications of the model are required. On the other hand, performance of purely data-driven methods is limited by the extent that the training data covers the possible space of inputs, since extrapolation outside of the observed data comes with significant uncertainty.

The third difficulty for real-world prognostics is the challenge of verification or validation, i.e. in field data there is typically a lack of ground truth against which to compare predictions. This is due to the lack of controlled RPTs, which makes capacity and end-of-life difficult to validate, and the fact that not all cells are cycled to end of life, reducing dramatically the data fidelity and availability. A related issue is how often to retrain a prognostic model, where a balance needs to be struck between having sufficiently rich data versus capturing gradual changes in usage and operating conditions.

In the remainder of this section, we review existing studies on this topic, then suggest approaches that can be used to overcome the issues and address the challenge of prognostics in the field. We separate these again into ‘data-driven’ and ‘physics-based’ methods, though most of these methods are hybrid in nature. The challenges in this section represent an exciting new frontier for battery system development and demonstration of the real world application of hybrid data-driven techniques. We finally discuss two practical aspects of implementation, namely data management and uncertainty.

Existing studies using field data In comparison to studies based on lab generated data, existing literature using field data for life modeling is sparse, and is mainly focused on diagnostics rather than lifetime prediction. However, some large databases of electric vehicle usage data have been collected and analyzed, as summarized in Table 1. These existing works focus on data-driven methods to estimate state of health. For example, Song et al. [99] used operating data from 700 electric vehicles to fit a feed-forward neural network using features calculated from usage history as inputs to estimate capacity, where the target output itself was estimated from partial charging curves. Alternatively, Wang et al. [100] used a larger dataset from 8032 electric taxis to track cell-to-cell variability at the pack level over time. The same data source was used by She et al. [101] to estimate capacity loss in 18 electric city buses by fitting a neural network with inputs consisting of usage features and the output of the peak value in the incremental capacity curve during charging. A different approach was proposed by Huo et al. [102], who installed data collectors on 16 electric taxis and performed periodic capacity calibration tests to obtain validation data for probabilistic

Ref	Dataset used	Aim	Method
[99]	700 B/HEVs, Shanghai	SOH estimation	Neural network
[100]	8032 EVs, Beijing	Evolution of cell variability	ECMs & Regression
[101]	18 EVs, Foshan	SOH estimation	ICA & neural network
[102]	16 EVs, Beijing	SOH estimation	Empirical model & Bayesian network

Table 1: Summary of existing state of health estimation papers using real-world pack-level battery data.

SOH estimates obtained through a Bayesian network with an empirical degradation model. These works give a glimpse of what is possible using large datasets, and they also encounter the challenges mentioned in this perspective, such as missing and noisy data and lack of validation data. These papers do not make their data and code available for others to download, so it is not possible to reproduce their results. The lack of availability of large publicly-accessible datasets, including validation data and real world operating conditions, remains a severe limitation for model development and performance quantification. Future studies on field data should strive to make their data and methods openly available [103] where possible to enable comprehensive quantitative comparison of algorithms.

Feature-based data-driven approaches. Data-driven approaches informed by physics show promise for use in field applications, particularly methods that have been demonstrated to work with relatively short bursts of data, data from controlled charging, and feature selection approaches that can handle varying operating conditions.

For SOH estimation, the Gaussian process method proposed by Richardson et al. [44] gives good results with only 10s of constant-current data. Similarly, methods that rely on features from partial SOC windows [47–49] are more likely to work with field data since conditions are more variable. Approaches that are based on extracting features from the battery charging phase (either the constant-current [37, 42] or constant-voltage phase [37–39, 42]) are of interest for field applications, since in some cases, such as electric vehicles, the charging phase can usually be closely controlled while the discharge phase depends on the power demands of the user. However, these methods still require the charge to begin from a small enough SOC for the features to exist, which may not be the case when microcycling (for example, regular short commutes in an electric vehicle). In this case internal resistance may be a more reliable health indicator than capacity.

For prognostics, methods based on incremental updates of capacity fade [35, 36, 45, 50, 51] may work well with field data since they are able to handle rapidly varying operating conditions, usually using some

form of averaging.

To be effective with field data, feature-based data-driven methods should demonstrate that extracted features are independent of instantaneous operating conditions such as C-rate, temperature, or SOC. If this is not the case, then the prognostics algorithm must account for such variations, which requires large amounts of training data. For example, if the map from ‘time-between-voltages’ to RUL depends on the C-rate, then enough tests need to be run to capture this dependency and the number of tests required grows exponentially with the number of such dependencies. In contrast, physics-based models directly encode these dependencies into the model. We now explore how such models can be used for life prediction in the field.

Hybrid physics-based/data-driven models. As mentioned previously, physics-based degradation modeling aims to simulate the physical mechanisms that cause the battery to age. In an ideal world, we could perfectly model each of these mechanisms and hence construct a digital twin of a battery which would tell us exactly how it will age in future.

In practice, most physics-based degradation models that have been proposed are only validated against data from a single operating condition, and studies that attempt validation using several operating conditions indicate that it is difficult to generalize [8]. This suggests that degradation models could be enhanced with data-driven methods in order to accurately predict the wide range of possible aging mechanisms and cell-to-cell variability [104]. To this end, promising frameworks exist, including the field inversion and machine learning paradigm [105], physics-informed neural networks [106–108], neural ordinary differential equations [109], universal differential equations [110], and neural differential equations [111]. A key advantage of these methods, as demonstrated for modeling the spread of COVID-19 by Dandekar et al. [112], is that they extrapolate beyond the training data much more accurately than pure data-driven methods. This suggests that these methods could be particularly well suited for battery lifetime estimation. Such hybrid methods have been applied to battery models by Bills et al. [113], who add unknown degradation terms to the model and then learn these terms using neural networks, and by Tait et al. [114] who use latent forcing functions to improve the accuracy of a simple electrochemical model. For a further in-depth review of these hybrid physics-based/machine-learning frameworks, we point the reader to [17].

Hybrid models outside of these frameworks are also possible. For example, Zhang et al. [61] use an equivalent circuit model to infer capacity from non-constant current data, then use this to parametrize a capacity fade curve. Aitio and Howey [82] use a Gaussian process to accurately infer the SOC dependency of resistance from synthetic drive-cycle data.

These studies demonstrate promising early results, with further research needed in this direction. A key

issue, however, is lack of openly available field data. While hybrid methods will be most useful when applied to field data, they should also be demonstrated using lab data where results can more easily be validated.

Data management. We have demonstrated that collecting and analyzing data from the field can be beneficial, but doing so requires investment. In our view, the costs and benefits of extensive data collection will vary depending on the application and stakeholder. For safety, lithium-ion systems already have extensive voltage, current and temperature monitoring; recording this data and/or uploading it to a server comes at an additional cost, and while small relative to the battery cost, this may still be prohibitive depending on the application business case. However, for those who choose to do this, there could be significant benefits in future (e.g. cell supplier selection, lifetime modeling, improved designs, second life use). In many applications, data is already collected and stored on a server, especially for warranty compliance of grid storage systems which have a lifetime of 10-25 years and may represent substantial (\$10M's) investments. There are indications from industry that more detailed field data collection is very valuable, especially for cell manufacturers. For example, Northvolt has a desire to collect extensive battery data from manufacturing through to end of life [115] including batteries deployed in real products, and we assume this will need to be cell level data. This is a lofty goal, but detailed cell level field data is obviously viewed as valuable for improving cell design and manufacturing. Off-grid energy suppliers are also interested in detailed data collection because their systems are often in rural areas that are difficult to reach, and therefore logistics and preventative maintenance (e.g. battery replacement) planning requires detailed remote monitoring and battery health prediction. Even if the benefit of collecting field data is not immediately clear (due to lack of existing methods to analyze it), historical data is invaluable for the validation of new algorithms.

Managing large volumes of data is an important aspect of dealing with battery field data as the number of cells grows from tens to millions. Significant computational resources are required to store, process, and retrieve data: over the course of its lifetime, a single cell may produce on the order of 100 MB¹ of time series data. This processing must be performed automatically and requires close collaboration between battery engineers and software engineers. Furthermore, battery data can come from a wide variety of sources: design, manufacturing, lab testing, field use, and second life use. To maximize the information that can be gleaned from the data, it is important to integrate all the data from this battery supply and operation chain onto the same platform. In particular, manufacturing metadata including materials and processing information is of equal importance to time series performance data and can be invaluable when trying to

¹assuming a thousand cycles of 1C/1C cycling while recording time, current, voltage, and temperature at a frequency of 1 Hz, and 4 bytes per single-precision data point

understand variations between different cells and modules. It may be that not every single product needs to be monitored in this way, just a sub-sample from which population behavior can be inferred.

Once the raw data has been recorded, it can be processed and aggregated in various ways by the battery management system itself before uploading to the cloud/server. Typically, time series data is recorded for each group of cells in parallel (a ‘brick’), and a module consists of several bricks in series with several modules together making a pack. Data can then be aggregated between these groups. There are two levels of aggregation: compression of time series data, and aggregation between bricks. For the first type, depending on the algorithm, it may not be necessary to upload all of the raw data to the cloud/server, and for each cycle kilobytes of time series data may be reduced to just a few bytes of aggregated data (Table 2), for example by extracting key features. Feature calculation is usually cheap and can be performed by the battery management system, then only the features need be uploaded to the cloud/server. Alternatively, if physics-based (or hybrid) models are simple enough to be run and fitted online, then only the key internal aging parameters need to be recorded at each cycle. In other cases, all the raw data must be uploaded to the cloud/server for processing, but once this has been done it can be archived or, in some cases, discarded.

Aggregation between cells can take several forms. The simplest approach would be to record data at the module level (i.e. a number of cell bricks in series) instead of the brick (i.e. parallel cell block) level, which does not take into account any heterogeneity in the module. Fast sampling at the module level or even polling [116] cell level data for individual cell level SOC and SOH estimation requires fast convergence that can only be achieved with the right excitation potentially during charging [117]. At the other extreme, the data from every brick can be uploaded to the cloud, but this may introduce a lot of redundant data. A good compromise, already done by some manufacturers, is to store only max/min/average brick voltages, as well as module level data, in order to capture the heterogeneity in the module. This requires measuring the data for every brick, which is usually already done for safety reasons. However, storing and transmitting this information to a central server may require firmware changes at the local level, and may require hardware changes too (depending on how the battery management system is configured). Although health estimation and lifetime prediction algorithms may be applied at either the cell, module or pack level, accuracy will likely be lower if only system level data is available [118].

Finally, data processing and storage should take into account privacy issues—for example, an individual’s daily routine could potentially be inferred from the power usage of their electric vehicle. Methods to anonymize and aggregate data without losing insights for lifetime prediction should be developed in order to address these issues.

Method	Data uploaded to cloud	Data per cycle per cell
Empirical aging map	Aggregated C-rate, temperature, SOC window	Bytes
Pure data-driven	Time series data	Kilobytes
Feature-based data-driven	Features	Bytes
Physics-based (offline processing)	Time series data	Kilobytes
Physics-based (online processing)	Degradation parameters	Bytes

Table 2: Size of data to be uploaded to the cloud. Some methods require high computing power and can only be performed in the cloud, so all of the time series data needs to be uploaded. For other methods, some aggregation can be performed by the battery management system, and only the aggregated data uploaded. The type of aggregation depends on the method.

Dealing with uncertainty. Regardless of the chosen method, a key requirement in model fitting and prediction with field data is the rigorous propagation of uncertainty through the estimation algorithms. Due to the constraints imposed by the data, the variance of model parameters in real-world scenarios is inherently large. Illustrated in Fig. 4, the uncertainty of health estimates is carried forward to lifetime predictions, where the compounding of uncertainty over future behavior is complex due to non-linear dependencies on model parameters, path-dependence, and unknown operating conditions. To this end, Bayesian methods for parameter estimation and prediction show promise because they enable principled inclusion of model and observation uncertainty. With these approaches, full parameter posterior probability distributions obtained either by analytical approximation [119] or using Monte Carlo techniques [120] may be used to construct battery performance estimates with realistic confidence bounds.

4 Life prediction for second life

When a battery can no longer meet the capacity or power requirements for its primary application, it can sometimes be re-purposed into a different ‘second life’ application. This is environmentally beneficial and economically viable, especially with the growing market for lithium-ion batteries for electric vehicles, which can then later be re-purposed for grid or stationary storage applications [122].

The challenge here is to predict the remaining useful life of the battery under its intended new application, which is likely very different to the previous use, in order to decide whether such new usage is viable and inform warranty decisions. There are two particular issues: Firstly, historical data to audit first life usage may be unavailable, for example due to commercial secrecy, and if available it may not be of the

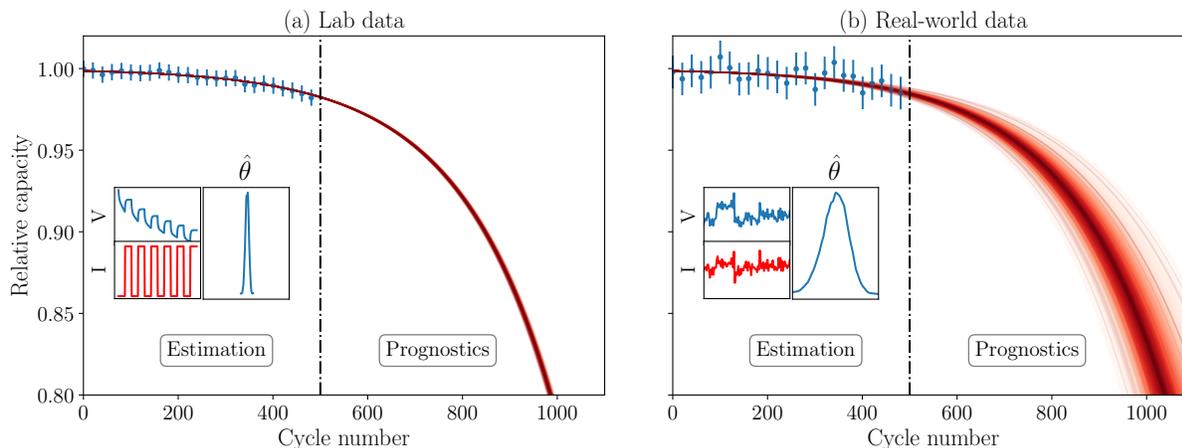


Figure 4: A key challenge in life prediction using field data is managing uncertainty. Controlled conditions and precise measurements in lab conditions can not be replicated in real-world operating environments. This results in less accurate and lower precision estimates of model parameters, $\hat{\theta}$, which have to be accounted for at the prognostic stage. Plots generated with PyMC3 [121].

required reliability or fidelity—though initiatives such as the Battery Passport [123] are being implemented to promote reliable sharing of data. In some cases, time series data from first life and manufacturing metadata are sold with the battery at a premium. Secondly, operating conditions can vary significantly from first life to second life, for example grid applications will typically have different thermal management systems and reduced SOC windows and C-rates compared to EVs, so a model of performance from first life may not be valid for second life. However, it is possible to perform controlled cycles at the re-purposing stage, such as slow deep discharges to determine capacity, pulse tests to determine internal resistance, or electrochemical impedance spectroscopy to determine the internal parameters of a model. The RUL estimation must account for the fact that the usage patterns in the second life will be different (usually, less intensive) than they were in the first life. Therefore, even if historical usage information is available, prognostics algorithms must be more advanced than simply extrapolating aging patterns and models from the first life.

Assessing state of health for second life. At its simplest, assessing the second-life viability of a particular battery pack consists of measuring its SOH, for example capacity under constant-current discharge at various C-rates. However, this may not tell the full story about the second-life viability of the pack, due to the intrinsic variability and path-dependence of capacity fade [73]. Cells that have the same SOH at a point in time may degrade at different rates in future even under the same operating conditions. In particular, Martinez-Laserna et al. [124] show that it is crucial to determine whether or not the battery has passed

the knee point in the capacity fade curve, since degradation does not slow down after this, even under less intensive usage conditions. If historical data are unavailable, it is not at the moment clear how to determine whether the cell is past the knee point. This is an interesting challenge problem for the community.

To address this, advanced RPTs may be possible to estimate parameters of a physics-based model at a level of detail that goes beyond just estimating capacity and resistance. It may be the case that some parameters of a physics-based model correlate well with whether or not the battery is past the knee point. Finding such parameters, and ways to easily identify them, would be extremely useful for assessing the second-life viability of individual batteries. In this context, for porous-electrode models, Park et al. [125] suggest a framework for devising current excitation profiles that maximizes the observability of the underlying model parameters, using a sensitivity analysis and Fisher information matrix. Alternatively, electrochemical impedance spectroscopy may be used to estimate parameters of a physics-based model [126], although without a reference electrode there may be ambiguity in attributing parameters to one or the other electrode. Differential voltage analysis is also a particularly promising framework here, since performing deep low C-rate charges or discharges is possible, and the degradation modes that are identified by differential voltage analysis have been shown to correlate well with knee points [59, 127]. A challenge with all of these techniques is that they are usually undertaken on individual cells—implementation at pack level is an open issue.

Cost-benefit analysis for second-life testing. When considering what tests should be performed to assess second-life viability, there is a cost versus accuracy trade-off, and the level of accuracy required for second-life assessment depends on the intended usage. A very simple assessment could consist of a binary go/no-go classification of whether a battery will be usable for the desired second-life application and duration, in which case relatively low test accuracy is sufficient. A more complex test could also identify the operating conditions that will maximize the battery lifetime in second life, such as SOC window and C-rate, but such tests will require higher accuracy and longer duration. If pack-to-module balancing is planned, very accurate tests are required in order to identify the strongest and weakest packs [128]. In each case, the cost of testing should be no greater than the additional revenue generated by improved test accuracy. While the cost of testing is relatively easy to quantify, calculating the revenue resulting from improved control is an open research question [4].

Finally, an important question to consider in second-life testing is cell-to-cell variability within a pack. It has been shown that the labor involved in opening up a pack is more expensive than the benefit from testing and screening the individual cells [129]. Therefore, second-life testing strategies should aim to predict the life and identify the performance variability of the cells within a pack without having to test each cell

independently. Since most packs measure voltage for every cell brick, this is in principle possible, but would require every single cell voltage to be recorded in a database—at present this is not necessarily the case.

5 Summary and outlook

Battery costs and customer confidence must be improved if we are to improve renewable energy integration by expanding grid storage and rapidly replace combustion engine vehicles with electric vehicles as part of the solution to maintain climate warming within 2°C of pre-industrial levels. To address this, there is a significant need to estimate battery health accurately, diagnose degradation, predict lifetime in different usage scenarios, and detect failures. Combining data-driven and physics-based models will outperform pure data-driven or pure physics-based approaches by allowing robust extrapolation of future behavior from available data.

Different approaches are suited to different use cases. For warranty design and certification of performance, consistent lab test data is used to inform empirical models for life prediction and may also be used for data-driven or physics-based modeling. On the other hand, in most real-world cases where data are less regular, the development of new hybrid models that combine the strengths of physics-based and data-driven methods is essential. Such models are claimed to require less training data than purely data-driven methods, have been shown to extrapolate more accurately [112], and are more flexible and generalizable than purely physics-based methods. Finally, ensembles of different models could also be considered [130], since they can outperform the best individual models.

Fleet data can also be used by companies to adjust their models developed with lab data. Developing and training predictive models for battery lifetime requires large amounts of data covering the whole range of operating conditions. Due to time and equipment constraints and the wide variety of operating conditions, lab data cannot completely cover the required range with realistic (non-accelerated) operating conditions. We believe that fleet data from devices in use augments lab data, leading to use-case-specific battery lifetime prediction models and improved designs.

Additionally, deeper characterization can be performed on cells that have reached the end of their life in order to determine their true SOH and cycle life, and hence build up a database for supervised learning. Individual cells that exhibit particularly interesting behavior can be selected for tear-down experiments to reveal how these cells aged and hence verify predictions from physics-based models. This can then be used to improve the models for cells that are still in operation. Building and maintaining large databases of field data is a challenging undertaking, but is a worthwhile endeavor to prepare for a time when algorithms to

analyze it will be available and will require training data. For example, a field data collection effort could begin with collecting data from taxis and buses, or vehicles used by company employees. Even data obtained during charging could be extremely useful to update lab-based models. This is already being done in some cities, as noted in Section 3. Open sharing of such data and analysis software between researchers would greatly accelerate efforts to improve algorithms for lifetime modeling. However collection and curation of this type of data is challenging due to trade secrets and industry competition concerns. A large cross-institutional research project involving national laboratories, universities, and others could facilitate widespread data collection and appropriate anonymization. Companies may also be willing to release small subsets of data openly, for example through competitions to assist in recruitment and marketing.

To enable these exciting opportunities, further research in several areas is required, as well as enhanced collaboration between disciplines [131]. We highlight the following items as particularly necessary:

- Development of consistent, large scale open-source databases for battery performance and lifetime data, to train and validate algorithms for lifetime prediction. Ideally, these should encompass lab data and field data and share common standards for data and metadata [103].
- Fast and accurate physics-based models, with a particular focus on understanding what are the key degradation sub-models that are important, and how to estimate and track key parameters. Traditionally, such complex models have required very high levels of expertise to implement, but this is becoming more accessible through open-source modeling packages such as PyBaMM [94].
- Scalable algorithms for combining data-driven methods with physics-based models. Some promising pathways for this are the field inversion and machine learning paradigm [105], neural differential equations [111], Gaussian process state space models [132], and universal differential equations [110].
- Methods to diagnose whether a cell has aged beyond the knee point, without requiring historical data, to assess whether the cell can be used in second-life applications.
- Quantification of value (e.g. additional revenue) resulting from improved state of health estimation and life prediction, for example for second life.
- Solutions for data privacy issues—Anonymization and intelligent aggregation of individual user data without loss of insights for diagnostics and prognostics.

There are many areas of battery research where modeling and data-driven techniques can add value, including materials discovery, battery design, and fast charging algorithms [133]. Achieving breakthroughs

in battery lifetime prediction in real applications will require new experimental approaches for lab tests that massively reduce the time required to understand degradation processes in batteries. It will also require new methods to map from accelerated aging tests to real world conditions—informed by a large amount of field data together with scaleable modeling approaches that also account for uncertainty in sensors and models. There is a significant opportunity through this to reduce costs by reducing pack sizing and stretching the usage envelope, in order to accelerate the roll-out of batteries, underpinning the rapid transition to clean energy and transport.

Acknowledgements

We are grateful for helpful conversations with Dr. Manfred Baldauf (Siemens), Mr. Igor Neiva Camargo (Continental Automotive), and Dr. Markus Schweizer-Berberich (Vitesco Technologies GmbH) about the topics in this paper. D.H. and A.A. acknowledge funding from the Faraday Institution (EP/S003053/1, Grant number FIRG003), and Shell Foundation (agreement 22077). A.S., V.S., and P.M. acknowledge support from the University of Michigan Battery Lab. This work was partially supported by the National Science Foundation under Grant No. 1762247.

Declaration of Interests

D.A.H. is co-founder of Brill Power Ltd., and is a technical advisor at Habitat Energy Ltd. All other authors declare no competing interests.

Author contributions

V.S. Conceptualization, Writing - Original Draft, Writing - Review and Editing. P.M. and A.A. Visualization, Writing - Review and Editing. S.L., Y.T.Y., F.S., M.U.K., J.W.L., J.B.S. Writing - Review and Editing. A.G.S. and D.A.H. Conceptualization, Writing - Original Draft, Writing - Review and Editing, Supervision, Project administration, Funding acquisition.

References

- [1] N. A. Samad, Y. Kim, and J. B. Siegel, “On power denials and lost energy opportunities in downsizing battery packs in hybrid electric vehicles,” *Journal of Energy Storage*, vol. 16, pp. 187–196, 2018.

- [2] National Academies of Sciences Engineering and Medicine, *Assessment of Technologies for Improving Light-Duty Vehicle Fuel Economy 2025-2035*. Washington, DC: The National Academies Press, 2021.
- [3] A. Dave, J. Mitchell, K. Kandasamy, S. Burke, B. Paria, B. Poczos, J. Whitacre, and V. Viswanathan, “Autonomous discovery of battery electrolytes with robotic experimentation and machine-learning,” *Cell Reports Physical Science*, vol. 1, no. 12, p. 100264, 2020.
- [4] J. M. Reniers, G. Mulder, and D. A. Howey, “Unlocking extra value from grid batteries using advanced models,” *Journal of Power Sources*, vol. 487, no. December 2020, p. 229355, 2021.
- [5] G. Harper, R. Sommerville, E. Kendrick, L. Driscoll, P. Slater, R. Stolkin, A. Walton, P. Christensen, O. Heidrich, S. Lambert, A. Abbott, K. Ryder, L. Gaines, and P. Anderson, “Recycling lithium-ion batteries from electric vehicles,” *Nature*, vol. 575, no. November, pp. 75–86, 2019.
- [6] G. L. Plett, *Battery Management Systems, Volume II: Equivalent-Circuit Methods*. Artech House, 2016.
- [7] J. Schmalstieg, S. Käbitz, M. Ecker, and D. U. Sauer, “A holistic aging model for Li(NiMnCo)O₂ based 18650 lithium-ion batteries,” *Journal of Power Sources*, vol. 257, pp. 325–334, 2014.
- [8] J. M. Reniers, G. Mulder, and D. A. Howey, “Review and Performance Comparison of Mechanical-Chemical Degradation Models for Lithium-Ion Batteries,” *Journal of The Electrochemical Society*, vol. 166, no. 14, pp. A3189–A3200, 2019.
- [9] X. Hu, L. Xu, X. Lin, and M. Pecht, “Battery Lifetime Prognostics,” *Joule*, vol. 4, no. 2, pp. 310–346, 2020.
- [10] R. Xiong, L. Li, and J. Tian, “Towards a smarter battery management system: A critical review on battery state of health monitoring methods,” *Journal of Power Sources*, vol. 405, pp. 18–29, 11 2018.
- [11] M. Bercibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van den Bossche, “Critical review of state of health estimation methods of Li-ion batteries for real applications,” *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 572–587, 4 2016.
- [12] A. Farmann, W. Waag, A. Marongiu, and D. U. Sauer, “Critical review of on-board capacity estimation techniques for lithium-ion batteries in electric and hybrid electric vehicles,” *Journal of Power Sources*, vol. 281, pp. 114–130, 5 2015.

- [13] S. M. Rezvanizani, Z. Liu, Y. Chen, and J. Lee, "Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility," *Journal of Power Sources*, vol. 256, pp. 110–124, 6 2014.
- [14] X. Han, L. Lu, Y. Zheng, X. Feng, Z. Li, J. Li, and M. Ouyang, "A review on the key issues of the lithium ion battery degradation among the whole life cycle," *eTransportation*, vol. 1, p. 100005, 2019.
- [15] M. Dubarry and G. Baure, "Perspective on commercial Li-ion battery testing, best practices for simple and effective protocols," *Electronics (Switzerland)*, vol. 9, no. 1, 2020.
- [16] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster, "Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review," *Renewable and Sustainable Energy Reviews*, vol. 113, p. 109254, 10 2019.
- [17] M. Aykol, C. Balaji Gopal, A. Anapolsky, P. Herring, B. van Vlijmen, M. Berliner, M. Bazant, R. Braatz, W. Chueh, and B. Storey, "Perspective—Combining Physics and Machine Learning to Predict Battery Lifetime," *Journal of The Electrochemical Society*, vol. in press, 2021.
- [18] B. Wu, W. D. Widanage, S. Yang, and X. Liu, "Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems," *Energy and AI*, vol. 1, p. 100016, 2020.
- [19] D. P. Finegan, J. Zhu, X. Feng, M. Keyser, M. Ulmefors, W. Li, M. Z. Bazant, and S. J. Cooper, "The Application of Data-Driven Methods and Physics-Based Learning for Improving Battery Safety," *Joule*, no. 2020, pp. 1–14, 2020.
- [20] S. Lee, P. Mohtat, J. B. Siegel, A. G. Stefanopoulou, J.-w. Lee, and T.-k. Lee, "Estimation Error Bound of Battery Electrode Parameters With Limited Data Window," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3376–3386, 2020.
- [21] I. Bloom, B. W. Cole, J. J. Sohn, S. A. Jones, E. G. Polzin, V. S. Battaglia, G. L. Henriksen, C. Motloch, R. Richardson, T. Unkelhaeuser, D. Ingersoll, and H. L. Case, "An accelerated calendar and cycle life study of Li-ion cells," *Journal of Power Sources*, vol. 101, no. 2, pp. 238–247, 2001.
- [22] J. Wang, P. Liu, J. Hicks-Garner, E. Sherman, S. Soukiazian, M. Verbrugge, H. Tataria, J. Musser, and P. Finamore, "Cycle-life model for graphite-LiFePO₄ cells," *Journal of Power Sources*, vol. 196, no. 8, pp. 3942–3948, 4 2011.

- [23] M. Schimpe, M. E. von Kuepach, M. Naumann, H. C. Hesse, K. Smith, and A. Jossen, "Comprehensive Modeling of Temperature-Dependent Degradation Mechanisms in Lithium Iron Phosphate Batteries," *Journal of The Electrochemical Society*, vol. 165, no. 2, pp. A181–A193, 2018.
- [24] M. De Gennaro, E. Paffumi, G. Martini, A. Giallonardo, S. Pedroso, and A. Loisel-Lapointe, "A case study to predict the capacity fade of the battery of electrified vehicles in real-world use conditions," *Case Studies on Transport Policy*, vol. 8, no. 2, pp. 517–534, 2020.
- [25] M. Ecker, J. B. Gerschler, J. Vogel, S. Käbitz, F. Hust, P. Dechent, and D. U. Sauer, "Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data," *Journal of Power Sources*, vol. 215, pp. 248–257, 2012.
- [26] H. Chaoui and C. C. Ibe-Ekeocha, "State of Charge and State of Health Estimation for Lithium Batteries Using Recurrent Neural Networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 8773–8783, 2017.
- [27] H. Chaoui, C. C. Ibe-Ekeocha, and H. Gualous, "Aging prediction and state of charge estimation of a LiFePO₄ battery using input time-delayed neural networks," *Electric Power Systems Research*, vol. 146, pp. 189–197, 2017.
- [28] Y. Ding, C. Lu, and J. Ma, "Li-ion battery health estimation based on multi-layer characteristic fusion and deep learning," *2017 IEEE Vehicle Power and Propulsion Conference, VPPC 2017 - Proceedings*, vol. 2018-Janua, pp. 1–5, 2018.
- [29] J. Wu, Y. Wang, X. Zhang, and Z. Chen, "A novel state of health estimation method of Li-ion battery using group method of data handling," *Journal of Power Sources*, vol. 327, pp. 457–464, 2016.
- [30] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, and L. Zhang, "Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach," *IEEE Access*, vol. 6, pp. 50 587–50 598, 2018.
- [31] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *Journal of Power Sources*, vol. 384, no. September 2017, pp. 387–395, 2018.
- [32] V. Klass, M. Behm, and G. Lindbergh, "A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation," *Journal of Power Sources*, vol. 270, pp. 262–272, 2014.

- [33] Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming, and A. A. Lee, "Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning," *Nature Communications*, vol. 11, no. 1, pp. 6–11, 2020.
- [34] X. Tang, K. Liu, X. Wang, F. Gao, J. MacRo, and W. D. Widanage, "Model Migration Neural Network for Predicting Battery Aging Trajectories," *IEEE Transactions on Transportation Electrification*, vol. 6, no. 2, pp. 363–374, 2020.
- [35] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *Journal of Power Sources*, vol. 357, pp. 209–219, 2017.
- [36] J. Wei, G. Dong, and Z. Chen, "Remaining Useful Life Prediction and State of Health Diagnosis for Lithium-Ion Batteries Using Particle Filter and Support Vector Regression," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5634–5643, 2018.
- [37] C. Hu, G. Jain, C. Schmidt, C. Strief, and M. Sullivan, "Online estimation of lithium-ion battery capacity using sparse Bayesian learning," *Journal of Power Sources*, vol. 289, pp. 105–113, 2015.
- [38] Z. Wang, S. Zeng, J. Guo, and T. Qin, "State of health estimation of lithium-ion batteries based on the constant voltage charging curve," *Energy*, vol. 167, pp. 661–669, 2019.
- [39] A. Eddahech, O. Briat, and J. M. Vinassa, "Determination of lithium-ion battery state-of-health based on constant-voltage charge phase," *Journal of Power Sources*, vol. 258, pp. 218–227, 2014.
- [40] Z. Wang, J. Ma, and L. Zhang, "State-of-Health Estimation for Lithium-Ion Batteries Based on the Multi-Island Genetic Algorithm and the Gaussian Process Regression," *IEEE Access*, vol. 5, pp. 21 286–21 295, 2017.
- [41] N. A. Samad, Y. Kim, J. B. Siegel, and A. G. Stefanopoulou, "Battery Capacity Fading Estimation Using a Force-Based Incremental Capacity Analysis," *Journal of The Electrochemical Society*, vol. 163, no. 8, pp. A1584–A1594, 2016.
- [42] N. Yang, T. Liu, and D. Zhong, "Online Estimation of State-Of-Health for Lithium Ion Batteries Based on Charge Curves," *11th International Conference on Reliability, Maintainability and Safety (ICRMS), IEEE*, pp. 1–8, 2016.
- [43] H.-T. Lin, T.-J. Liang, and S.-M. Chen, "Estimation of Battery State of Health Using Probabilistic Neural Network," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 2, pp. 679–685, 2013.

- [44] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. A. Howey, "Gaussian Process Regression for In Situ Capacity Estimation of Lithium-Ion Batteries," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 127–138, 1 2019.
- [45] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Battery health prediction under generalized conditions using a Gaussian process transition model," *Journal of Energy Storage*, vol. 23, no. March, pp. 320–328, 2019.
- [46] J. Yang, B. Xia, W. Huang, Y. Fu, and C. Mi, "Online state-of-health estimation for lithium-ion batteries using constant-voltage charging current analysis," *Applied Energy*, vol. 212, no. August 2017, pp. 1589–1600, 2018.
- [47] B. Chinomona, C. Chung, L.-K. Chang, W.-C. Su, and M.-C. Tsai, "Long Short-Term Memory Approach to Estimate Battery Remaining Useful Life Using Partial Data," *IEEE Access*, vol. 8, pp. 165 419–165 431, 2020.
- [48] Z. Yun and W. Qin, "Remaining useful life estimation of lithium-ion batteries based on optimal time series health indicator," *IEEE Access*, vol. 8, pp. 55 447–55 461, 2020.
- [49] S. Greenbank and D. A. Howey, "Automated feature selection for data-driven models of rapid battery capacity fade and end of life," *arXiv*, 2021.
- [50] M. Lucu, E. Martinez-Laserna, I. Gandiaga, K. Liu, H. Camblong, W. D. Widanage, and J. Marco, "Data-driven nonparametric Li-ion battery ageing model aiming at learning from real operation data - Part B: Cycling operation," *Journal of Energy Storage*, vol. 30, no. May, p. 101410, 2020.
- [51] ———, "Data-driven nonparametric Li-ion battery ageing model aiming at learning from real operation data – Part A: Storage operation," *Journal of Energy Storage*, vol. 30, no. April, p. 101409, 2020.
- [52] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fragedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, and R. D. Braatz, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.
- [53] P. Fermín-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M. F. Anjos, J. Sylvester, and G. dos Reis, "Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells," *Energy and AI*, vol. 1, p. 100006, 2020.

- [54] V. Sulzer, P. Mohtat, S. Lee, J. B. Siegel, and A. G. Stefanopoulou, “Promise and Challenges of a Data-Driven Approach for Battery Lifetime Prognostics,” *arXiv preprint arXiv:2010.07460*, 2020.
- [55] I. Bloom, A. N. Jansen, D. P. Abraham, J. Knuth, S. A. Jones, V. S. Battaglia, and G. L. Henriksen, “Differential voltage analyses of high-power, lithium-ion cells 1. Technique and application,” *Journal of Power Sources*, vol. 139, no. 1-2, pp. 295–303, 2005.
- [56] H. M. Dahn, A. J. Smith, J. C. Burns, D. A. Stevens, and J. R. Dahn, “User-Friendly Differential Voltage Analysis Freeware for the Analysis of Degradation Mechanisms in Li-Ion Batteries,” *Journal of The Electrochemical Society*, vol. 159, no. 9, pp. A1405–A1409, 2012.
- [57] J. Wang, J. Purewal, P. Liu, J. Hicks-Garner, S. Soukazian, E. Sherman, A. Sorenson, L. Vu, H. Tataria, and M. W. Verbrugge, “Degradation of lithium ion batteries employing graphite negatives and nickel-cobalt-manganese oxide + spinel manganese oxide positives: Part 1, aging mechanisms and life estimation,” *Journal of Power Sources*, vol. 269, pp. 937–948, 2014.
- [58] S. Lee, J. B. Siegel, A. G. Stefanopoulou, J.-W. Lee, and T.-K. Lee, “Electrode State of Health Estimation for Lithium Ion Batteries Considering Half-cell Potential Change Due to Aging,” *Journal of The Electrochemical Society*, vol. 167, no. 9, p. 090531, 2020.
- [59] M. Dubarry, G. Baure, and A. Devie, “Durability and Reliability of EV Batteries under Electric Utility Grid Operations: Path Dependence of Battery Degradation,” *Journal of The Electrochemical Society*, vol. 165, no. 5, pp. A773–A783, 2018.
- [60] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, “Degradation diagnostics for lithium ion cells,” *Journal of Power Sources*, vol. 341, pp. 373–386, 2017.
- [61] D. Zhang, S. Dey, H. E. Perez, and S. J. Moura, “Remaining Useful Life Estimation of Lithium-Ion Batteries based on Thermal Dynamics,” *American Control Conference*, 2017.
- [62] A. Chu, A. Allam, A. Cordoba Arenas, G. Rizzoni, and S. Onori, “Stochastic capacity loss and remaining useful life models for lithium-ion batteries in plug-in hybrid electric vehicles,” *Journal of Power Sources*, vol. 478, no. September, p. 228991, 2020.
- [63] Y. Li, M. Vilathgamuwa, T. Farrell, S. S. Choi, N. T. Tran, and J. Teague, “A physics-based distributed-parameter equivalent circuit model for lithium-ion batteries,” *Electrochimica Acta*, vol. 299, pp. 451–469, 2019.

- [64] M. Ouyang, X. Feng, X. Han, L. Lu, Z. Li, and X. He, "A dynamic capacity degradation model and its applications considering varying load for a large format Li-ion battery," *Applied Energy*, vol. 165, pp. 48–59, 2016.
- [65] S. Atalay, M. Sheikh, A. Mariani, Y. Merla, E. Bower, and W. D. Widanage, "Theory of battery ageing in a lithium-ion battery: Capacity fade, nonlinear ageing and lifetime prediction," *Journal of Power Sources*, vol. 478, p. 229026, 2020.
- [66] X. G. Yang, Y. Leng, G. Zhang, S. Ge, and C. Y. Wang, "Modeling of lithium plating induced aging of lithium-ion batteries: Transition from linear to nonlinear aging," *Journal of Power Sources*, vol. 360, pp. 28–40, 2017.
- [67] J. Keil and A. Jossen, "Electrochemical Modeling of Linear and Nonlinear Aging of Lithium-Ion Cells," *Journal of The Electrochemical Society*, vol. 167, no. 11, p. 110535, 2020.
- [68] G. Ning, R. E. White, and B. N. Popov, "A generalized cycle life model of rechargeable Li-ion batteries," *Electrochimica Acta*, vol. 51, no. 10, pp. 2012–2022, 2006.
- [69] H. Ekström and G. Lindbergh, "A Model for Predicting Capacity Fade due to SEI Formation in a Commercial Graphite/LiFePO₄ Cell," *Journal of The Electrochemical Society*, vol. 162, no. 6, pp. A1003–A1007, 2015.
- [70] M. Safari and C. Delacourt, "Simulation-Based Analysis of Aging Phenomena in a Commercial Graphite/LiFePO₄ Cell," *Journal of The Electrochemical Society*, vol. 158, no. 12, p. A1436, 2011.
- [71] P. Mohtat, S. Lee, J. B. Siegel, and A. G. Stefanopoulou, "Towards better estimability of electrode-specific state of health: Decoding the cell expansion," *Journal of Power Sources*, vol. 427, no. January, pp. 101–111, 2019.
- [72] A. G. Hsieh, S. Bhadra, B. J. Hertzberg, P. J. Gjeltema, A. Goy, J. W. Fleischer, and D. A. Steingart, "Electrochemical-acoustic time of flight: In operando correlation of physical dynamics with battery charge and health," *Energy and Environmental Science*, vol. 8, no. 5, pp. 1569–1577, 2015.
- [73] T. Raj, A. A. Wang, C. W. Monroe, and D. A. Howey, "Investigation of Path-Dependent Degradation in Lithium-Ion Batteries," *Batteries & Supercaps*, pp. 1–10, 9 2020.
- [74] L. Su, J. Zhang, J. Huang, H. Ge, Z. Li, F. Xie, and B. Y. Liaw, "Path dependence of lithium ion cells aging under storage conditions," *Journal of Power Sources*, vol. 315, pp. 35–46, 5 2016.

- [75] K. L. Gering, S. V. Sazhin, D. K. Jamison, C. J. Michelbacher, B. Y. Liaw, M. Dubarry, and M. Cugnet, “Investigation of path dependence in commercial lithium-ion cells chosen for plug-in hybrid vehicle duty cycle protocols,” *Journal of Power Sources*, vol. 196, no. 7, pp. 3395–3403, 4 2011.
- [76] Z. Ma, J. Jiang, W. Shi, W. Zhang, and C. C. Mi, “Investigation of path dependence in commercial lithium-ion cells for pure electric bus applications: Aging mechanism identification,” *Journal of Power Sources*, vol. 274, pp. 29–40, 1 2015.
- [77] Anonymous, “Research and Development of High-Power and High-Energy Electrochemical Storage Devices,” United States Advanced Battery Consortium, Tech. Rep., 2014.
- [78] K. Smith, A. Saxon, M. Keyser, B. Lundstrom, Z. Cao, and A. Roc, “Life prediction model for grid-connected Li-ion battery energy storage system,” *Proceedings of the American Control Conference*, pp. 4062–4068, 2017.
- [79] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, and D. U. Sauer, “Online capacity estimation of lithium-ion batteries with deep long short-term memory networks,” *Journal of Power Sources*, vol. 482, p. 228863, 2021.
- [80] W. Diao, S. Saxena, B. Han, and M. Pecht, “Algorithm to determine the knee point on capacity fade curves of lithium-ion cells,” *Energies*, vol. 12, no. 15, 2019.
- [81] R. Drummond, S. Zhao, D. A. Howey, and S. R. Duncan, “Circuit synthesis of electrochemical supercapacitor models,” *Journal of Energy Storage*, vol. 10, pp. 48–55, 2017.
- [82] A. Aitio and D. Howey, “Combining non-parametric and parametric models for stable and computationally efficient battery health estimation,” *Proceedings of the ASME 2020 Dynamic Systems and Control Conference*, 2020.
- [83] S. G. Marquis, V. Sulzer, R. Timms, C. P. Please, and S. Jon Chapman, “An asymptotic derivation of a single particle model with electrolyte,” *Journal of the Electrochemical Society*, vol. 166, no. 15, pp. A3693–A3706, 2019.
- [84] V. Sulzer, S. J. Chapman, C. P. Please, D. A. Howey, and C. W. Monroe, “Faster Lead-Acid Battery Simulations from Porous-Electrode Theory: Part II. Asymptotic Analysis,” *Journal of The Electrochemical Society*, vol. 166, no. 12, pp. A2372–A2382, 2019.

- [85] D. Di Domenico, A. G. Stefanopoulou, and G. Fiengo, “Lithium-Ion Battery State of Charge and Critical Surface Charge Estimation Using an Electrochemical Model-Based Extended Kalman Filter,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 132, no. 6, p. 061302, 2010.
- [86] F. Single, A. Latz, and B. Horstmann, “Identifying the Mechanism of Continued Growth of the Solid–Electrolyte Interphase,” *ChemSusChem*, vol. 11, no. 12, pp. 1950–1955, 2018.
- [87] M. Tang, S. Lu, and J. Newman, “Experimental and Theoretical Investigation of Solid-Electrolyte-Interphase Formation Mechanisms on Glassy Carbon,” *Journal of The Electrochemical Society*, vol. 159, no. 11, pp. A1775–A1785, 2012.
- [88] P. Ramadass, B. Haran, P. M. Gomadam, R. White, and B. N. Popov, “Development of First Principles Capacity Fade Model for Li-Ion Cells,” *Journal of The Electrochemical Society*, vol. 151, no. 2, p. A196, 1 2004.
- [89] S. E. J. O’Kane, I. D. Campbell, M. W. J. Marzook, G. J. Offer, and M. Marinescu, “Physical Origin of the Differential Voltage Minimum Associated with Lithium Plating in Li-Ion Batteries,” *Journal of The Electrochemical Society*, vol. 167, no. 9, p. 090540, 5 2020.
- [90] P. Arora, M. Doyle, and R. E. White, “Mathematical Modeling of the Lithium Deposition Overcharge Reaction in Lithium-Ion Batteries Using Carbon-Based Negative Electrodes,” *Journal of The Electrochemical Society*, vol. 146, no. 10, pp. 3543–3553, 1999.
- [91] W. Ai, L. Kraft, J. Sturm, A. Jossen, and B. Wu, “Electrochemical Thermal-Mechanical Modelling of Stress Inhomogeneity in Lithium-Ion Pouch Cells,” *Journal of The Electrochemical Society*, vol. 167, no. 1, p. 013512, 2020.
- [92] C.-H. Chen, F. Brosa Planella, K. O’Regan, D. Gastol, W. D. Widanage, and E. Kendrick, “Development of Experimental Techniques for Parameterization of Multi-scale Lithium-ion Battery Models,” *Journal of The Electrochemical Society*, vol. 167, no. 8, p. 080534, 5 2020.
- [93] M. Ecker, S. Käbitz, I. Laresgoiti, and D. U. Sauer, “Parameterization of a Physico-Chemical Model of a Lithium-Ion Battery,” *Journal of The Electrochemical Society*, vol. 162, no. 9, pp. A1849–A1857, 2015.
- [94] V. Sulzer, S. G. Marquis, R. Timms, M. Robinson, and S. J. Chapman, “Python Battery Mathematical Modelling (PyBaMM),” *ECSarXiv. February*, vol. 7, pp. 1–9, 2020.

- [95] Y. Hui, M. Li, A. Downey, S. Shen, V. Pavan, H. Ye, C. Vanelzen, G. Jain, S. Hu, S. Laflamme, and C. Hu, “Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction,” *Journal of Power Sources*, vol. 485, no. December 2020, p. 229327, 2021.
- [96] F. An and M. Ross, “A model of fuel economy and driving patterns. sae technical paper 930328,” *Society of Automotive Engineers*, 1993.
- [97] B. Y. Liaw and M. Dubarry, “From driving cycle analysis to understanding battery performance in real-life electric hybrid vehicle operation,” *Journal of power sources*, vol. 174, no. 1, pp. 76–88, 2007.
- [98] A. Saxena, J. R. Celaya, I. Roychoudhury, B. Saha, S. Saha, and K. Goebel, “Designing Data-Driven Battery Prognostic Approaches for Variable Loading Profiles: Some Lessons Learned,” in *First European Conference of the Prognostics and Health Management Society 2012*, 2012, p. 10.
- [99] L. Song, K. Zhang, T. Liang, X. Han, and Y. Zhang, “Intelligent state of health estimation for lithium-ion battery pack based on big data analysis,” *Journal of Energy Storage*, vol. 32, no. July, p. 101836, 2020.
- [100] Q. Wang, Z. Wang, L. Zhang, P. Liu, and Z. Zhang, “A novel consistency evaluation method for series-connected battery systems based on real-world operation data,” *IEEE Transactions on Transportation Electrification*, vol. 7, no. 2, pp. 1–1, 2020.
- [101] C. She, Z. Wang, F. Sun, P. Liu, and L. Zhang, “Battery Aging Assessment for Real-World Electric Buses Based on Incremental Capacity Analysis and Radial Basis Function Neural Network,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3345–3354, 2020.
- [102] Q. Huo, Z. Ma, X. Zhao, T. Zhang, and Y. Zhang, “Bayesian Network Based State-of-Health Estimation for Battery on Electric Vehicle Application and its Validation through Real-World Data,” *IEEE Access*, vol. 9, pp. 11 328–11 341, 2021.
- [103] V. De Angelis, Y. Preger, and B. R. Chalama, “Battery Lifecycle Framework: A Flexible Repository and Visualization Tool for Battery Data from Materials Development to Field Implementation,” *ECSarXiv*, 2021.
- [104] M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, “Fusing physics-based and deep learning models for prognostics,” *arXiv preprint arXiv:2003.00732*, 2020.

- [105] E. J. Parish and K. Duraisamy, “A paradigm for data-driven predictive modeling using field inversion and machine learning,” *Journal of Computational Physics*, vol. 305, pp. 758–774, 2016.
- [106] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics informed deep learning (Part II): Data-driven discovery of nonlinear partial differential equations,” *arXiv*, no. Part II, pp. 1–19, 2017.
- [107] ———, “Physics informed deep learning (Part II): Data-driven discovery of nonlinear partial differential equations,” *arXiv*, no. Part I, pp. 1–22, 2017.
- [108] ———, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [109] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, “Neural Ordinary Differential Equations,” *32nd Conference on Neural Information Processing Systems*, 2018.
- [110] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, and A. Ramadhan, “Universal differential equations for scientific machine learning,” *arXiv*, 2020.
- [111] C. Rackauckas, M. Innes, Y. Ma, J. Bettencourt, L. White, and V. Dixit, “DiffEqFlux.jl — A julia library for neural differential equations,” *arXiv*, pp. 1–17, 2019.
- [112] R. Dandekar, C. Rackauckas, and G. Barbastathis, “A Machine Learning-Aided Global Diagnostic and Comparative Tool to Assess Effect of Quarantine Control in COVID-19 Spread,” *Patterns*, vol. 1, no. 9, p. 100145, 2020.
- [113] A. Bills, S. Sripad, W. L. Fredericks, M. Guttenberg, D. Charles, E. Frank, and V. Viswanathan, “Universal Battery Performance and Degradation Model for Electric Aircraft,” *arXiv preprint arXiv:2008.01527*, 2020.
- [114] D. J. Tait, F. Brosa Planella, T. Damoulas, and W. Dhammika Widanage, “Scalable Multitask Latent Force Models with Applications to Predicting Lithium-ion Concentration,” *NeurIPS*, 2020.
- [115] Northvolt, “Setting a new standard in digitalization of battery assets,” 2019.
- [116] C. Speltino, A. Stefanopoulou, and G. Fiengo, “Cell equalization in battery stacks through State Of Charge estimation polling,” *Proceedings of the 2010 American Control Conference, ACC 2010*, pp. 5050–5055, 2010.

- [117] Z. Song, J. Hou, X. Li, X. Wu, X. Hu, H. Hofmann, and J. Sun, “The sequential algorithm for combined state of charge and state of health estimation of lithium-ion battery based on active current injection,” *Energy*, vol. 193, p. 116732, 2020.
- [118] R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, “Critical Review on the Battery State of Charge Estimation Methods for Electric Vehicles,” *IEEE Access*, vol. 6, pp. 1832–1843, 2017.
- [119] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, “Variational Inference: A Review for Statisticians,” *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.
- [120] R. V. Craiu and J. S. Rosenthal, “Bayesian computation via markov chain monte carlo,” *Annual Review of Statistics and Its Application*, vol. 1, pp. 179–201, 2014.
- [121] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, “Probabilistic programming in Python using PyMC3,” *PeerJ Computer Science*, vol. 2, p. e55, 2016.
- [122] E. Martinez-Laserna, I. Gandiaga, E. Sarasketa-Zabala, J. Badedo, D. I. Stroe, M. Swierczynski, and A. Goikoetxea, “Battery second life: Hype, hope or reality? A critical review of the state of the art,” *Renewable and Sustainable Energy Reviews*, vol. 93, no. April, pp. 701–718, 2018.
- [123] European Commission, “Green Deal: Sustainable batteries for a circular and climate neutral economy,” 2020.
- [124] E. Martinez-Laserna, E. Sarasketa-Zabala, I. Villarreal Sarria, D. I. Stroe, M. Swierczynski, A. Warnecke, J. M. Timmermans, S. Goutam, N. Omar, and P. Rodriguez, “Technical Viability of Battery Second Life: A Study from the Ageing Perspective,” *IEEE Transactions on Industry Applications*, vol. 54, no. 3, pp. 2703–2713, 2018.
- [125] S. Park, D. Kato, Z. Gima, R. Klein, and S. Moura, “Optimal experimental design for parameterization of an electrochemical lithium-ion battery model,” *Journal of the Electrochemical Society*, vol. 165, no. 7, pp. A1309–A1323, 2018.
- [126] A. M. Bizeray, J. H. Kim, S. R. Duncan, and D. A. Howey, “Identifiability and parameter estimation of the single particle lithium-ion battery model,” *IEEE Transactions on Control Systems Technology*, vol. 27, no. 5, pp. 1862–1877, 2019.
- [127] C. Zhang, Y. Wang, Y. Gao, F. Wang, B. Mu, and W. Zhang, “Accelerated fading recognition for

- lithium-ion batteries with Nickel-Cobalt-Manganese cathode using quantile regression method,” *Applied Energy*, vol. 256, no. March, p. 113841, 2019.
- [128] X. Cui, J. Siegel, A. Stefanopoulou, and A.-T. Avestruz, “Grid Interfaces to Electric Vehicle Chargers Using Statistically-Structured Power Conversion Architectures for Second-Use Batteries as Energy Buffering,” *Submitted to: The IEEE Transactions on Smart Grid*, 2021.
- [129] J. Neubauer, K. Smith, E. Wood, and A. Pesaran, “Identifying and Overcoming Critical Barriers to Widespread Second Use of PEV Batteries,” *National Renewable Energy Laboratory (NREL)*, no. February, pp. 23–62, 2 2015.
- [130] D. Opitz and R. Maclin, “Popular Ensemble Methods: An Empirical Study,” *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 1999.
- [131] J. A. Baker, M. Beuse, S. C. Decaluwe, L. W. Jing, E. Khoo, S. Sripad, U. Ulissi, A. Verma, A. A. Wang, Y. T. Yeh, N. Yiu, D. A. Howey, and V. Viswanathan, “Fostering a Sustainable Community in Batteries,” *ACS Energy Letters*, vol. 5, no. 7, pp. 2361–2366, 2020.
- [132] R. Frigola, Y. Chen, and C. E. Rasmussen, “Variational Gaussian process state-space models,” *Advances in Neural Information Processing Systems*, vol. 4, no. January, pp. 3680–3688, 2014.
- [133] D. A. Howey, S. A. Roberts, V. Viswanathan, A. Mistry, M. Beuse, E. Khoo, S. C. Decaluwe, and V. Sulzer, “Free Radicals : Making a Case for Battery Modeling,” *The Electrochemical Society Interface*, vol. 29, no. 30, 2020.