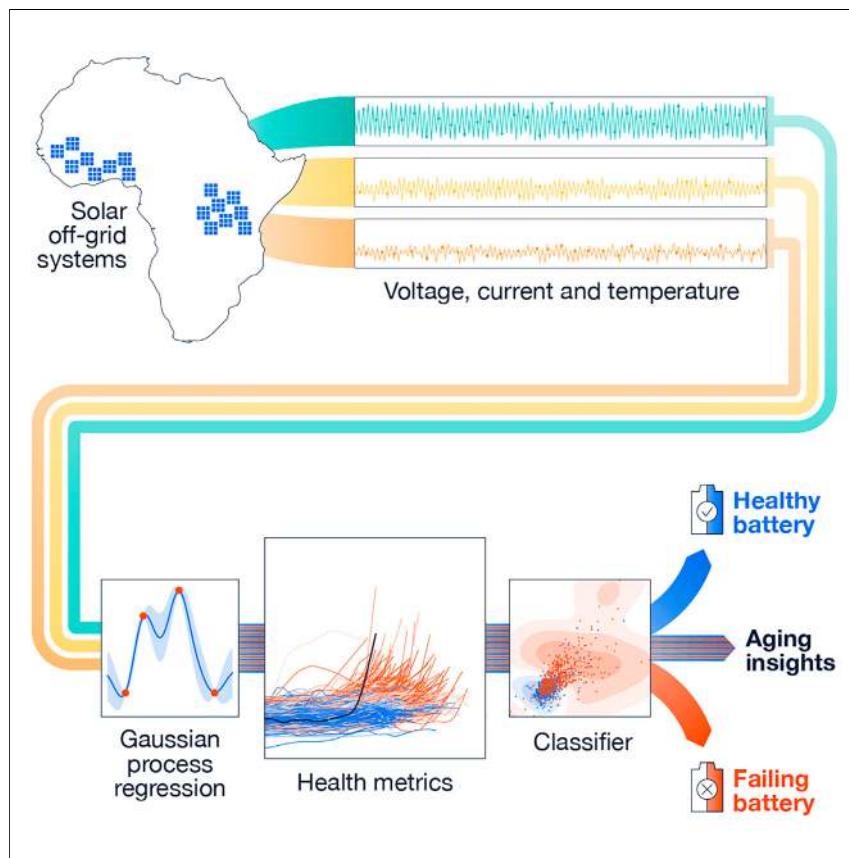


Article

Predicting battery end of life from solar off-grid system field data using machine learning



Off-grid solar-battery systems provide clean electricity, enabling education and enterprise. However, these systems are in remote areas, and it can be difficult to replace failed batteries. To improve reliability and cost-effectiveness, a non-invasive method to estimate battery health is required. We demonstrate how real-world operating data may be used to infer health and detect end of life. This work highlights the opportunity to analyze field data with machine learning to understand battery aging.

Antti Aitio, David A. Howey

david.howey@eng.ox.ac.uk

Highlights

Off-grid solar-battery systems provide clean electricity to millions of people

Battery replacement can be difficult due to remoteness

We demonstrate non-invasive estimation of battery health from field data

We estimate end-of-life probability with 82% accuracy from 1,027 batteries



Article

Predicting battery end of life from solar off-grid system field data using machine learning

Antti Aitio¹ and David A. Howey^{1,2,*}

SUMMARY

Hundreds of millions of people lack access to electricity. Decentralized solar-battery systems are key for addressing this while avoiding carbon emissions and air pollution but are hindered by relatively high costs and rural locations that inhibit timely preventive maintenance. Accurate diagnosis of battery health and prediction of end of life from operational data improves user experience and reduces costs. However, lack of controlled validation tests and variable data quality mean existing lab-based techniques fail to work. We apply a scalable probabilistic machine learning approach to diagnose health in 1,027 solar-connected lead-acid batteries, each running for 400–760 days, totaling 620 million data rows. We demonstrate 73% accurate prediction of end of life, 8 weeks in advance, rising to 82% at the point of failure. This work highlights the opportunity to estimate health from existing measurements using “big data” techniques, without additional equipment, extending lifetime and improving performance in real-world applications.

INTRODUCTION

To achieve universal electricity access, the number of decentralized solar-battery systems and solar mini grids in areas without grid access will need to increase 10-fold,¹ but this is inhibited by the relatively high costs and uncertain lifetimes of batteries.² Diagnosis and prediction of battery state of health (SoH) in real-world operating environments is required for operational safety, warranties, and planning of maintenance, as well as for improving designs by understanding the impact of varying usage on battery life. Diagnosis in the field is challenging because direct measurement of SoH using standardized performance tests is usually not possible due to the costs of the required service interruption and testing equipment. Therefore, health diagnosis should be performed directly from monitored operational data, for example, battery terminal voltage, temperature, and current. However, this means that the controlled operating conditions that would normally ensure consistent health estimation in laboratory tests are missing. Further complication arises because the most common battery health metrics—capacity and internal resistance—are influenced by operating conditions. Finally, the current, voltage, and temperature sensors used in many real-world applications are not of the same accuracy as those in laboratories, and data recording is often incomplete, resulting in gaps in the time series.

Techniques for battery health estimation can broadly be classified into model-based and data-driven methods.^{3–6} Model-based approaches typically use an electrical equivalent circuit model combined with techniques from feedback control to track internal states, such as state of charge (SoC), and parameters, such as resistance and capacity. Gradual evolution in the parameters enables SoH estimation using Bayesian filtering^{7–9} or adaptive observers.^{10–12} The choice of battery model is a

Context & scale

Off-grid solar-battery systems provide clean electricity to those who lack access, enabling education and enterprise. However, these systems are in remote areas, and when their batteries fail, it can be difficult to replace them quickly. To improve reliability, uptime, and cost-effectiveness, a scalable and non-invasive method to diagnose and predict battery health is required. We demonstrate how real-world battery operating data may be used to infer health and detect end-of-life failure using only the measured voltage, current, and temperature data from users in the field, with no additional sensors or requirements to take systems offline for specific tests. This work highlights the opportunity to complement laboratory battery data with large field datasets analyzed through machine learning to improve performance and understanding of aging.



trade-off between parsimony/computational resources and accuracy/flexibility. Equivalent circuit models are ubiquitous and widely employed^{7,8} but may suffer from a lack of accuracy across the wide range of operating points experienced in real-world usage.¹³ Higher fidelity “physics-based” models derived from porous electrode theory are also available^{14,15} but are generally considered too complex and computationally demanding for SoH tracking despite recent improvements.^{16–20}

In contrast, data-driven methods try to estimate SoH either directly from raw current, voltage, and temperature measurements or from features such as time spent in certain operating regions. Machine learning techniques used for this purpose include Gaussian process (GP) regression,^{21,22} support vector machines,²³ and neural networks.^{24,25} Data-driven health estimation has mainly been investigated using laboratory data under controlled conditions at relatively small scales. Many publicly available experimental laboratory datasets exist,²⁶ but they are limited in size.

Supplementing laboratory aging data, field data from end users enables failure detection, gives insight into real-world performance, and improves understanding of manufacturing and usage variability.²⁷ There have been few studies of battery SoH estimation using field data, even though this presents a significant opportunity to broaden understanding of performance.^{28,29} Although existing demonstrators show promise, their “black-box” nature reduces interpretability in comparison with model-driven approaches. Additionally, if charging and discharging patterns are highly dynamic, it might not be possible to calculate consistent features over the battery lifetime. So-called “hybrid” models that combine data-driven and physics-based methods³⁰ offer promise by balancing robustness, flexibility, and transparency.

In real-world applications, constraints are imposed by changing usage conditions, less accurate sensors, lack of controlled usage, and lack of prior knowledge of model parameters. Here, we implement physics-informed probabilistic machine learning for SoH estimation that is robust to changing operating conditions and data gaps. Our approach, illustrated in [Figure 1](#), is demonstrated using measured current, voltage, and temperature data from lead-acid batteries but is applicable to any chemistry that can be represented with a low-order electrical model, including lithium-ion cells. We employ GP regression^{31–33} to construct health trajectories of 1,027 batteries connected to photovoltaic systems in sub-Saharan Africa, using internal resistance as a health metric, since it may be estimated from data more accurately than discharge capacity because the observable impact of capacity changes on measured voltage is very small. Resistance is estimated as a function of instantaneous operating conditions, such as temperature and SoC, enabling calibration of data, resulting in smooth estimates. Next, an end-of-life failure detection algorithm based on a GP classifier³¹ predicts the probability of future failure by combining the estimated health trajectories with stress factors also extracted from the raw data. To train the classifier, batteries were labeled as failed or healthy using repair data from field workshops. We demonstrate, using 5-fold stratified cross validation, that this technique gives 82% balanced accuracy of end-of-life failure prediction at the time of failure versus a 66% benchmark, and 73% accuracy 8 weeks in advance of failure versus a 50% benchmark.

RESULTS AND DISCUSSION

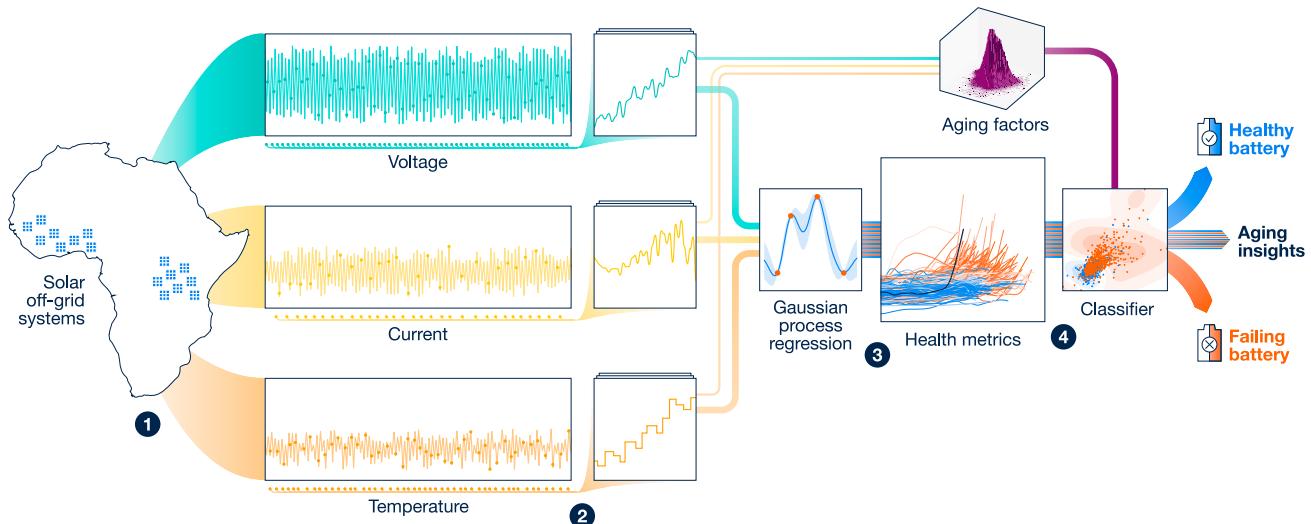
Real-world dataset

Our dataset was generated by 1,027 lead-acid batteries, each with nominal voltage 12 V (internally comprising 6 cells in series), nominal capacity 20 Ah, and attached to

¹Department of Engineering Science, Oxford University, Parks Road, OX1 3PJ Oxford, UK

²Lead contact

*Correspondence: david.howey@eng.ox.ac.uk
<https://doi.org/10.1016/j.joule.2021.11.006>

**Figure 1. Workflow**

Left to right: (1) Data are retrieved from PV-connected batteries—the raw dataset contains 620 million rows of current, voltage, and temperature measurements. (2) Data are sub-sampled for computational efficiency to yield 39 million rows from an average of 261 charging segments per battery. (3) GP regression is applied to obtain estimates for health metrics (calibrated internal resistance) over the lifetime of each battery. (4) Batteries are classified as healthy or failing up to 2 months prior to end of life using these metrics plus other aging factors calculated from the raw data; this also gives insight into the key factors driving aging.

a 50 Wp photovoltaic panel. These systems are used for lighting, phone charging, and small appliances, and are located across sub-Saharan Africa. Each battery was in use for 400–760 days, giving a total dataset size of 620 million rows (49 GB). This dataset is a small subset of the total number of systems deployed and was selected to ensure that each time series was at least 400 days in length and also that the set contained approximately the same number of failed versus healthy batteries. A full explanation of the data selection process is given in the [supplemental information](#). The distributions of terminal voltage, current (where, by convention, charging is positive), temperature, and sampling times are shown in [Figure 2](#). Sampling times are irregular with a mode of 60 s, with temperature recorded less frequently on average.

There are some particular challenges for battery health diagnosis from the measured field data in this application. First, the depth of discharge is commonly quite small, for example, the majority of usage in our dataset is between 46% and 100% SoC, making it difficult to observe changes in the discharge voltage curve caused by capacity changes as the cell ages. Second, since average currents are small (99th percentile approximately 0.1C, where C-rate is defined as the current divided by the nominal Ampere-hour capacity of the battery) and often constant, estimation of the internal resistance is numerically poorly conditioned, in other words small errors in measured voltage cause large errors in estimated resistance. These difficulties are compounded by unknown sensor accuracy.

Furthermore, the average currents, temperatures, and depths of discharge vary over time and across the population of batteries. We therefore require a methodology for health diagnosis that is robust to changes in usage and that does not require controlled diagnostic tests. In most battery systems, either capacity or internal resistance (or both) may be used as a health metric.^{34,35} In lead-acid systems, internal resistance is a clear health indicator,³⁶ although the techniques outlined here can

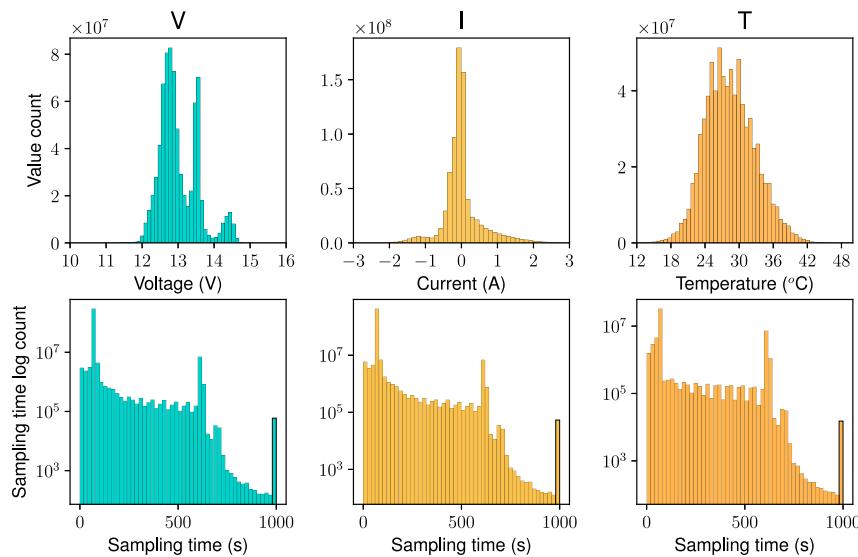


Figure 2. Solar off-grid systems experience varying states of charge, currents, and temperatures, and the systems we used had non-uniform data sampling

Shown here (top) are distributions of voltage, current, temperature, and (bottom) histograms of their sampling times. Sampling time has a mode at $t = 60$ s for each measurement (counts are plotted on a logarithmic scale).

also be used to estimate capacity if sufficiently deep discharges are available. In our dataset, charging segments typically provide a more varied and higher amplitude input signal for resistance estimation compared with discharging segments, where the current is on average small and relatively constant. We therefore chose to estimate the resistance during charging as the SoH metric.

Data-driven modeling

We now describe the modeling and inference process, outlined in [Figure 1](#), in detail. First, measured current, voltage, and temperature data with non-uniform sampling were obtained from a central database, and suitable charging segments selected (see [experimental procedures](#))—on average, 261 such segments were identified per battery. Then, within each charging segment, data were interpolated onto a 1 min uniform grid in time, and electrolyte acid concentration as a SoC metric was estimated using Coulomb counting, with initial conditions calibrated by a known open circuit voltage (OCV) versus SoC relationship (see [experimental procedures](#)). This process resulted in an average of 38,000 rows of data for each battery, reduced 16-fold down from 603,000 rows per battery.

The internal resistance of all batteries depends on age, current, temperature, and SoC. In lead-acid cells the reasons for this include nonlinear kinetics,^{18,19} nucleation and dissolution of lead sulfate,³⁷ hydrolysis during charging,^{38,39} and degradation mechanisms such as sulfation, loss of active material, and electrode corrosion.^{40,41} Because of this, resistance estimates are not a reliable SoH metric unless they are first calibrated to remove the impacts of current, temperature, and SoC. Comprehensively modeling all the underlying physics would result in a model with a large number of parameters,³⁸ and these are challenging to estimate from data.¹⁹ However, machine learning techniques can be used to learn from data the dependency of internal resistance on other factors. To ensure principled treatment of uncertainty associated with field data, we used a Bayesian approach, expressing

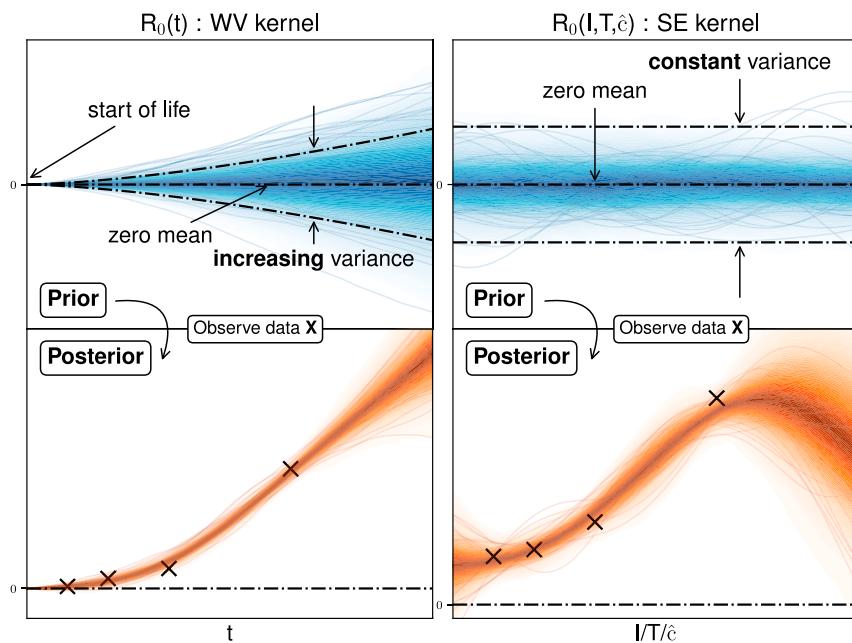


Figure 3. Illustration of the two kernels used to model battery internal resistance as a SoH metric

A GP defines a prior distribution over functions and is conditioned with observations to produce a posterior distribution. Here, random draws from priors (top row) and posteriors (bottom row) are shown for a non-stationary Wiener velocity (WV) kernel used to model the time dependency of resistance, and a stationary squared exponential (SE) kernel with short scale smoothness used to model the dependence of resistance on operating point. Dotted lines show 1σ in the priors. The WV kernel anchors the aging process to zero at start of life and extrapolates its last known trajectory, whereas the SE kernel is non-zero everywhere and reverts to the mean when extrapolated.

the battery internal resistance as a GP over applied current, temperature, estimated electrolyte concentration (i.e., SoC), and time.

In contrast to parametric models,³⁵ GPs offer a flexible probabilistic technique that makes fewer assumptions about the structure of the underlying data.³¹ We assumed that internal resistance over the lifetime of each battery consisted of a sum of two independent GPs, the first capturing the dependency on the instantaneous operating point, and the second the degradation. This improves computational efficiency, although it assumes that the dependence of resistance on age can be decoupled from the dependence of resistance on temperature, SoC, and current. We modeled the operating point dependency using a standard squared exponential kernel and the degradation dependency using a Wiener velocity kernel,⁴² both shown in Figure 3. The former reflects our assumption that the variability of resistance with instantaneous temperature, SoC, and current should be relatively smooth. The latter is a non-stationary kernel,⁴³ enabling degradation at the beginning of life to be zero for each individual battery, and then to grow as the battery ages. This means that extrapolation of future battery health, beyond available data, follows the trajectory learned from the data more accurately.

GP models have “hyperparameters” that describe the smoothness, magnitude, and periodicity of the data being modeled. Fitting to data and then using the resulting model involves two steps. First, the hyperparameter values must be estimated, then the posterior distribution (i.e., mean and variance) of the internal resistance must be evaluated. Both of these steps can be computationally expensive, each

by default scaling with $\mathcal{O}(n^3)$, where n is the number of data points being fitted,³¹ equal to approximately 38,000 per battery. To overcome this challenge, we applied recursive techniques^{32,33} (see [experimental procedures](#)). To benchmark our approach against an existing method,^{7,8} we considered a separate case where R_0 follows a random walk through time and is assumed to be independent of operating conditions. This process is also controlled by a set of hyperparameters for which we found maximum likelihood estimates using the same methodology.

Constructing a consistent battery health metric

To ensure that comparisons of health between batteries were like-for-like, we estimated internal resistance as a function of temperature, SoC, current, and time in the down-selected dataset and then chose a “single” constant set of values of the operating conditions at which to evaluate SoH, for all batteries. This can be thought of as learning a function $R_0 = f(T, I, \hat{c}, t)$ from data, using the GP technique described in the previous section, and then evaluating or slicing through this function at fixed values of all independent variables apart from time.

[Figure 4A](#) shows the variability of internal resistance as a function of temperature, current, and SoC estimated at the population level (see [experimental procedures](#)). A wide range of operating points are observed in the down-selected dataset, as shown in the [Figure 4A](#) insets, and consequently the estimated value of R_0 changes substantially with the instantaneous operating conditions. If this is not accounted for, then these effects would mask the variability caused solely by degradation.

The function describing R_0 that is learned from data is consistent with battery physics. First, the inverse relationship between internal resistance and temperature is due to the Arrhenius dependence of reaction rate on temperature, and results in a variation of up to 0.16Ω in estimated resistance. Second, the shape of internal resistance with respect to acid concentration is due to competing effects. The exchange current density is an increasing function of electrolyte acid concentration,¹⁸ but the transport limitation caused by the reduced rate of dissolution of lead sulfate during charging³⁷ becomes dominant at higher states of charge. The size of this SoC effect was estimated to be approximately 0.09Ω across our dataset, with relatively high uncertainty at lower SoC values, which are visited less often, as shown by the wider credible intervals. Finally, internal resistance reduces as current increases because the relationship between reaction rate and overpotential is nonlinear, specifically, the activation overpotential increases more slowly at higher current densities, as described by Butler-Volmer kinetics.¹⁸ Accounting for this dependency on applied current is key because it has the largest impact on R_0 in the observed operating range, resulting in changes of up to 1.58Ω , although there is high uncertainty associated with the estimate at higher applied currents. This is caused by the batteries spending relatively little time at higher current, resulting in fewer data points and therefore a wider variance of the posterior predictive distribution obtained through GP regression.

To calibrate estimates of battery internal resistance across all batteries, we must choose a reference set of operating conditions (temperature, current, and SoC). As indicated in [Figure 4B](#), there is a trade-off between choosing a single fixed set of calibration conditions for the entire population of batteries versus using separate calibration conditions for each battery trajectory. To obtain a low variance estimate of the resistance of an individual battery, it is best to calibrate using the mean operating conditions for that individual battery, since conditions further from this are associated with higher uncertainty. However, a standardized population-wide calibration condition allows truly like-for-like comparisons to be made between

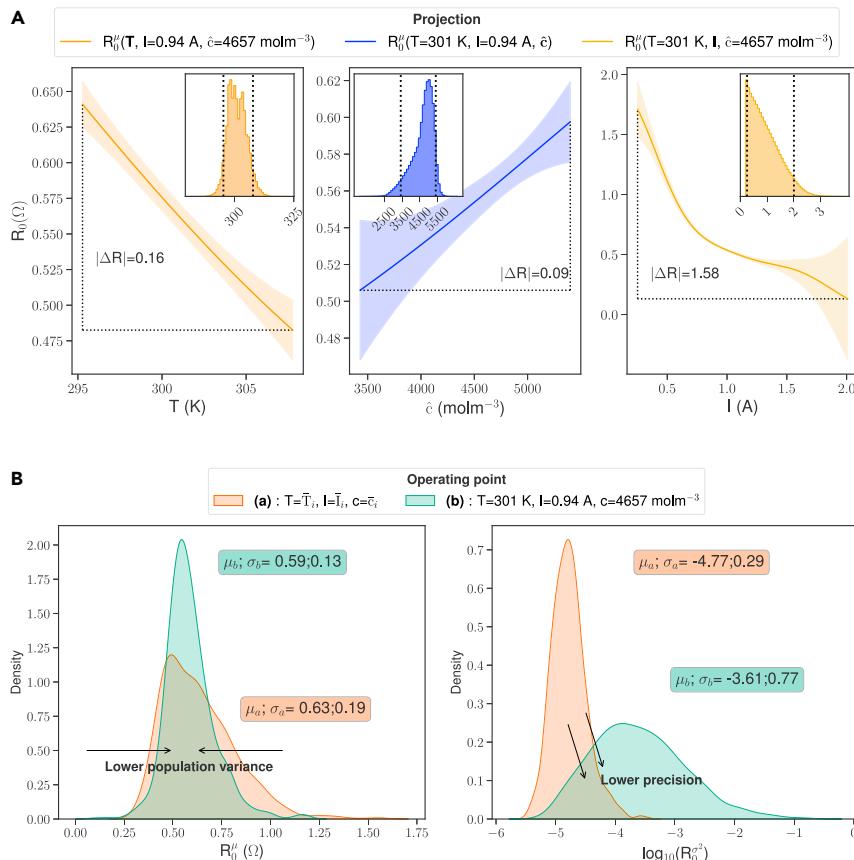


Figure 4. Internal resistance is impacted by operating conditions and population variability

(A) Internal resistance varies significantly with operating conditions, as shown by projections of estimated R_0 as a function of temperature, applied current, and SoC (acid concentration) between 5th and 95th percentiles of each. Shaded regions show $\pm 2\sigma$ credible intervals. Insets show histograms of the independent variables for the down-selected dataset.

(B) There is a trade-off between using a single population-wide calibration point versus a local calibration point per battery. Left-hand image shows lower variance in estimated mean of R_0 values with a single calibration point. However, in this case the uncertainty associated with each R_0 estimate increases, as right-hand image shows.

batteries and gives 37% lower overall standard deviation, so this is what we chose. The remaining population variance reflects cell-to-cell variability at beginning of life resulting from manufacturing variations, storage time, and conditions prior to field deployment.

As a result of the calibration process, we produce trajectories of R_0 that are only a function of time, all referred to a common operating point which was the population mean of temperature, current, and acid concentration. The trajectories are shown in Figure 5. In addition to the absolute value of resistance R_0 , the derivative with respect to time $\partial R_0 / \partial t$ is also a useful health indicator (calculation given in the [supplemental information](#)). Together these two metrics capture whether the battery degradation is beyond the so-called “knee point,” i.e., the onset of accelerated degradation toward end of life.

Predicting and understanding battery end of life

Having constructed consistent battery health trajectories from raw measured data by estimating internal resistance and calibrating it for operating conditions, we can now

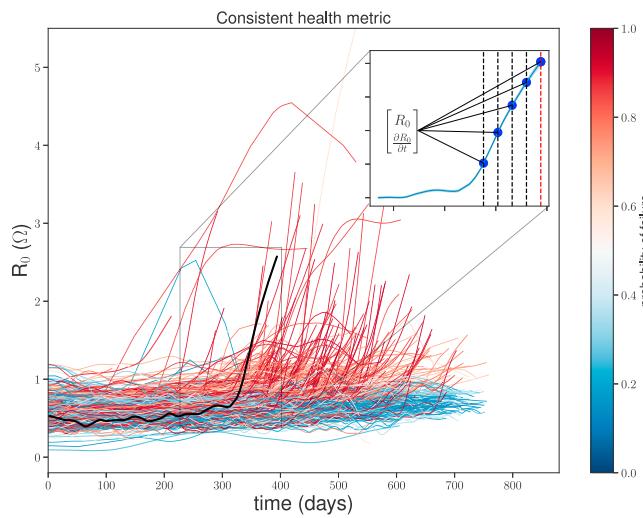


Figure 5. Our approach results in smooth SoH trajectories across the entire dataset of 1,027 batteries

Shown here are calibrated internal resistance profiles estimated from measured data. Color indicates the probability of a battery having failed at the end of each time line, as predicted by a classifier using the features calculated as inputs. The “knee point” is a key indicator of end of life, and as the inset shows, after this the resistance R_0 and its gradient $\partial R_0 / \partial t$ increase.

(1) validate the health metrics by using them to predict end-of-life failure and (2) gain insight into the factors that drive battery aging. To predict battery end of life due to aging, we used repair data provided by BBOXX to label each battery as failed or healthy at the end its respective data time series and trained a GP classifier³¹ to predict this based on health and stress factors at various different prediction horizons. This provides an indirect method of validating the health estimates, although there is uncertainty associated with the labeling of healthy and failed batteries, i.e., there could be false positives and false negatives since repairs are driven by customer decisions. The available data were split into training and test sets, and the classification performance was measured by predicting whether a repair would occur in the test set.

We assessed the performance of this end-of-life classifier under multiple scenarios. First, we used only the health features R_0 and $\partial R_0 / \partial t$ described above as inputs. Second, to benchmark the performance, we used a health metric consisting solely of internal resistance fitted with a random walk, as is common in the literature,^{7,8} without consideration of operating-condition dependency. Third, we considered the case where the health metrics R_0 and $\partial R_0 / \partial t$ were augmented with stress factors known to affect lead-acid battery health.⁴⁰ The latter are features indicative of usage that are also extracted from the raw data, and they consisted of calendar age, charge throughput, cycle count, cumulative time spent at float charge voltage, as well as average temperature and voltage (see the [supplemental information](#) for calculation details). Fourth, we tested the performance using solely the stress factors, omitting information relating to the current SoH as indicated by resistance. To demonstrate the ability to predict end of life in advance with different horizons from 0 to 8 weeks, and with varying ratios of failed versus healthy batteries, we structured the classification performance tests with nested cross validation (see [experimental procedures](#)).

The classifier performance as a function of prediction horizon and according to the various scenarios is shown in [Figure 6A](#). The best overall end-of-life diagnosis and

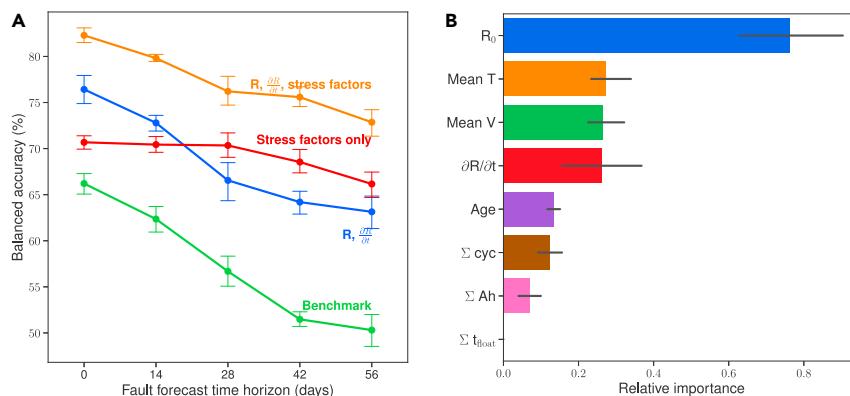


Figure 6. End-of-life prediction performance varies with prediction horizon and model configuration

(A) Combining estimated SoH (resistance and gradient of resistance) with aging stress factors gives best overall end-of-life prediction. Classification performance is quantified using balanced accuracy, for prediction horizons from 0 to 56 days. Error bars show the standard deviation in the average test accuracy grouped by test set, including a varying proportion of failed batteries.

(B) Internal resistance, mean temperature, rate of change of internal resistance with time, and mean voltage are important predictors of end of life. The relative importance of various inputs was quantified using the inverse length scales of the hyperparameters averaged across training sets. The standard deviation of each inverse length scale is indicated by the error bar.

prediction performance occurs when we combine both the current SoH, as indicated by estimated resistance and its gradient, with the stress factors. In this case 82% balanced accuracy (see [Equation 12](#) in [experimental procedures](#) for definition) is achieved at end of life and 73% when the time horizon for prediction is extended to 8 weeks. The [supplemental information](#) gives the full confusion matrix showing classifier performance. In comparison, the benchmark approach and the approach using only estimated SoH—without stress factors—give a lower accuracy which rapidly drops with increasing prediction horizon. This is due to the absence of knee points in SoH as one moves further back in time from the end of life; 2 months before they fail, fewer batteries have experienced the onset of accelerated degradation. At this point the benchmark model performance is near 50% balanced accuracy, which is equivalent to random classification. The classifier using only the stress factors—without resistance estimates—performs well in comparison with the benchmark, achieving an average accuracy of 69% over all horizons. This indicates that aging is being driven by usage conditions and calendar life and is fairly consistent across the population, although introducing battery-specific SoH information near end of life significantly improves predictive accuracy.

The classification technique also gives insight into the factors driving aging. Relative importance was quantified using the inverses of the length-scale hyperparameters of the classifier. Resistance and its rate of change over time are important end-of-life indicators, but other factors, in particular mean temperature and mean voltage, are significant, as shown in [Figure 6B](#). Temperature is known to have a complex impact on lead-acid battery degradation, and the full dataset spans more than 20°C range. Elevated temperatures, especially during charging, may improve lifetime due to improved solubility of lead sulfate,⁴⁰ although they also increase electrode grid corrosion. Similarly, mean voltage is indicative of mechanisms such as sulfation, which occurs at low voltage.⁴⁰ Although some aging factors, such as cycle count, are considered unimportant, this is likely because they are correlated with other factors such as charge throughput and age (see [supplemental information](#)).

Conclusions

Real-world operating data from battery systems in the field may be used to detect end-of-life failure before it happens, improving maintenance, value, safety, and customer experience and giving insights into degradation and performance. We have developed and demonstrated at large scale a data-driven approach for battery SoH estimation and end-of-life prediction using only measured current, voltage, and temperature data while in use, without requiring controlled cycling or systems to be taken offline. Using data from 1,027 solar off-grid lead-acid batteries, each running for 400–760 days, we obtain an end-of-life prediction performance of 73% balanced accuracy, 8 weeks prior to end of life, rising to 82% accuracy at end of life. This accuracy was achieved by combining estimates of SoH with aging stress factors also calculated from the measured data. We used probabilistic machine learning techniques to learn internal resistance as a function of current, temperature, SoC, and time, enabling us to calibrate SoH to consistent reference conditions across the entire population dataset. The success of the approach may be explained by the combination of a population-wide health model with a battery-specific indicator that becomes increasingly important toward end of life. These techniques are applicable to any battery that can be represented with a simple electrical circuit model assuming that sufficient data are available, that degradation manifests as a resistance change, and that C-rates are small. Similar to any data-driven approach, a limitation of the work is that a sufficiently large dataset is required to train the models and demonstrate effectiveness, although a Bayesian approach provides a principled framework for relating uncertainty to the amount of data available. Broadly, this work highlights the opportunity to complement laboratory battery data with large field datasets analyzed through machine learning to improve performance and understanding. An immediate practical benefit is cost savings in the supply chain for battery spare parts through reducing storage times and inventory requirements. Additionally, more accurate lifetime models can be built from the population data—this would be particularly powerful if a large number of solar-battery system organizations are willing to share operational data. Finally, these algorithms could be used for second life screening if sufficient first life battery data are available.

EXPERIMENTAL PROCEDURES**Resource availability***Lead contact*

Inquiries regarding the data and code associated with this paper can be directed to David Howey (david.howey@eng.ox.ac.uk).

Materials availability

This study did not generate new materials.

Data and code availability

Data can be obtained from the Oxford Research Archive: <https://doi.org/10.5287/bodleian:aVR4oDV4N>. Code may be found at the Oxford University Innovation Software Store: <https://process.innovation.ox.ac.uk/software/p/18642/able—advanced-battery-lifetime-estimation—academic-use-only/1>.

Data selection

To obtain a clean dataset and reduce the computational load required for model training, several pre-processing steps were necessary. First, segments of charging data were chosen over the lifetime of each battery. To be eligible, segments had to meet the conditions listed in [Table 1](#).

Table 1. Qualifying charging segment conditions

Criterion	Accepted range
Charging segment duration	>6,000 s
Starting voltage range	11.5–12.9 V
Starting current	<0.1 A
Max(voltage in segment)	>14 V
Max(time gap in recorded data)	<610 s

These conditions ensured that charging segments covered a reasonable SoC range. Additionally, each charging segment was truncated to include voltages only up to 14 V due to increased uncertainty in estimating SoC at higher voltages when using the method described in [inferring acid concentration from current and voltage data](#) because the magnitude of the side reactions increases exponentially with terminal voltage. After down-selecting appropriate charging segments, data were interpolated to a 1-min time grid using piece-wise cubic hermite interpolation.⁴⁴ During charging, data points where $I_t < 0.2$ A, or the estimated OCV was above the measured terminal voltage (which may occur sometimes due to inaccuracy of the SoC estimate), were removed to improve numerical conditioning and ensure resistance estimates were always positive.

Inferring acid concentration from current and voltage data

In lead-acid batteries, electrolyte acid concentration is a direct measure of state of charge—the latter is the normalized version of the former, but requires the maximum concentration to be known, and this changes with age. We therefore simply use acid concentration as a SoC indicator throughout and use the two terms interchangeably in this paper. To infer acid concentration from measured data, we first measured the battery OCV in a laboratory (see [supplemental information](#)) then used this to determine initial acid concentration from measured voltage at points where the charging current was at a minimum. The required OCV curve was measured experimentally using the galvanostatic intermittent titration technique (Biologic SP-150 potentiostat), placing the battery in a thermal chamber (Binder MK53) at 25°C. From this data, the electrolyte volume was also inferred by a least squares fit using Bode's well-known result.⁴⁵ Finally, the trajectory of acid concentration for each charging segment was obtained by Coulomb counting from the initial concentration and using the measured current data, accounting for the known side reactions in lead-acid systems^{38,39} by using a lumped term for the gassing reactions⁴¹ according to

$$\frac{d\hat{c}_t}{dt} = \frac{I_t - I_{\text{gas},0} e^{c_T(T_t - T_0) + c_V(V_t - V_0)}}{FV_{\text{elec}}}, \quad (\text{Equation 1})$$

where I_t , T_t , V_t are the measured current, temperature, and terminal voltage, F is Faraday's constant, V_{elec} is the estimated electrolyte volume, and gassing current parameters $I_{\text{gas},0}$, c_T , T_0 , c_V , V_0 were from literature.⁴¹ The substantial uncertainty due to the variation in these parameters over the lifetime of the battery was taken into consideration by projecting the input uncertainty in the acid concentration \hat{c} to be measurement noise variance in GP regression ([Equations 6A and 6B](#)).

Input data normalization

To ensure comparability of hyperparameters, the current, temperature, and estimated acid concentration I_t , T_t , \hat{c}_t were normalized using the population level moments according to

$$X = \frac{x - \bar{x}}{\sigma_x}, x \in \{T_t, I_t, \hat{c}_t\}, \quad (\text{Equation 2})$$

where \bar{x} and σ_x represent the population mean and variance of the down-selected dataset. A normalized timescale was obtained by dividing the time since beginning of life in seconds by 34,560,000 s, i.e., a nominal 400 days life expressed in seconds, bringing it to a similar range as the other inputs. This method of normalization ([Equation 2](#)) was also used for the inputs for the GP classifier.

Modeling internal resistance as a GP

Battery-specific estimates

Given that the C-rates during charging are low (<0.2C), we may ignore the effect of concentration overpotentials.^{18,19} Therefore the terminal voltage is given by the sum of the OCV and a lumped linearized internal resistance term:

$$V_t = V_0(\hat{c}_t) + R_0(t, I_t, T_t, \hat{c}_t)I_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_{n,t}^2) \quad (\text{Equation 3})$$

where V_t is the terminal voltage and t, I_t, T_t , and \hat{c}_t are the (normalized) time since beginning of life, applied current (positive for charging by convention), measured temperature, and estimated bulk sulfuric acid concentration at time t , respectively. The experimentally obtained OCV as a function of acid concentration is V_0 and the dependency $R_0(I_t, T_t, \hat{c}_t, t)$ is modeled by a zero-mean GP:

$$R_0(x) \sim \mathcal{GP}(0, k(x)), \quad x \in \{t, I, T, \hat{c}_t\}, \quad (\text{Equation 4})$$

where k is the covariance function. We chose to model R_0 as a sum of two GPs, where the degradation process as a function of time is described by a Wiener velocity (WV) kernel,⁴³ and the dependency of R_0 on operating point (I_t, T_t, \hat{c}_t) is described by a squared exponential (SE) kernel, such that

$$k(I, T, \hat{c}, t) = \sigma_{f,0}^2 \left(\frac{\min^3(t, t')}{3} + |t - t'| \frac{\min^2(t, t')}{2} \right) + \sigma_{f,1}^2 \exp \left(\sum_{x \in \{I, T, \hat{c}\}} -\frac{|x - x'|^2}{2l_x^2} \right), \quad (\text{Equation 5})$$

where $|.|$ denotes the absolute distance between two points. The choice of the WV process, used in target tracking applications,⁴⁶ allows for better extrapolation outside the observed data because the kernel is non-stationary. In comparison, using the zero-mean SE kernel would result in extrapolation over longer time horizons tending to the prior distribution. Examples of random draws from the prior distributions for each kernel are shown in [Figure 3](#).

Extrapolation over time is necessary in our case because the down-selected charging segments are not evenly distributed over time across the population. Extrapolation is thus required to give estimates of R_0 at points 0–56 days preceding the data series end for each battery. Additionally, expressing R_0 as a sum of kernels, [Equation 5](#), makes the assumption that the degradation is purely additive, and independent of the operating point chosen. This significant simplification reduces the degrees of freedom of the system in comparison with a product of kernels over all inputs and makes inference computationally lower cost.

To fit the model defined by [Equations 3, 4](#), and [5](#), the hyperparameters to be estimated are the two process variances, $\sigma_{f,0}^2, \sigma_{f,1}^2$, and length scales l_x across the inputs, (I_t, T_t, \hat{c}_t) . The model is fitted by adjusting these hyperparameters according to a cost function (defined in [Equation S15](#) of the [supplemental information](#)), which includes the error between the modeled and measured voltage. The fitted function is given by the posterior distribution of R_0 for each battery. Special treatment was given to

the noise ε_t in [Equation 3](#). There may be considerable uncertainty in the OCV function $V_0(\hat{c})$, for example, due to hysteresis, as well as in the estimate of acid concentration \hat{c} , due to uncertainty in the parameters in [Equation 1](#). Additionally, our experimental estimation of the OCV curve and electrolyte volume (see [supplemental information](#)) was done using a single battery, so variations in these across the population are another source of uncertainty. To account for this, we assumed a 10% uncertainty in $d\hat{c}/dt$, together with a 150-mV standard deviation caused by voltage measurement and OCV uncertainty, giving a total variance per charge segment as

$$\sigma_{n,t}^2 = 0.0225 + \text{Var}(\hat{c}_t) \left(\frac{dV_0}{d\hat{c}} \right)^2 \quad (\text{Equation 6A})$$

$$\text{Var}(\hat{c}_t) = \sum_t 0.01 \Delta \hat{c}_t^2, \quad (\text{Equation 6B})$$

where $\Delta \hat{c}_t = (dc_t/dt)\Delta t$, assuming dc_t/dt is constant in each time segment Δt . In summary, the GP regression includes heteroskedastic noise $\sigma_{n,t}^2$ which we pre-calculate rather than estimate as a hyperparameter in the fitting process.

To impose smoothness on the function R_0 over all input dimensions, we assumed a prior distribution over hyperparameters:

$$p(\sigma_{f,0}, \sigma_{f,1}, l_T, l_c) = \prod_{m \in \{0,1\}} \chi(\sigma_{f,m}, k=1, s=0.2) \prod_{x \in \{l, T, \hat{c}\}} \Gamma^{-1}(l_x, \alpha=1, \beta=2), \quad (\text{Equation 7})$$

where χ represents the chi distribution and Γ^{-1} the inverse gamma distribution. The chi distribution with the degree of freedom $k=1$ is equal to a half-normal distribution with standard deviation $s=0.2$. We assumed an inverse gamma prior for the length scales, with shape and scale parameters α and β chosen to give a mode of 1.

To estimate the posterior distribution of R_0 for each battery and to recover maximum-a-posteriori (MAP) estimates of the hyperparameters, we employed a recursive estimation framework for GPs,[32,33,47](#) where they are interpreted to be the solution of a stochastic partial differential equation of the form

$$\frac{\partial R_0(t, l, T, \hat{c})}{\partial t} = FR_0(t, l, T, \hat{c}) + L\omega(t, l, T, \hat{c}), \quad (\text{Equation 8})$$

where the transition matrix F , dispersion matrix L , and the properties of the white noise process $\omega(t, l, T, \hat{c})$ are determined by the kernel function, [Equation 5](#). This framework allows the use of standard Kalman filtering and smoothing techniques[43,48](#) to estimate both the posterior distribution of R_0 and the so-called energy function that is the negative unnormalized logarithm of the posterior probability of the hyperparameter vector. Crucially, this method scales as $\mathcal{O}(n)$ over the number of data rows, which in our case is order 10^4 for each battery.

To obtain a finite-dimensional representation of the system in [Equation 8](#), we used a similar approach to that of Särkkä et al.[32](#) First, we applied a k-means algorithm to choose 20 representative points across (l_t, T_t, \hat{c}_t) , at which to estimate R_0 through the lifetime of the battery. In order to obtain estimates for R_0 for the observed (l_t, T_t, \hat{c}_t) through all charging segments, the predictive distribution of the GP over the operating points was calculated and added to the value predicted by the degradation GP.

Using these recursive techniques, we fitted regression hyperparameters across the entire population of batteries, consisting of 39 million rows of data, in approximately 80 minutes using an Apache Spark cluster running 30 cores.

Population level model over operating points

Population level estimates of R_0 as a function of all the operating points were calculated by fitting a GP over the individual battery R_0 estimates (i.e., the solutions of [Equation 8](#)) at each battery's mean operating point at the beginning of life. This was done using the standard batch GP approach,³¹ with the SE kernel and hyperprior for the magnitude and length scales equivalent to the recursive case. The hyperprior including the measurement noise was then

$$p(\sigma_f, \sigma_n, l_t, l_i, l_c) = \chi(\sigma_n, k=1, s=0.1)\chi(\sigma_f, k=1, s=0.2) \prod_{x \in \{l, T, c\}} \Gamma^{-1}(l_x, \alpha=1, \beta=2). \quad (\text{Equation 9})$$

In addition, we added the variance estimates for each R_0 value to the measurement noise. Given MAP estimates retrieved by the L-BFGS-B algorithm,⁴⁹ the posterior means and variances for the population level R_0 function were given by

$$\mu_p = k(x^*, X) [k(X, X) + \sigma_n^2 I + \sigma_R^2(X)]^{-1} R_0(X) \quad (\text{Equation 10A})$$

$$\sigma_p^2 = k(x^*, x^*) - k(x^*, X) [k(X, X) + \sigma_n^2 I + \sigma_R^2(X)]^{-1} k(X, x^*), \quad (\text{Equation 10B})$$

where k is the SE kernel function and $R_0(X)$ and $\sigma_R^2(X)$ denote the mean and variance of the battery-wise R_0 estimates at operating point X —in this case X was the mean operating for the battery. The predictive distribution was calculated on a grid x^* such that the range for each input variable (temperature, applied current, and estimated acid concentration) was varied between its 5th and 95th percentile in turn while keeping the other two constant at the population mean.

Benchmark model

To benchmark our approach, we used the same model ([Equation 3](#)) but without the dependency of R_0 on the operating point. Furthermore, we assumed that R_0 followed a random walk through time, an approach commonly taken in the literature to adapt parameters to data.⁵⁰ Using this recursive approach, the tuning parameters are the process and noise covariances as well as the initial variance estimate for R_0 , which are often set manually. For consistency with the main approach, we estimated the process noise covariance using maximum likelihood and set the initial variance to be a constant multiple of the process covariance and the noise covariance was calculated in the same way ([Equations 6A](#) and [6B](#)) as for the main method.

End-of-life prediction using Gaussian process classification

To classify the batteries as failed or healthy at various time horizons up to the end date of each battery data series, we used a standard GP classification framework.³¹ The inputs $R_0, \frac{\partial R_0}{\partial t}$ were extrapolated as required from the last observed charging segment to the appropriate point in time preceding the end of the time series. In all cases, inputs were fed into a GP classifier using a SE covariance function with automatic relevance detection:

$$k_{\text{class}}(x, x') = \sigma_f^2 \exp\left(\sum_x -\frac{|x - x'|^2}{2l_x^2}\right) \quad (\text{Equation 11})$$

with a uniform hyperprior. Maximum likelihood estimates for the hyperparameters were then obtained by minimizing the negative log marginal likelihood of the data, using the GP classifier implementation in the Scikit-learn toolbox.⁵¹ The performance metrics chosen for the classifier were such that the unevenness in labeling (536 healthy, 491 failed) was taken into account. The balanced accuracy metric was calculated as the average of classifier sensitivity and specificity:

$$\text{Balanced accuracy} = \frac{1}{2} (\text{Sensitivity} + \text{Specificity}) = \frac{1}{2} \left(\frac{\text{TP}}{\text{TP} + \text{FN}} + \frac{\text{TN}}{\text{TN} + \text{FP}} \right), \quad (\text{Equation 12})$$

where TP, TN, FP, and FN represent the true positive, true negative, false positive, and false negative rates, respectively.

To quantify classifier performance, first, we used 5-fold stratified cross validation to split the dataset into training and test sets. The training set in each case had a 48%/52% proportion of failed versus healthy batteries using data up to the end of each time series, training 20 classifiers altogether (4 sets of inputs and 5 training sets). Then, for each outer cross validation stage, test subsets were selected, each containing varying proportions of failed batteries. These subsets were chosen by keeping the healthy battery set fixed and then randomly sampling 40%, 60%, and 80% of the failed batteries in the test set, repeated ten times, plus the case when 100% of the failed batteries was used, giving a total of 3,100 test cases. All performance metrics reported are the average for the test sets using 5-fold cross validation for each test case.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.joule.2021.11.006>.

ACKNOWLEDGMENTS

The authors thank BBOXX for access to data from PV-connected battery energy storage systems; D. Lovell for assistance with Figure 1; and V. Sulzer, R. Drummond, J. Reniers, and S. Cooper for manuscript feedback. We acknowledge funding from the Faraday Institution (EP/S003053/1, grant number FIRG003) and Shell Foundation (agreement 22077).

AUTHOR CONTRIBUTIONS

Conceptualization, D.A.H.; data curation, software, and visualization, A.A.; writing – original draft, review, and editing, A.A. and D.A.H.; supervision, project administration, and funding acquisition, D.A.H.

DECLARATION OF INTERESTS

The authors have filed a patent related to this work: GB Application no. 2105995.1, dated 27 April 2021. D.A.H. is co-founder of Brill Power and is a technical adviser at Habitat Energy. A.A. declares no competing interests.

Received: August 20, 2021

Revised: September 24, 2021

Accepted: November 17, 2021

Published: December 15, 2021

REFERENCES

- ESMAP (2019). Mini Grids for Half a Billion People: Market Outlook and Handbook for Decision Makers. Technical Report (World Bank). <https://sdgs.un.org/sites/default/files/2021-05/Mini%20Grids%20For%20Half%20A%20Billion%20People%20Market%20Outlook%20And%20Handbook%20For%20Decision%20Makers.pdf>.
- Lee, J.T., and Callaway, D.S. (2018). The cost of reliability in decentralized solar power systems in sub-Saharan Africa. *Nat. Energy* 3, 960–968. <https://doi.org/10.1038/s41560-018-0240-y>.
- Farmann, A., Waag, W., Marongiu, A., and Sauer, D.U. (2015). Critical review of on-board capacity estimation techniques for lithium-ion batteries in electric and hybrid electric vehicles. *J. Power Sources* 281, 114–130. <https://doi.org/10.1016/j.jpowsour.2015.01.129>.
- Berecibar, M., Gandiaga, I., Villarreal, I., Omar, N., Van Mierlo, J., and Van Den Bossche, P. (2016). Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renew. Sustain. Energy Rev.* 56, 572–587. <https://doi.org/10.1016/j.rser.2015.11.042>.
- Xiong, R., Li, L., and Tian, J. (2018). Towards a smarter battery management system: a critical review on battery state of health monitoring methods. *J. Power Sources* 405, 18–29. <https://doi.org/10.1016/j.jpowsour.2018.10.019>.
- Li, Y., Liu, K., Foley, A.M., Zülke, A., Berecibar, M., Nanini-Maury, E., et al. (2019). Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renew. Sustain. Energy Rev.* 113, 109254. <https://doi.org/10.1016/j.rser.2019.109254>.
- Plett, G.L. (2004). Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. State and parameter estimation. *J. Power Sources* 134, 277–292. <https://doi.org/10.1016/J.JPOWSOUR.2004.02.033>.
- Plett, G.L. (2006). Sigma-point Kalman filtering for battery management systems of LiPB-based HEV battery packs. Part 2: simultaneous state and parameter estimation. *J. Power Sources* 161, 1369–1384. <https://doi.org/10.1016/j.jpowsour.2006.06.004>.
- Baba, A., and Adachi, S. (2015). Simultaneous state of charge and parameter estimation of lithium-ion battery using log-normalized unscented Kalman filter. In 2015 American Control Conference (ACC) (IEEE Publications), pp. 311–316.
- Kim, I.S. (2010). A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer. *IEEE Trans. Power Electron.* 25, 1013–1022. <https://doi.org/10.1109/TPEL.2009.2034966>.
- Kim, J., Lee, S., and Cho, B.H. (2012). Complementary cooperation algorithm based on DEKF combined with pattern recognition for SOC/capacity estimation and SOH prediction. *IEEE Trans. Power Electron.* 27, 436–451. <https://doi.org/10.1109/TPEL.2011.2158554>.
- Moura, S.J., Chaturvedi, N.A., and Krstić, M. (2014). Adaptive partial differential equation observer for battery state-of-charge/state-of-health estimation via an electrochemical model. *J. Dyn. Syst. Meas. Control* 136, 011015. <https://doi.org/10.1115/1.4024801>.
- Gomez, J., Nelson, R., Kalu, E.E., Weatherspoon, M.H., and Zheng, J.P. (2011). Equivalent circuit model parameters of a high-power Li-ion battery: thermal and state of charge effects. *J. Power Sources* 196, 4826–4831. <https://doi.org/10.1016/j.jpowsour.2010.12.107>.
- Doyle, M., Fuller, T.F., and Newman, J. (1993). Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *J. Electrochem. Soc.* 140, 1526. <https://doi.org/10.1149/1.2221597>.
- Fuller, T.F., Doyle, M., and Newman, J. (1994). Simulation and optimization of the dual lithium ion insertion cell. *J. Electrochem. Soc.* 141, 1. <https://doi.org/10.1149/1.2054684>.
- Jokar, A., Rajabloo, B., Désilets, M., and Lacroix, M. (2016). Review of simplified pseudo-two-dimensional models of lithium-ion batteries. *J. Power Sources* 327, 44–55. <https://doi.org/10.1016/j.jpowsour.2016.07.036>.
- Marquis, S.G., Sulzer, V., Timms, R., Please, C.P., and Chapman, S.J. (2019). An asymptotic derivation of a single particle model with electrolyte. *J. Electrochem. Soc.* 166, A3693–A3706. <https://doi.org/10.1149/2.0341915jes>.
- Sulzer, V., Chapman, S.J., Please, C.P., Howey, D.A., and Monroe, C.W. (2019a). Faster lead-acid battery simulations from porous-electrode theory: Part I. Physical model. *J. Electrochem. Soc.* 166, A2363–A2371. <https://doi.org/10.1149/2.0301910jes>.
- Sulzer, V., Chapman, S.J., Please, C.P., Howey, D.A., and Monroe, C.W. (2019b). Faster lead-acid battery simulations from porous-electrode theory: Part II. Asymptotic analysis. *J. Electrochem. Soc.* 166, A2372–A2382. <https://doi.org/10.1149/2.0441908jes>.
- Chu, Z., Plett, G.L., Trimboli, M.S., and Ouyang, M. (2019). A control-oriented electrochemical model for lithium-ion battery, Part I: Lumped-parameter reduced-order model with constant phase element. *J. Energy Storage* 25, 100828. <https://doi.org/10.1016/j.est.2019.100828>.
- Richardson, R.R., Birk, C.R., Osborne, M.A., and Howey, D.A. (2019). Gaussian process regression for *in situ* capacity estimation of lithium-ion batteries. *IEEE Trans. Ind. Inform.* 15, 127–138. <https://doi.org/10.1109/TII.2018.2794997>.
- Yang, D., Zhang, X., Pan, R., Wang, Y., and Chen, Z. (2018). A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *J. Power Sources* 384, 387–395. <https://doi.org/10.1016/j.jpowsour.2018.03.015>.
- Klass, V., Behm, M., and Lindbergh, G. (2014). A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. *J. Power Sources* 270, 262–272. <https://doi.org/10.1016/j.jpowsour.2014.07.116>.
- Chaoui, H., Ibe-Ekeocha, C.C., and Gualous, H. (2017). Aging prediction and state of charge estimation of a LiFePO₄ battery using input time-delayed neural networks. *Electr. Power Syst. Res.* 146, 189–197. <https://doi.org/10.1016/j.epsr.2017.01.032>.
- Li, W., Sengupta, N., Dechant, P., Howey, D., Annaswamy, A., and Sauer, D.U. (2021). Online capacity estimation of lithium-ion batteries with deep long short-term memory networks. *J. Power Sources* 482, 228863. <https://doi.org/10.1016/j.jpowsour.2020.228863>.
- dos Reis, G., Strange, C., Yadav, M., and Li, S. (2021). Lithium-ion battery data and where to find it. *Energy and AI* 5. <https://doi.org/10.1016/j.egyai.2021.100081>.
- Sulzer, V., Mohtat, P., Aitio, A., Lee, S., Yeh, Y.T., Steinbacher, F., Khan, M.U., Lee, J.W., Siegel, J.B., David, A., and Stefanopoulou, A.G. (2021). The challenge and opportunity of battery lifetime prediction from field data. *Joule* 5, 1934–1955. <https://doi.org/10.1016/j.joule.2021.06.005>.
- Song, L., Zhang, K., Liang, T., Han, X., and Zhang, Y. (2020). Intelligent state of health estimation for lithium-ion battery pack based on big data analysis. *J. Energy Storage* 32, 101836. <https://doi.org/10.1016/j.est.2020.101836>.
- Wang, Q., Wang, Z., Zhang, L., Liu, P., and Zhang, Z. (2020). A novel consistency evaluation method for series-connected battery systems based on real-world operation data. *IEEE Trans. Transp. Electrif.* 7, 437–451. <https://doi.org/10.1109/TTE.2020.3018143>.
- Aykol, M., Gopal, C.B., Anapolksky, A., Herring, P.K., van Vlijmen, B., Berliner, M.D., Bazant, M.Z., Braatz, R.D., Chueh, W.C., and Storey, B.D. (2021). Perspective-combining physics and machine learning to predict battery lifetime. *J. Electrochem. Soc.* 168, 030525. <https://doi.org/10.1149/1945-7111/abec55>.
- Rasmussen, C.E., and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. (MIT Press).
- Särkkä, S., and Hartikainen, J. (2012). Infinite-dimensional Kalman filtering approach to spatio-temporal Gaussian process regression. In Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, N.D. Lawrence and M. Girolami, eds. (PMLR), pp. 993–1001.
- Särkkä, S., Solin, A., and Hartikainen, J. (2013). Spatiotemporal learning via infinite-dimensional bayesian filtering and smoothing: a look at Gaussian process regression through Kalman filtering. *IEEE Signal Process. Mag.* 30, 51–61. <https://doi.org/10.1109/MSP.2013.2246292>.
- Plett, G.L. (2011). Recursive approximate weighted total least squares estimation of battery cell total capacity. *J. Power Sources* 196, 2319–2331. <https://doi.org/10.1016/j.jpowsour.2010.09.048>.
- Remmlinger, J., Buchholz, M., Meiler, M., Bernreuter, P., and Dietmayer, K. (2011). State-

- of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *J. Power Sources* 196, 5325–5331. <https://doi.org/10.1002/anie.201504913>.
36. Kollmeyer, P.J., and Jahns, T.M. (2019). Aging and performance comparison of absorbed glass matte, enhanced flooded, PbC, NiZn, and LiFePO₄ 12V start stop vehicle batteries. *J. Power Sources* 441, 227139. <https://doi.org/10.1016/j.jpowsour.2019.227139>.
37. Huck, M., and Sauer, D.-U. (2020). Modeling transient processes in lead-acid batteries in the time domain. *J. Energy Storage* 29, 101430. <https://doi.org/10.1016/j.est.2020.101430>.
38. Newman, J., and Tiedemann, W. (1997). Simulation of recombinant lead-acid batteries. *J. Electrochem. Soc.* 144, 3081. <https://doi.org/10.1149/1.1837963>.
39. Bernardi, D.M., and Carpenter, M.K. (1995). A mathematical model of the oxygen-recombination lead-acid cell. *J. Electrochem. Soc.* 142, 2631–2642. [https://doi.org/10.1016/0021-9169\(92\)90162-9](https://doi.org/10.1016/0021-9169(92)90162-9).
40. Ruetschi, P. (2004). Aging mechanisms and service life of lead-acid batteries. *J. Power Sources* 127, 33–44. <https://doi.org/10.1016/j.jpowsour.2003.09.052>.
41. Schiffer, J., Sauer, D.U., Bindner, H., Cronin, T., Lundsager, P., and Kaiser, R. (2007). Model prediction for ranking lead-acid batteries according to expected lifetime in renewable energy systems and autonomous power-supply systems. *J. Power Sources* 168, 66–78. <https://doi.org/10.1016/j.jpowsour.2006.11.092>.
42. Särkkä, S., and Solin, A. (2019). *Applied Stochastic Differential Equations* (Cambridge University Press).
43. Solin, A. (2016). *Stochastic differential equation methods for spatio-temporal Gaussian process regression* (Aalto University), Ph.D. thesis.
44. Fritsch, F.N., and Carlson, R.E. (1980). Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* 17, 238–246.
45. Bode, H. (1977). *Lead-Acid Batteries* (Wiley).
46. Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation* (Wiley).
47. Hartikainen, J., and Särkkä, S. (2010). Kalman filtering and smoothing solutions to temporal Gaussian process regression models. In *Proceedings of the 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010)*, pp. 379–384. <https://doi.org/10.1109/MLSP.2010.5589113>.
48. Särkkä, S. (2013). *Bayesian Filtering and Smoothing* (Cambridge University Press).
49. Byrd, R.H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* 16, 1190–1208. <https://doi.org/10.1137/0916069>.
50. Plett, G.L. (2015). *Battery Management Systems, Volume II: Equivalent-Circuit Methods* (Artech House Publishers).
51. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.