Data-driven battery health diagnostics

David Howey, Antti Aitio
University of Oxford
david.howey@eng.ox.ac.uk

ECC Workshop, Future Automotive Systems, July 2022

Battery Intelligence Lab

University of Oxford

Outline

The challenge of state of health estimation from field data

- Why do we care?
- Why is it so difficult?
- Existing methods

Outline

The challenge of state of health estimation from field data

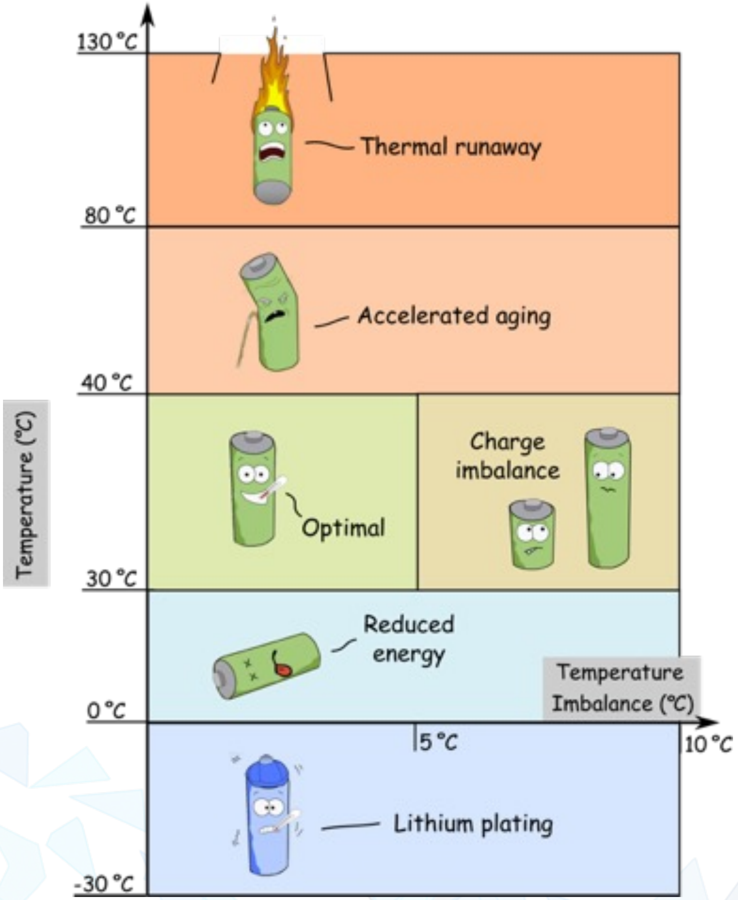
- Why do we care?
- Why is it so difficult?
- Existing methods

Our approach: joining model- and data-driven methods

- Gaussian process regression and equivalent circuit models
- SOH diagnosis in PV-connected batteries in sub-Saharan Africa
- Parameterisation of more complex models using drive cycle data

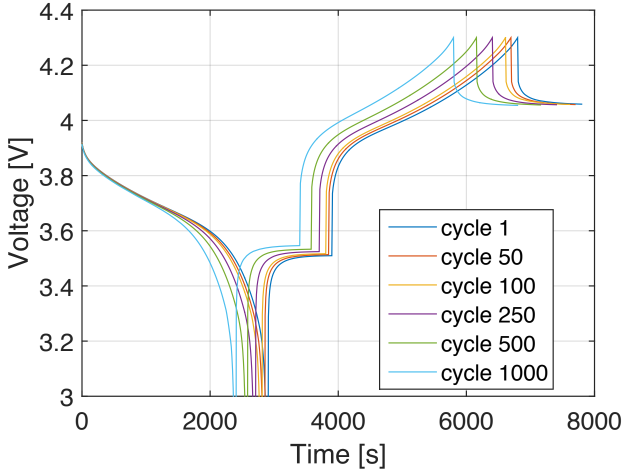
Batteries are a lot like people: they need looking after

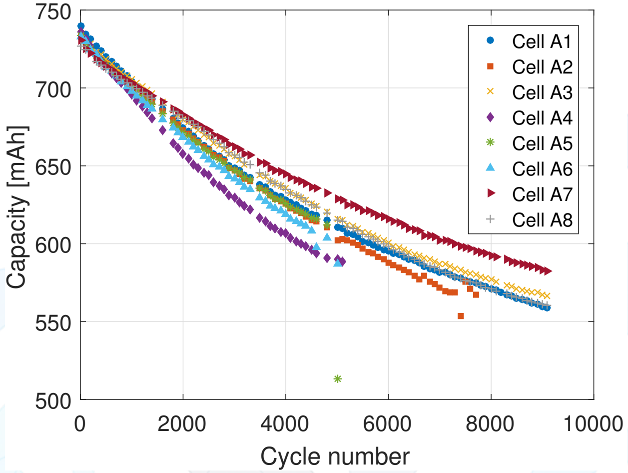
Thermally fragile



Mechanically fragile

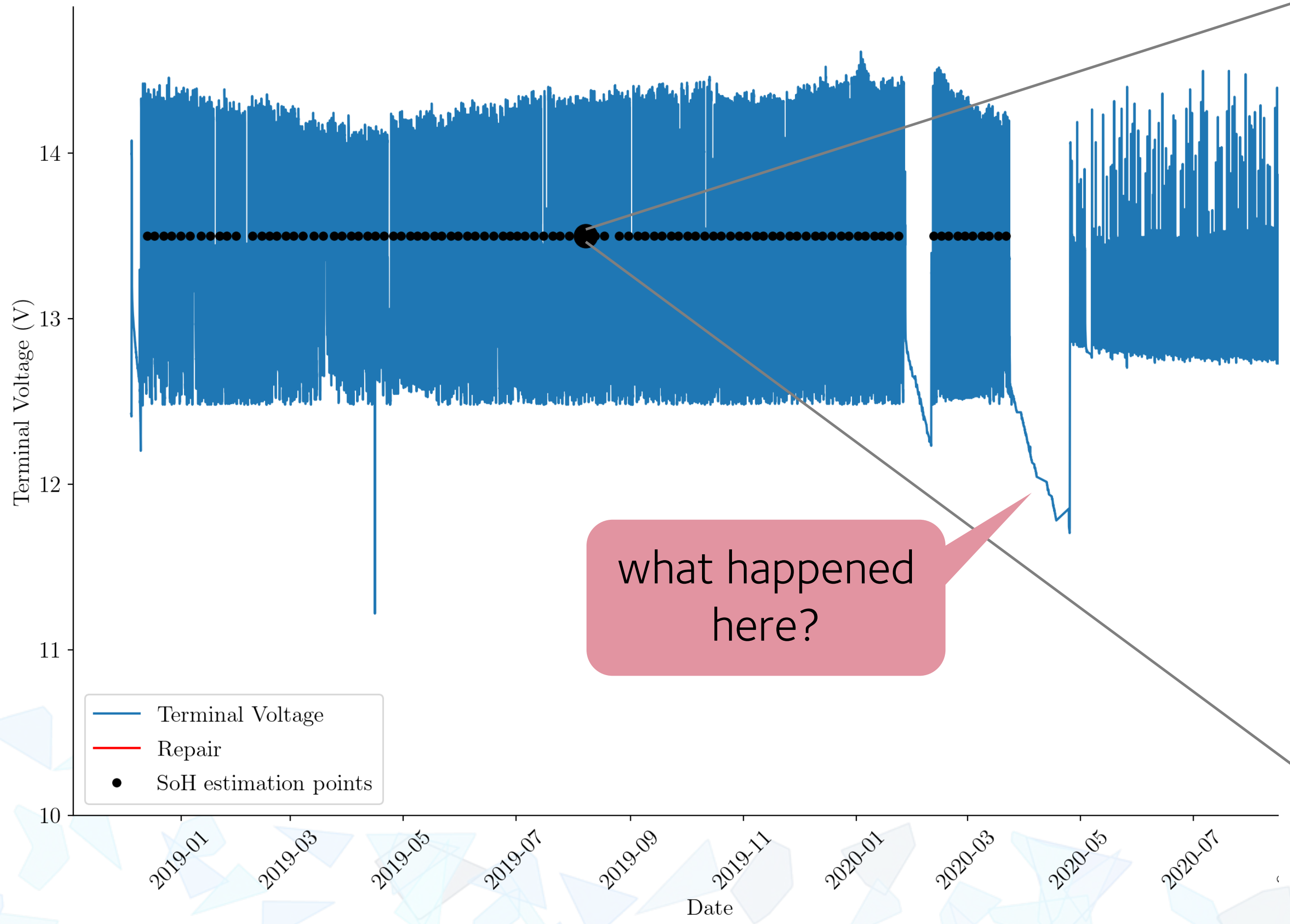
Degrade over time

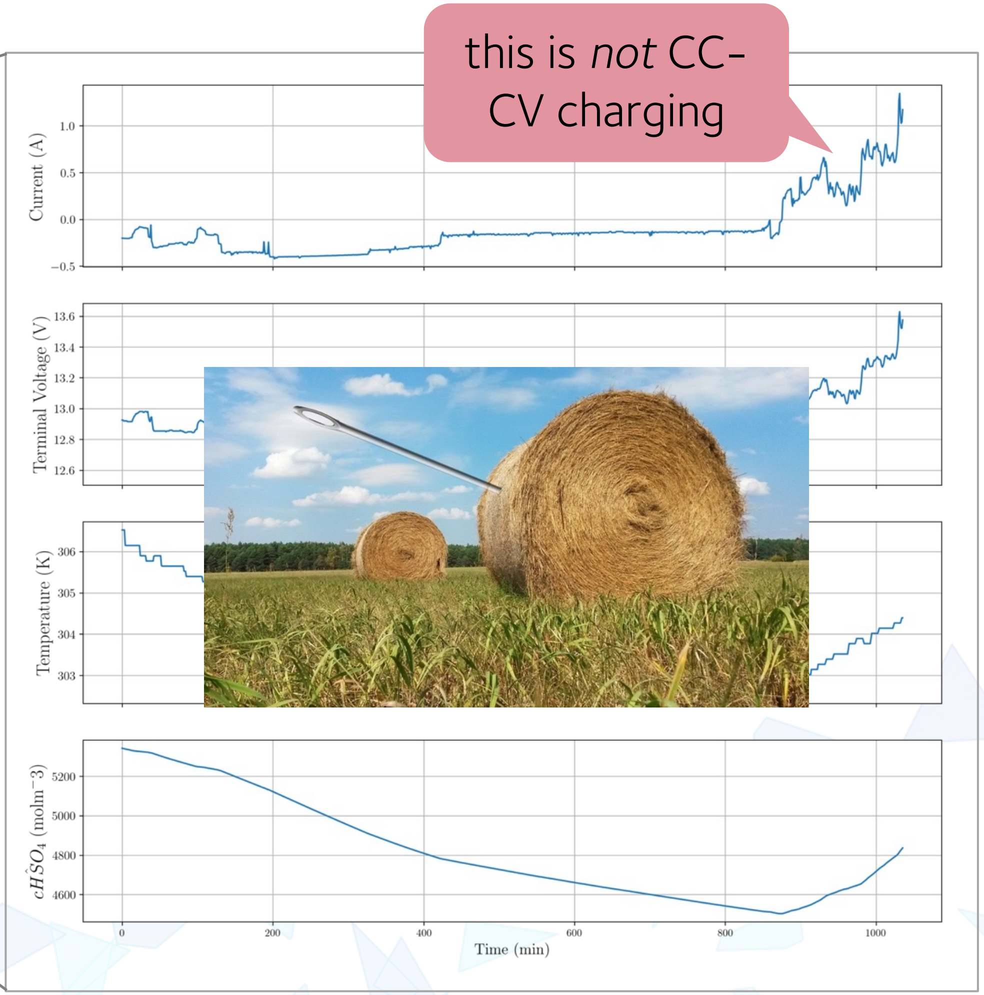




Pictures: Rob Richardson, Wikimedia commons, Adrien Bizeray, Christoph Birkl

Estimating SOH from field data is difficult



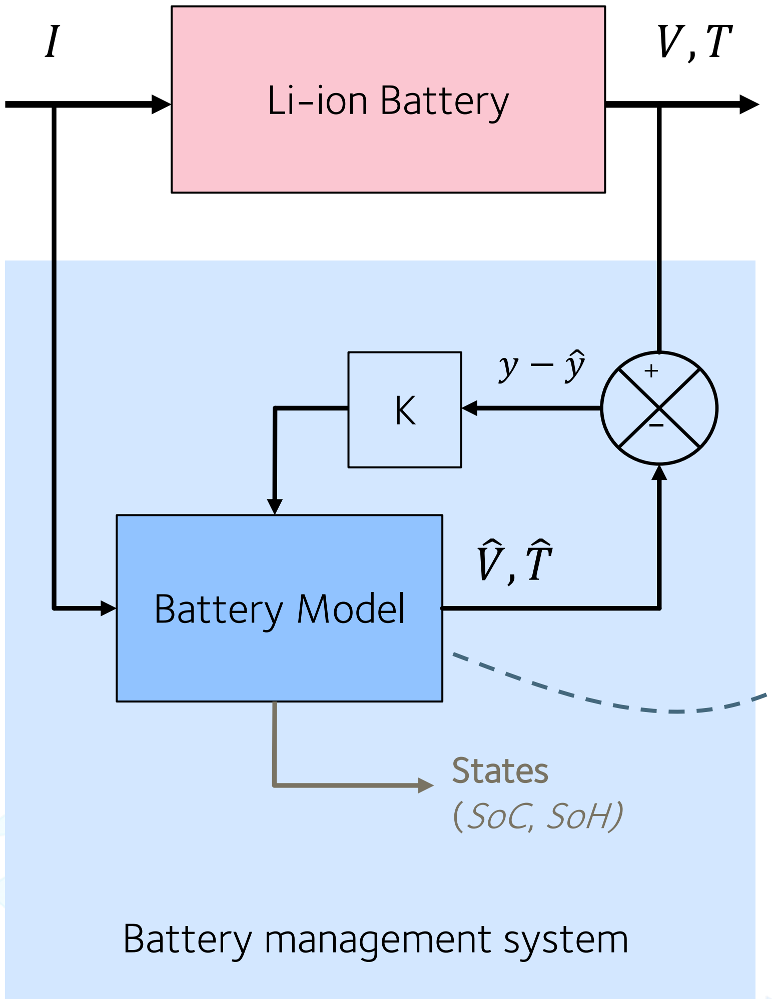


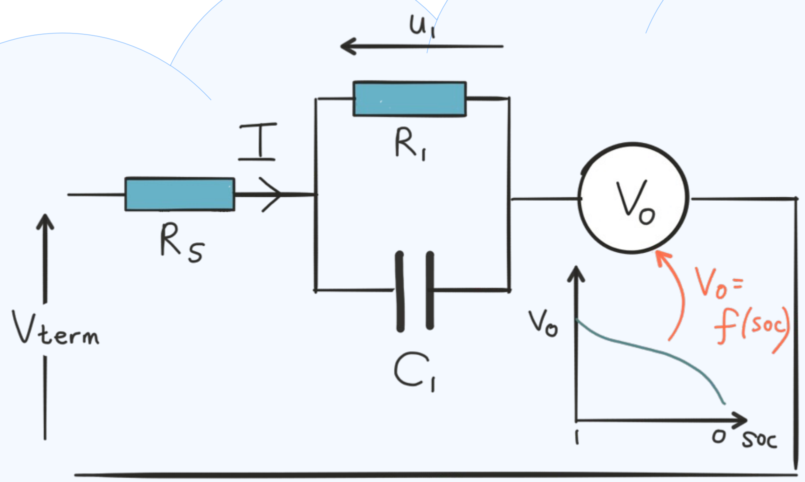
Lead-acid battery data from BBOXX Ltd.

Pictures: Antti Aitio, Pixabay

Existing methods of battery health diagnosis

The model/observer-based approach for SOH diagnosis is common





$$\begin{bmatrix}\dot{u}_1\\ \dot{u}_2\end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1C_1} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix} + \begin{bmatrix}\frac{1}{C_1}\\ \frac{1}{Q}\end{bmatrix} I$$

SOC

capacity

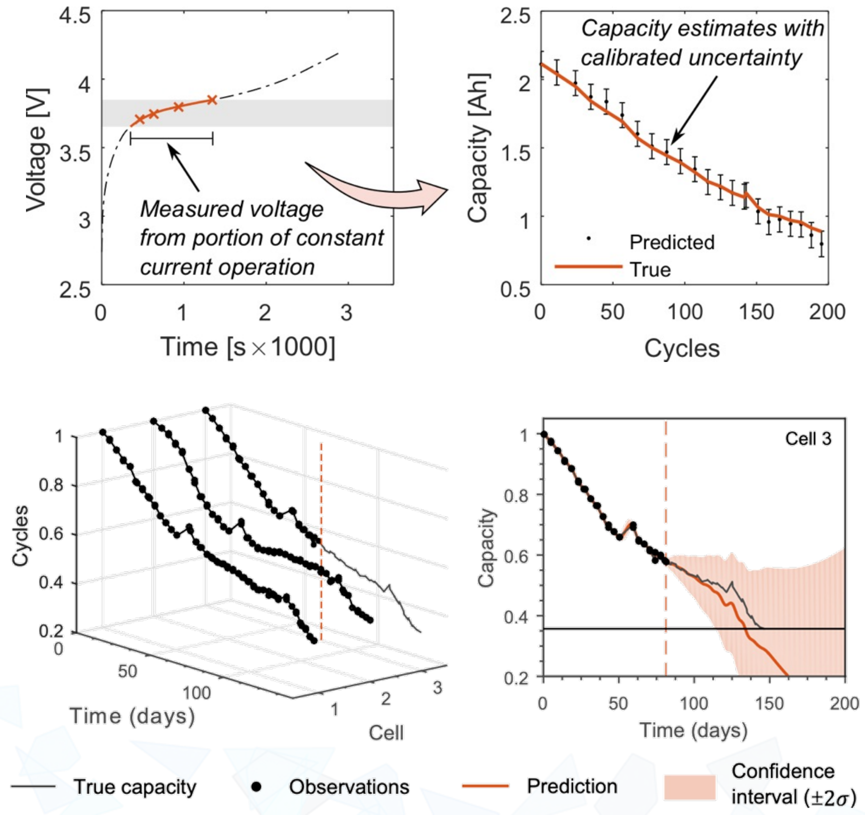
$$V_{term} = V_0(u_2) + \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix} + R_S I$$

resistance

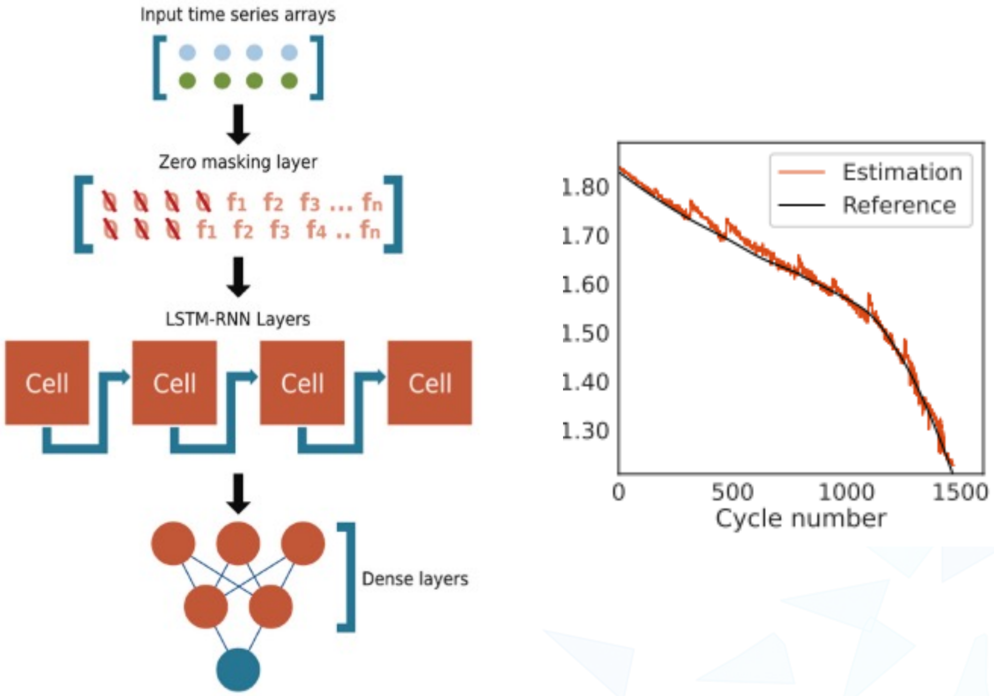
Adapt parameters: $\theta_k = \theta_{k-1} + r_{k=1}$

'Pure' data-driven methods work but may struggle to generalize

- Use non-linear mapping from operating data to SOH
- Flexible – but can it generalize?



Richardson, R.R. et al. (2017)
Richardson, R.R. et al. (2018).

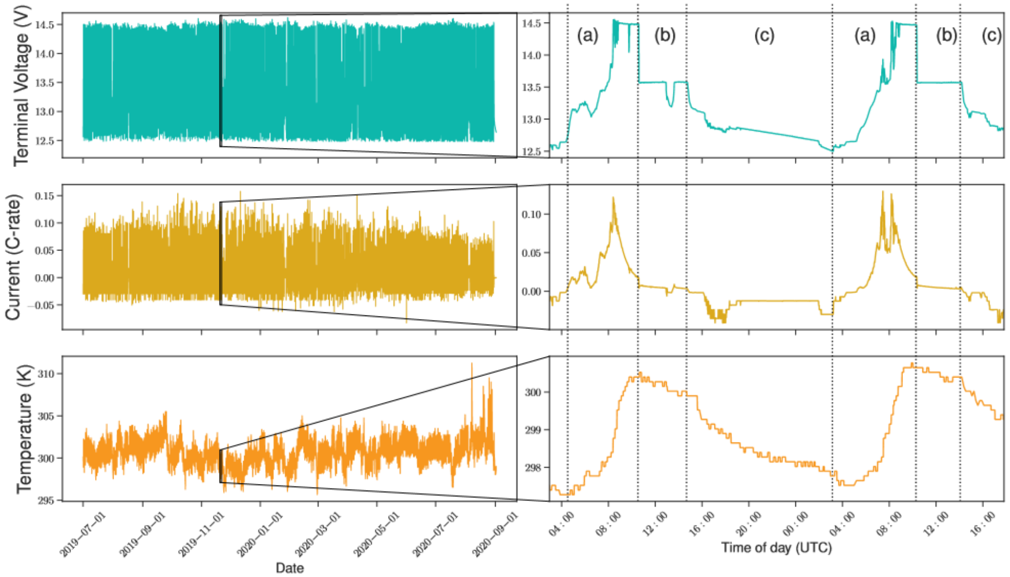


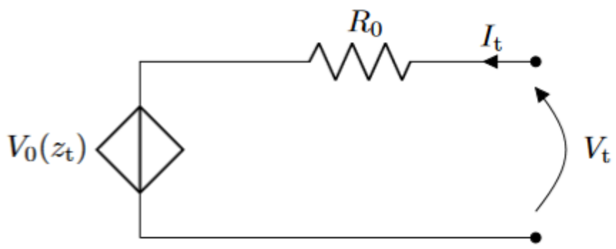
Li et al. (2021)

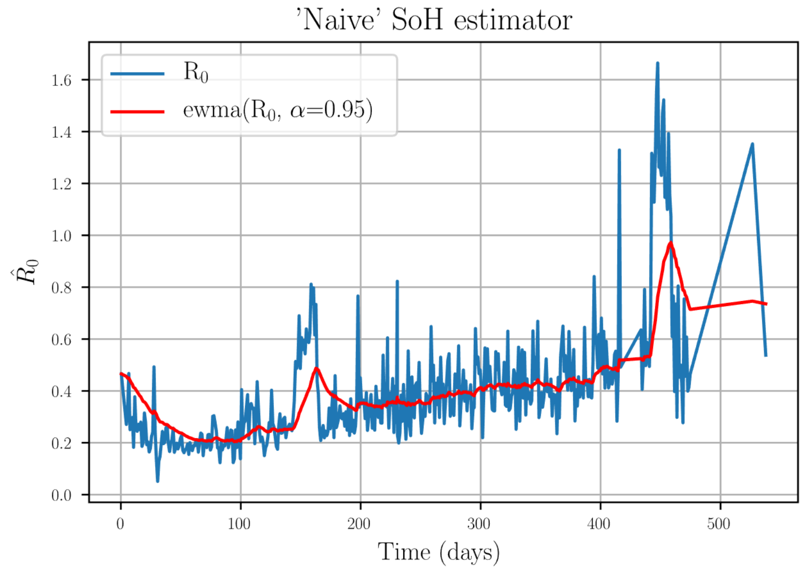
Combining data- and model-driven frameworks

Our initial attempts to estimate model parameters gave noisy results

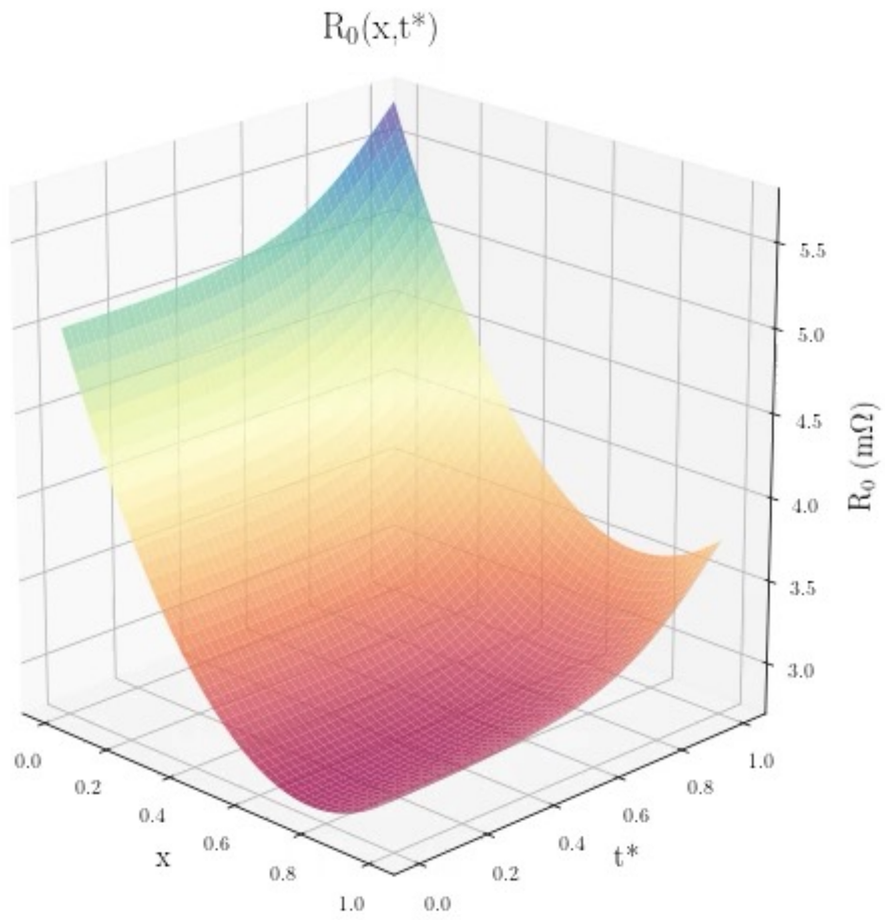
- Dealing with ECM parameter dynamics and their inherent instability

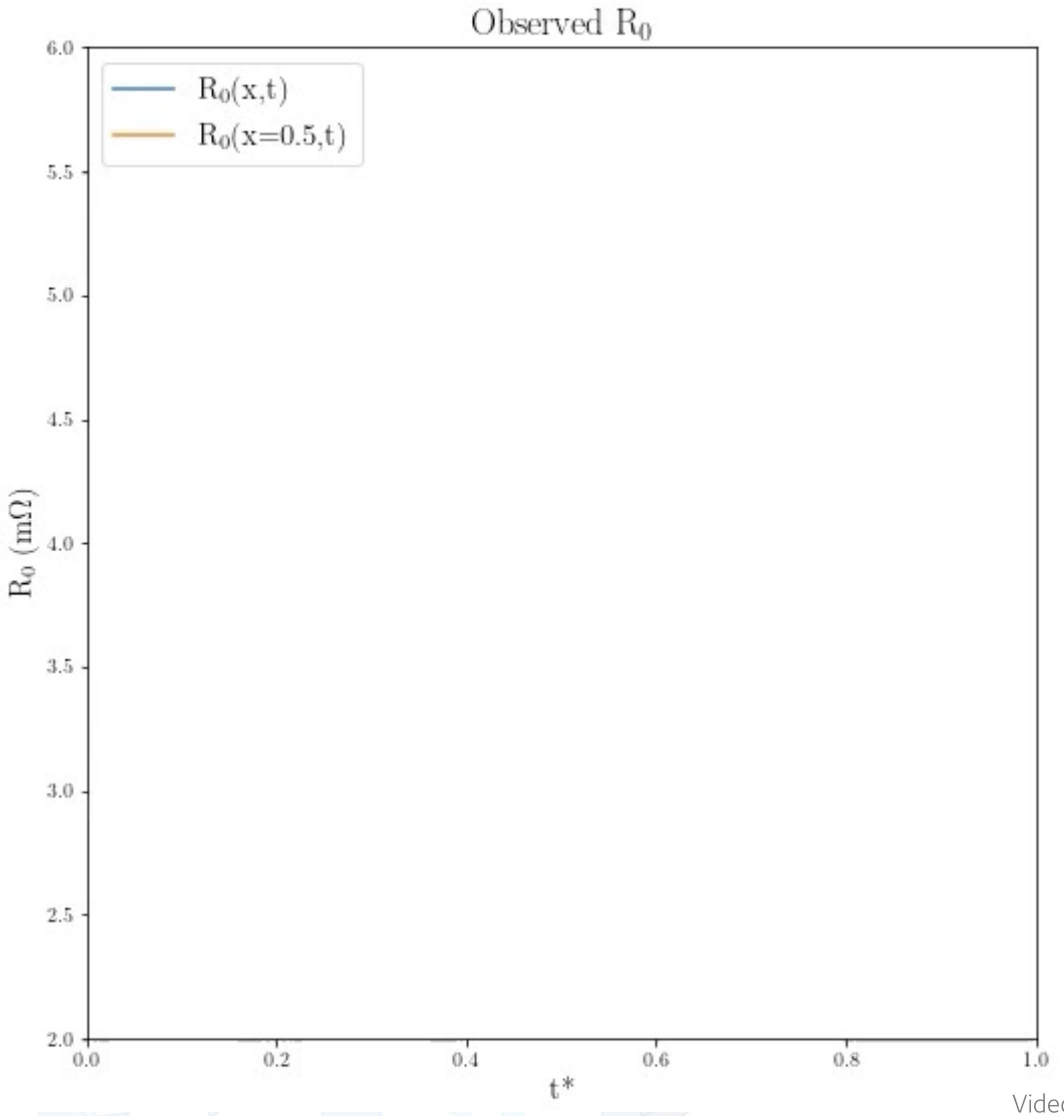






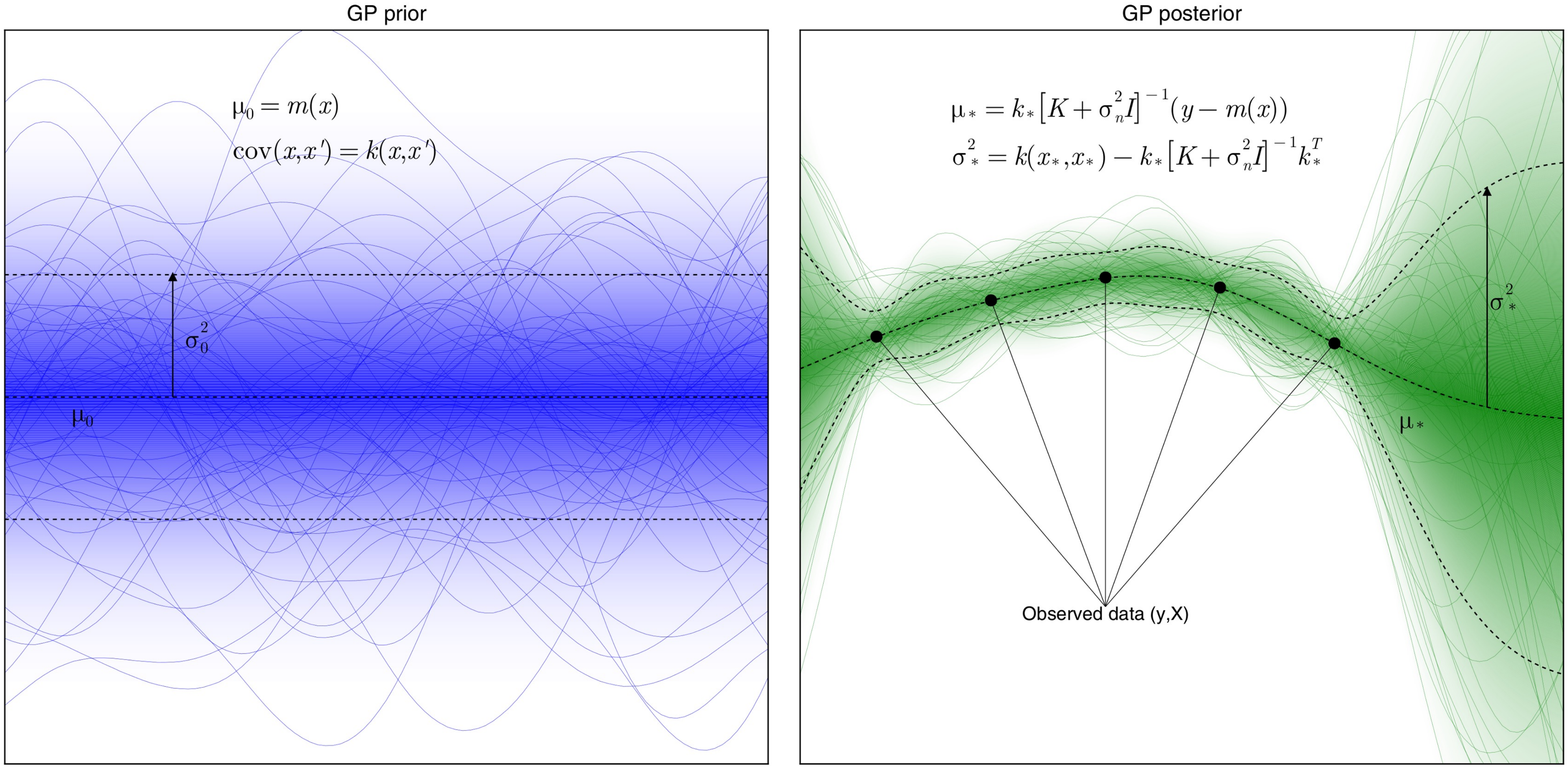
ECM Parameters, e.g. resistance are *functions* (SOC, T, I..)



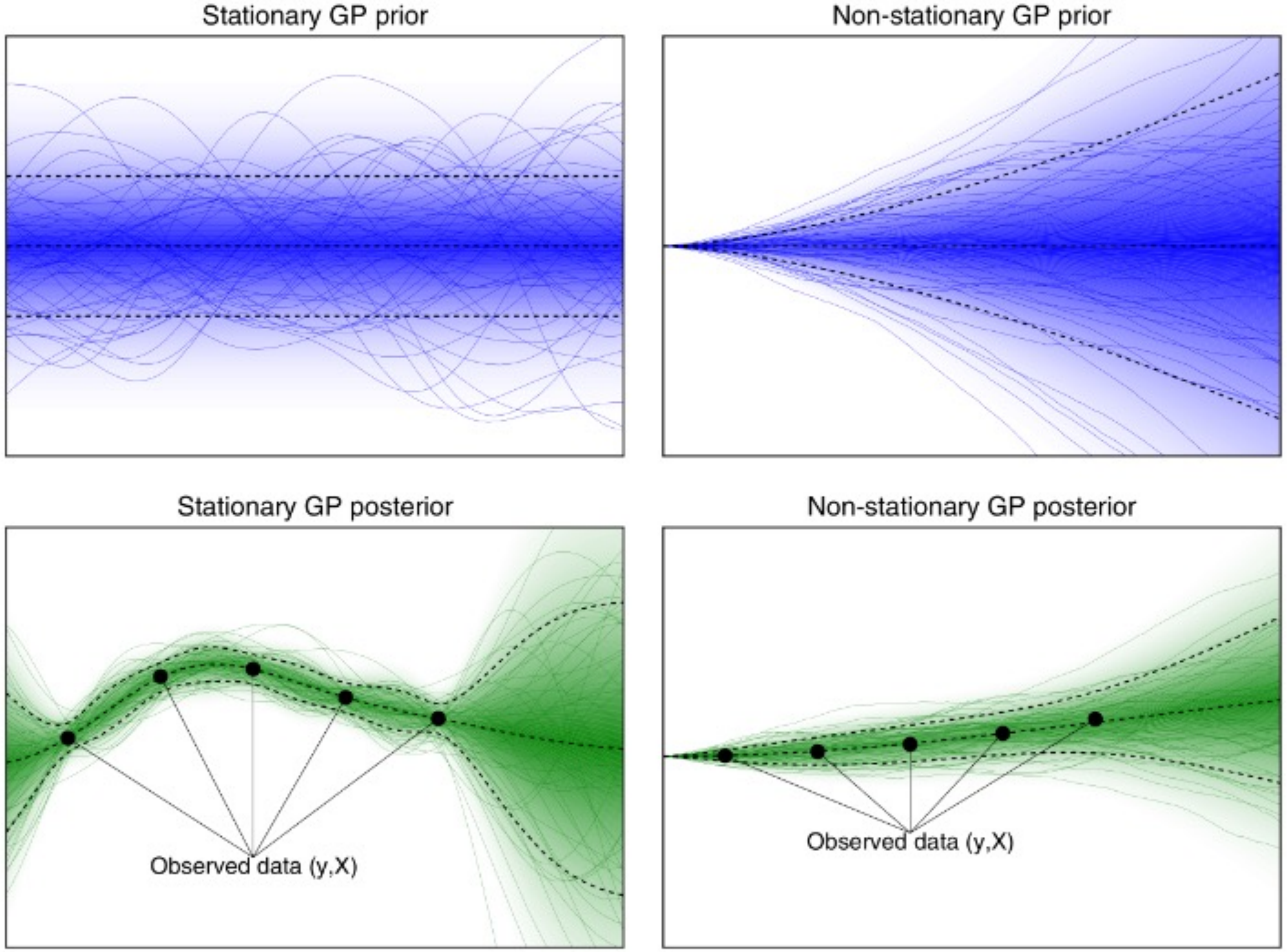


Video: Antti Aitio

GPs are a principled, flexible Bayesian approach for estimating functions



GPs are a principled, flexible Bayesian approach for estimating functions



However, there is a 'big-n' problem with GP regression

- Predictive distribution for new inputs $\boldsymbol{X}_*$ has a closed form solution!

$\boldsymbol{K} \in \mathbb{R}^{\boldsymbol{n} \times \boldsymbol{n}}$!!

$$\mu^* = k_*[K + \sigma_n^2 I]^{-1} (y - m(x)),$$
$$\mathbb{V}^* = k(x^*, x^*) - k_*[K + \sigma_n^2 I]^{-1} k_*^T$$

- To optimise **hyperparameters**, maximise log marginal likelihood:

$$\log p(y|X, \theta) = -y^T [K_\theta + \sigma_n^2 I]^{-1} y - \frac{1}{2} \log |K_\theta + \sigma_n^2 I| - \frac{n}{2} \log 2\pi$$

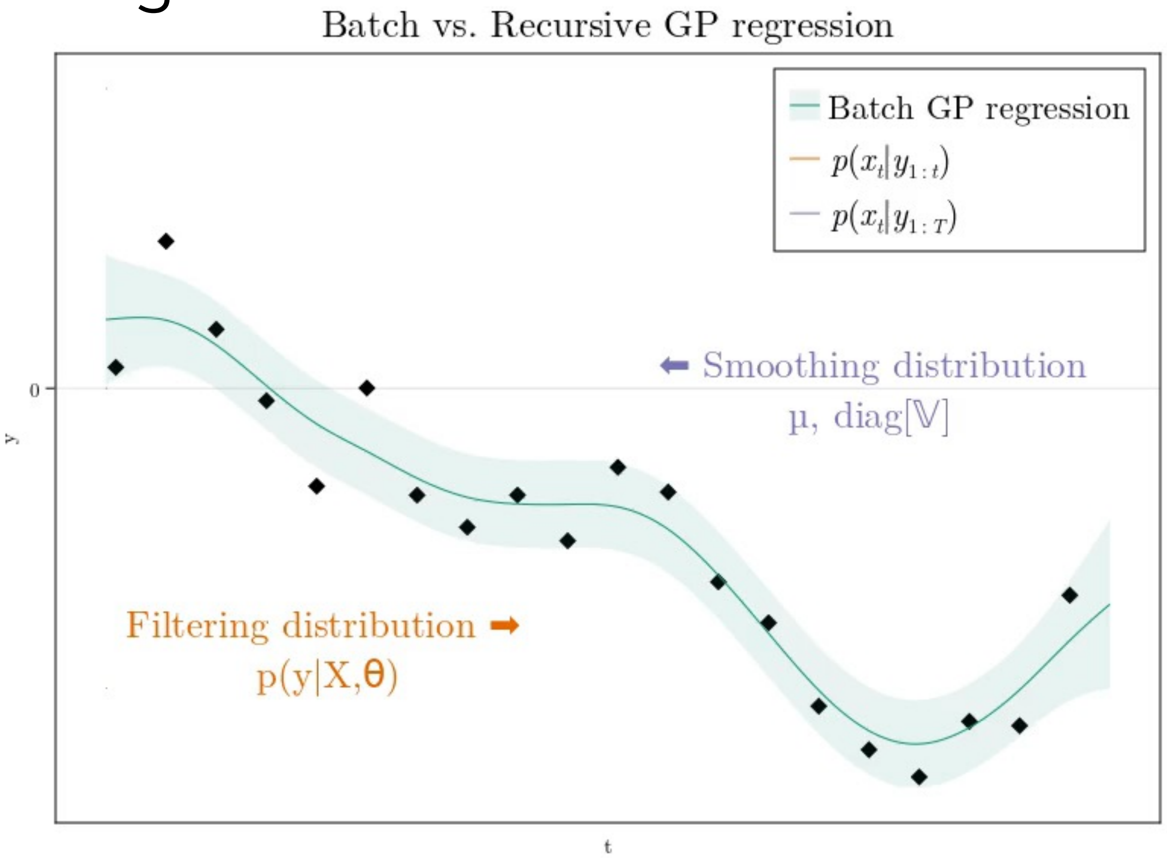
A recursive approach can tackle upscaling

- **A Gaussian process is a solution to a stochastic differential equation** (Särkkä, Solin et al. 2013)

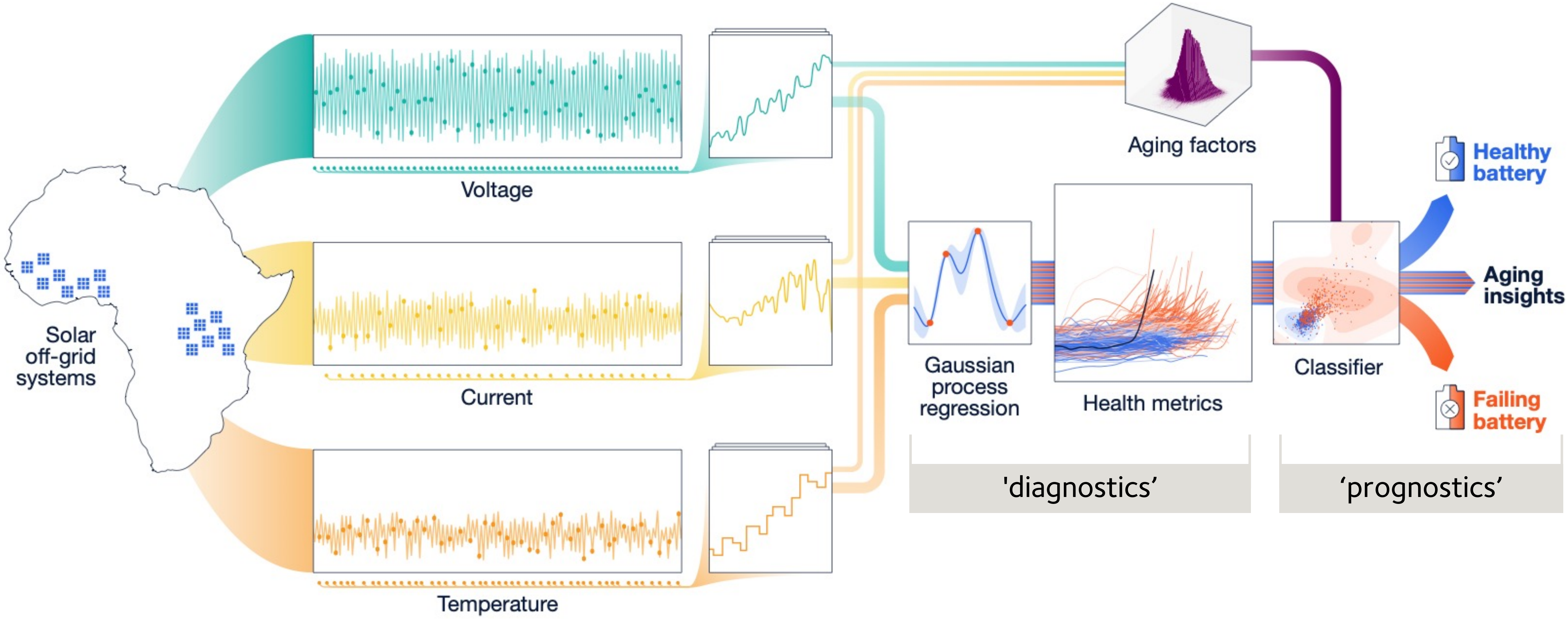
$$f(\mathbf{x},t) \sim \mathcal{GP}\left(m(\mathbf{x},t), k(\mathbf{x},t,\mathbf{x}',t')\right) \qquad \frac{\partial f(\mathbf{x},t)}{\partial t} = \mathcal{F} f(\mathbf{x},t) + L w(\mathbf{x},t)$$

$$\mathbf{y} = f(\mathbf{x},t) + \varepsilon \qquad \Longleftrightarrow \qquad \mathbf{y}_t = \mathcal{H}(\mathbf{x}_t) + \varepsilon_t$$

- The solution to a linear SDE with Gaussian noise is found using the **Kalman Filter & RTS smoother**



Overall pipeline for battery health from field data



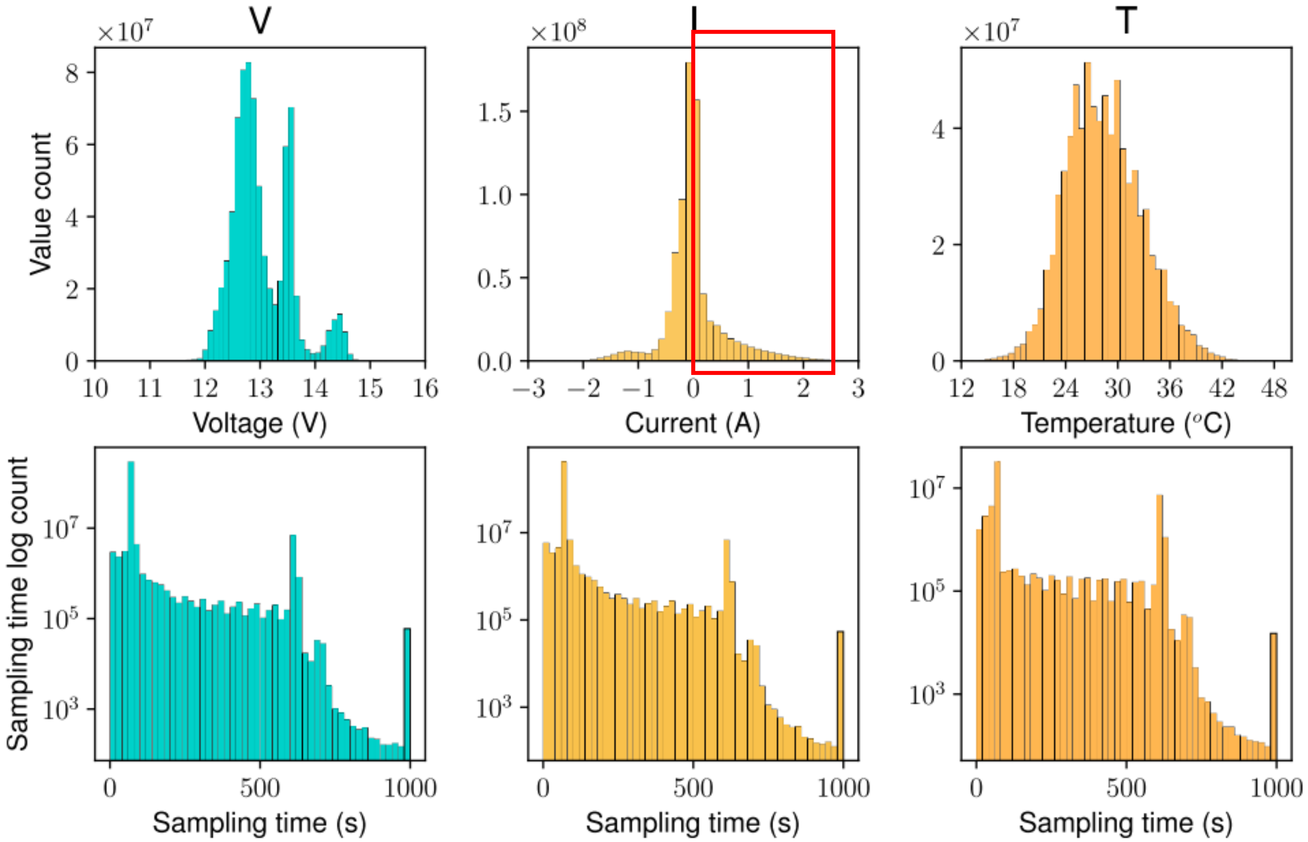
Aitio and Howey (2021)

Model the battery series resistance with a Gaussian process

- Battery model:

$$V_t = V_0(z_t) + R_0(\zeta_t, u_t, z_t, T_t)u_t + \varepsilon \ , \ \varepsilon \sim \mathcal{N}(0, \sigma^2_{n,t})$$

$$R_0 \sim \mathcal{GP}\left(0, k(\mathbf{x}, \mathbf{x}')\right) \ , \ \mathbf{x} = [\zeta \ u \ z \ T]$$



Aitio and Howey (2021)

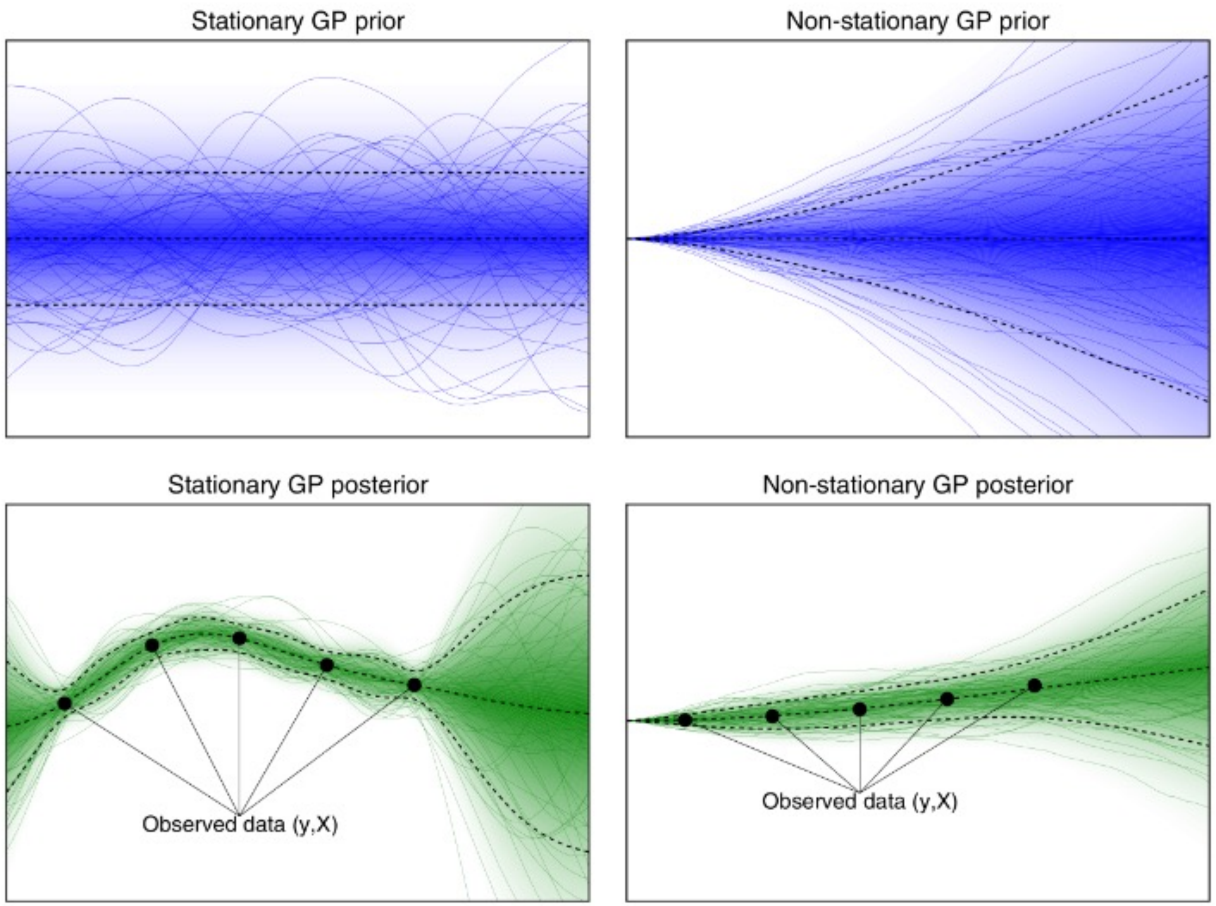
Model the battery series resistance with a Gaussian process

- GP Kernel functions:

$$k_{\mathbf{x}}(\mathbf{x},\mathbf{x}') = k_{\zeta}(\zeta,\zeta') + k_{\mathbf{x}_{\mathrm{OP}}}(\mathbf{x}_{\mathrm{OP}},\mathbf{x}'_{\mathrm{OP}})$$

$$k_{\zeta}(\zeta,\zeta') = \sigma_{\mathrm{f},0}^2\left(\frac{\min^3(\zeta,\zeta')}{3} + |\zeta-\zeta'|\frac{\min^2(\zeta,\zeta')}{2}\right)$$

$$k_{\mathbf{x}_{\mathrm{OP}}}(\mathbf{x}_{\mathrm{OP}},\mathbf{x}'_{\mathrm{OP}}) = \sigma_{\mathrm{f},1}^2 \exp\left(-\frac{1}{2}(\mathbf{x}_{\mathrm{OP}}-\mathbf{x}'_{\mathrm{OP}})\Sigma^{-1}(\mathbf{x}_{\mathrm{OP}}-\mathbf{x}'_{\mathrm{OP}})^T\right)$$



Fitting the Gaussian process to data

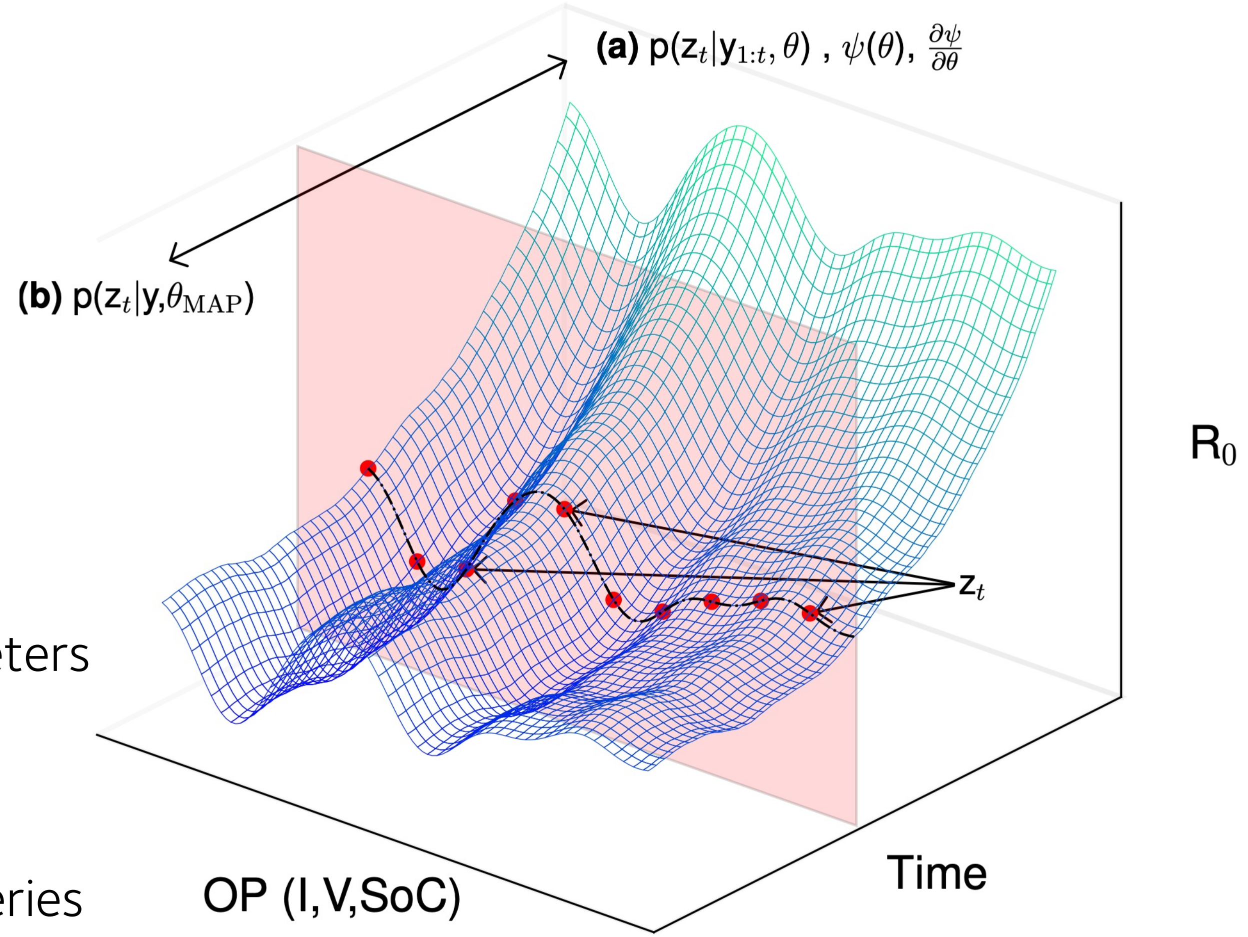
- Problem ends up looking rather familiar!

'battery dynamics'

$$\begin{bmatrix} z_{\mathrm{Batt}} \\ z_{\mathrm{GP}} \end{bmatrix}_{\mathrm{t+1}} = \underbrace{\begin{bmatrix} A_{\mathrm{Batt}} & 0 \\ 0 & \mathrm{expm}(Ft_{\mathrm{s}}^{*}) \end{bmatrix}}_{A_{\mathrm{sys}}} \begin{bmatrix} z_{\mathrm{Batt}} \\ z_{\mathrm{GP}} \end{bmatrix}_{\mathrm{t}} + BI_{\mathrm{t}} + \varepsilon,$$

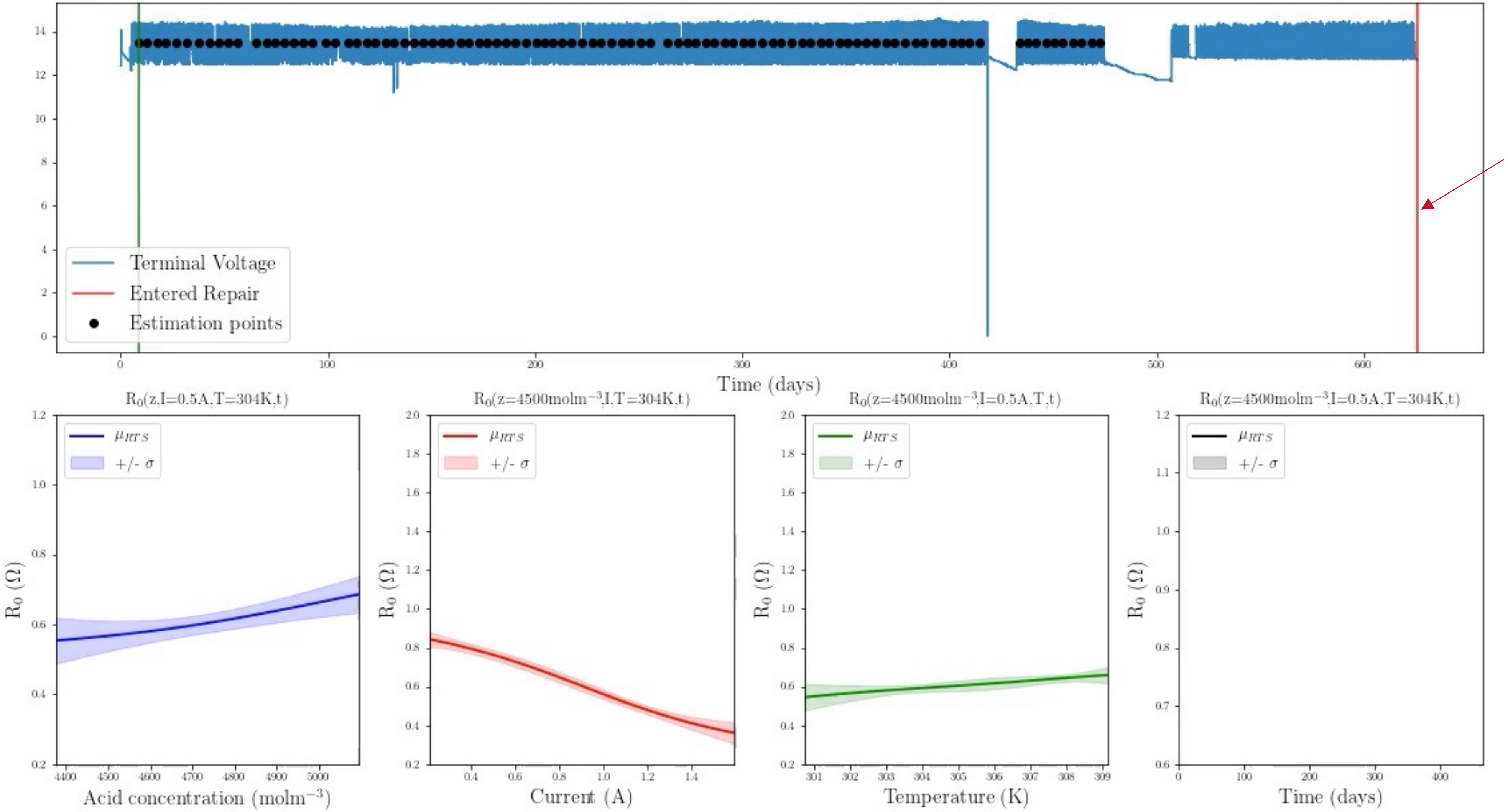
'parameter dynamics'

- Discretize 'non-time' dimensions **x** using k-means
- Estimate *maximum a posteriori* (**MAP**) hyperparameters (length scales, magnitudes) using forward pass to calculate "energy function"
- Use backward pass to get GP posteriors for all batteries



Figures reprinted from Aitio and Howey, Joule 5(12):3204-3220, 2021

From field data, learn the dependence of R_S on SOC, *T*, *I*, *t*

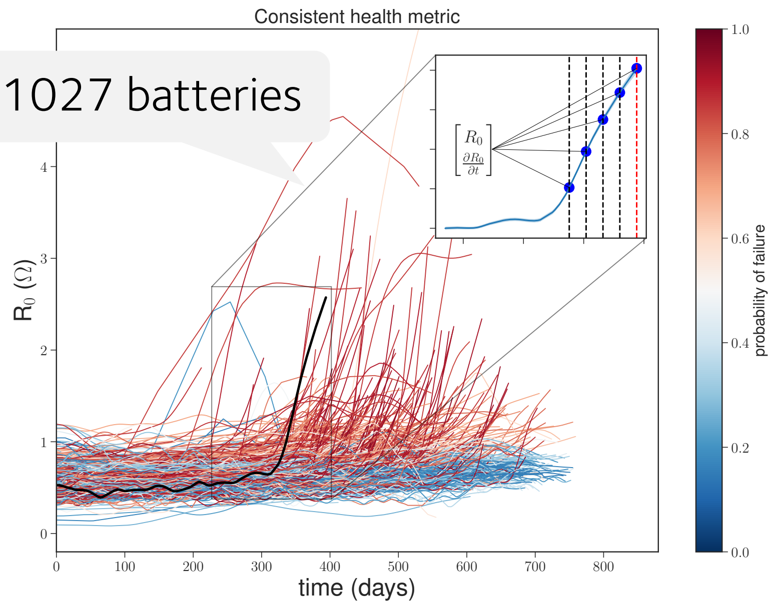


unit sent for repair

Voltage data: BBOXX; Video: Antti Aitio

To predict failure, train a classifier with independent validation data

1027 batteries



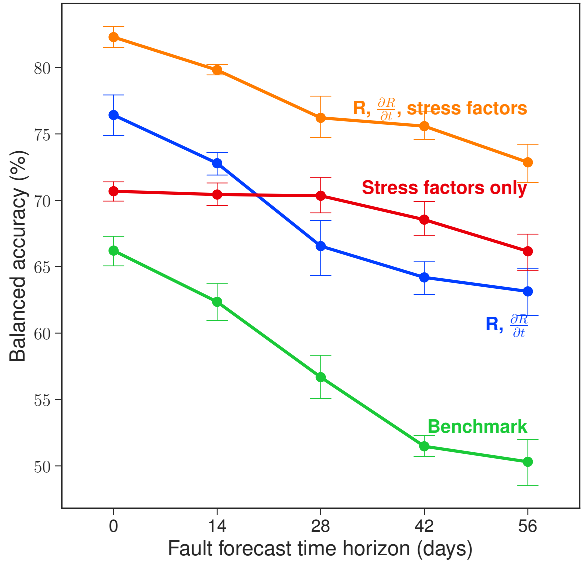
Stress factors, i.e. cumulative:

- Age
- Charge throughput
- Cycles
- Mean temperature
- Mean voltage
- …

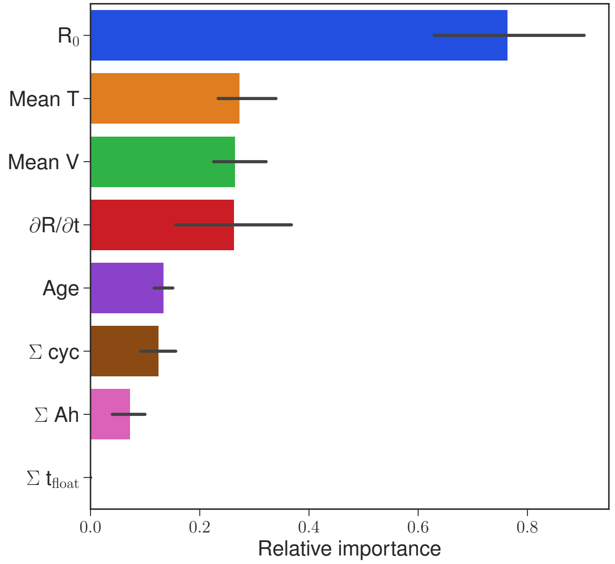
Classifier

fault predictor:

$$\text{Balanced accuracy} = \frac{1}{2}(\text{Sensitivity+ Specifity}) = \frac{1}{2}\left(\frac{\text{TP}}{\text{TP+FN}} + \frac{\text{TN}}{\text{TN+FP}}\right)$$



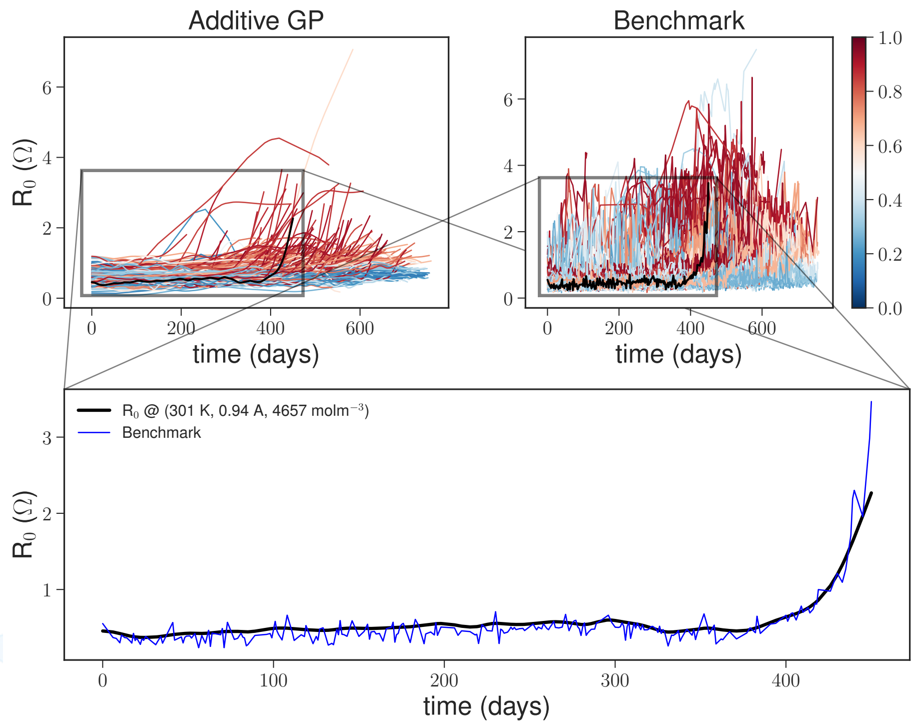
aging model:



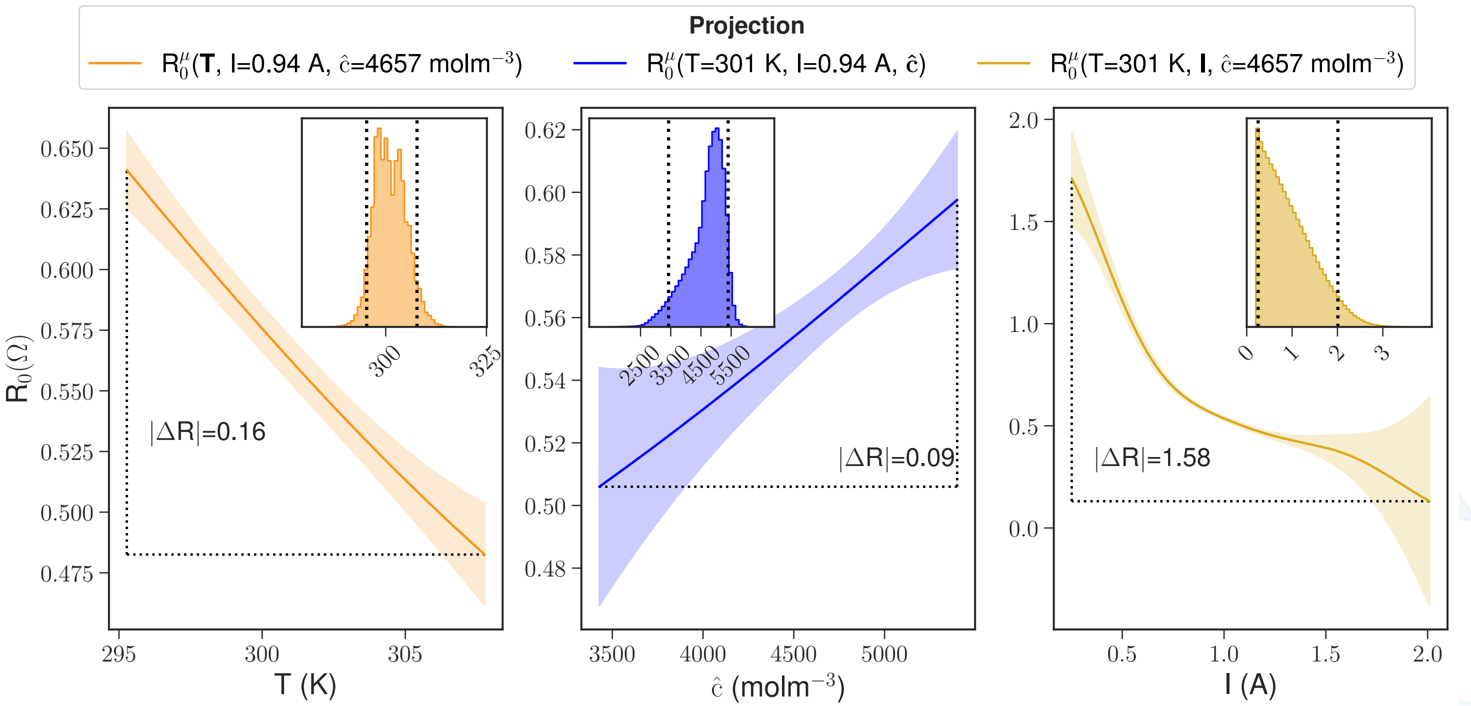
Figures reprinted from Aitio and Howey, Joule 5(12):3204-3220, 2021

Population-wide results show the importance of calibration

Benchmark results are not smooth



This is due to widely varying field conditions



Figures reprinted from Aitio and Howey, Joule 5(12):3204-3220, 2021

Summary

- Data- and model- driven SOH estimation methods have strengths and weaknesses – we should take aspects from both paradigms to get the best of both worlds

Model-driven

- Insight
- Interpretability
- Reliability of extrapolation

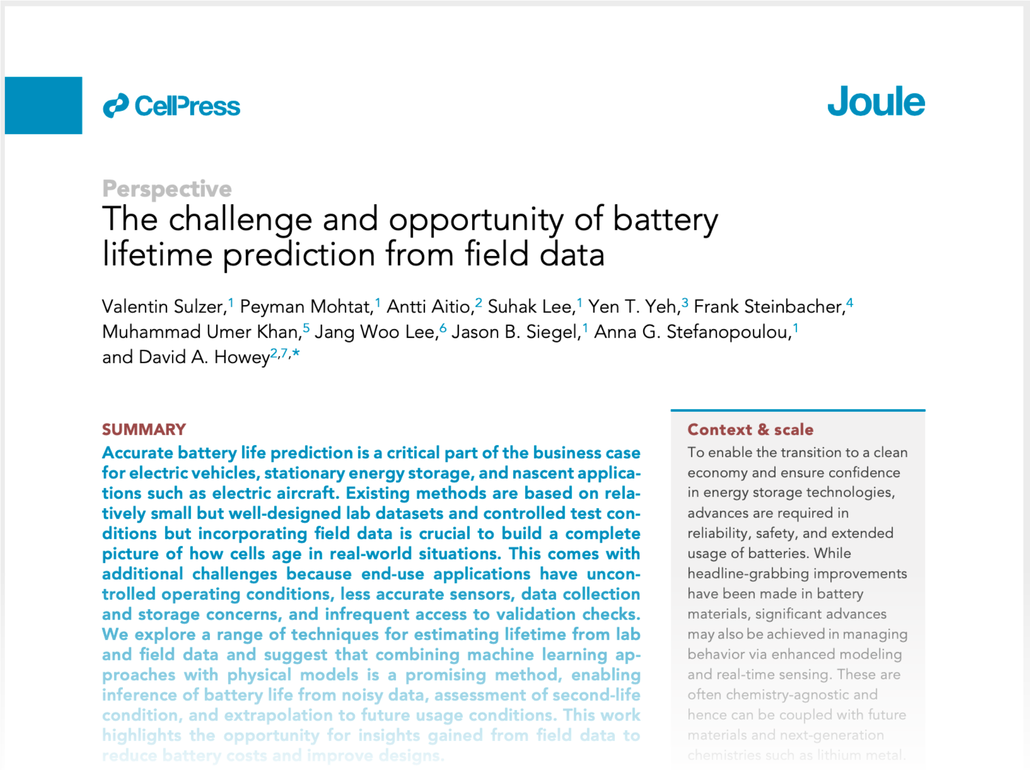
The 'Grey box' approach

Data-driven

- Flexibility
- Fewer battery parameters required

- Simple empirical models extended with **Bayesian** ML methods offer robustness when using field data
- We can extend the concept!
  - fleet-level data offer a **prior mean** for new cells?
  - battery parameters **co-evolve** – partial charge / discharge data to estimate capacity?

A quick advert for a couple of papers…



CellPress Joule

Perspective

The challenge and opportunity of battery lifetime prediction from field data

Valentin Sulzer,[1] Peyman Mohtat,[1] Antti Aitio,[2] Suhak Lee,[1] Yen T. Yeh,[3] Frank Steinbacher,[4] Muhammad Umer Khan,[5] Jang Woo Lee,[6] Jason B. Siegel,[1] Anna G. Stefanopoulou,[1] and David A. Howey[2,7,*]

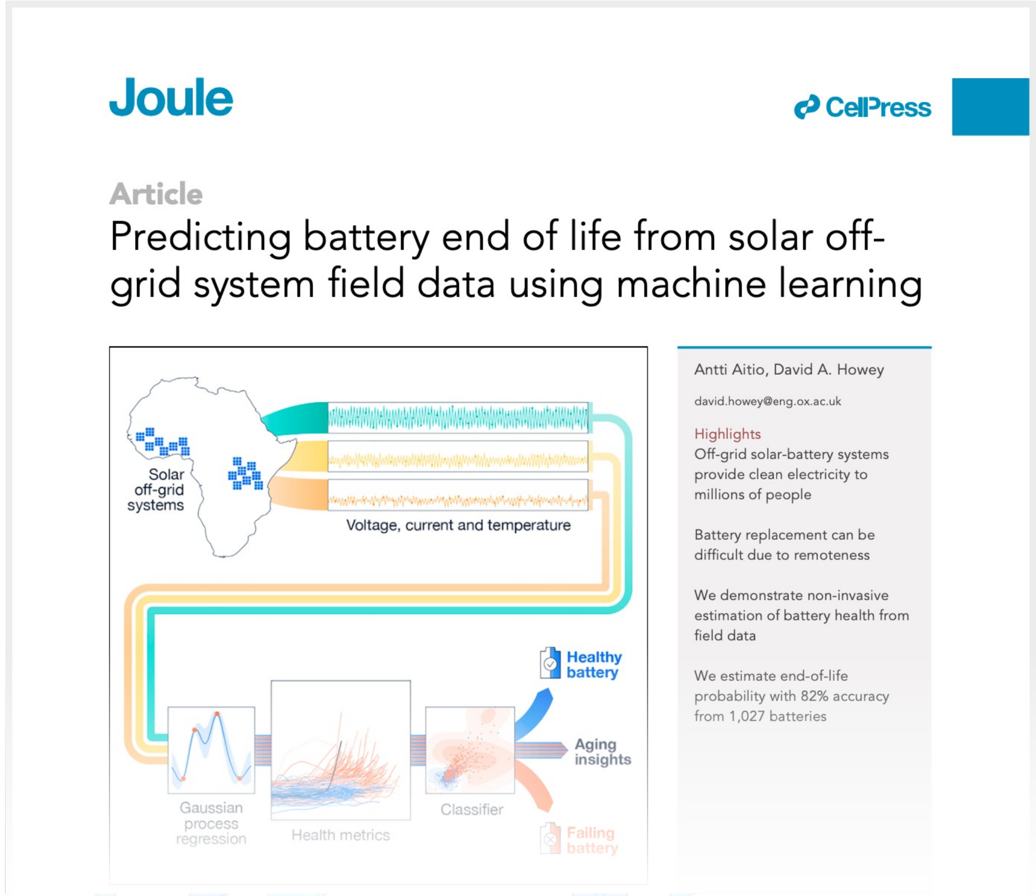
SUMMARY

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.

Context & 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.

Sulzer, Mohtat et al. Joule 5(8):1934-1955, 2021



Joule CellPress

Article

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

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

Aitio & Howey, Joule 5(12):3204-3220, 2021

Battery Intelligence Lab

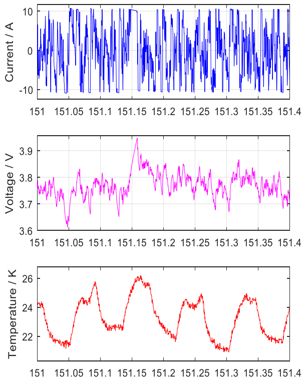
University of Oxford

Outlook: Data availability, model parsimony and scaleup are open issues

1. Data-driven approaches are **only as good as the available data!**

2. There should be more focus on the need for **simple but interpretable** models

3. Scaleup is a difficult, and many companies are still not convinced about the **financial benefits** of these approaches



Life: 3-5 years
Data rows: 1 million

Bboxx THE FARADAY INSTITUTION MULTI-SCALE MODELLING Modelling

david.howey@eng.ox.ac.uk
http://howey.eng.ox.ac.uk
@davidhowey

Pictures: Sam Greenbank, BBOXX