



# Bat492: Machine Learning for Accelerated Life Prediction and Cell Design 2021

June 2021

*Changing the World's Energy Future*

Eric J Dufek



*INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC*

#### **DISCLAIMER**

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

# **Bat492: Machine Learning for Accelerated Life Prediction and Cell Design 2021**

**Eric J Dufek**

**June 2021**

**Idaho National Laboratory  
Idaho Falls, Idaho 83415**

**<http://www.inl.gov>**

**Prepared for the  
U.S. Department of Energy  
Under DOE Idaho Operations Office  
Contract DE-AC07-05ID14517**

# Bat492: Machine Learning for Accelerated Life Prediction and Cell Design

Eric Dufek

Idaho National Laboratory

June 21-25, 2021



This presentation does not contain any proprietary, confidential, or otherwise restricted information  
INL-CON-21-62775



# Overview

## Timeline

- Start: October 1, 2020
- End: September 30, 2022
- Percent Complete: 35%

## Budget

- Funding for FY21 – \$1.2M

## Barriers

- Time needed to predict life and understand failure modes
- Lack of tools and methods which readily cascade across programs
- Distinct need to link physics to enhance the technology development process

## Partners

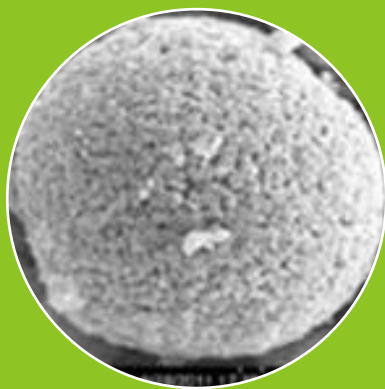
- Idaho National Laboratory
- National Renewable Energy Laboratory
- Close collaboration with Behind-the-meter-storage (bat442), and Extreme Fast Charge and Cell Evaluation of Lithium-ion Batteries (XCEL, bat 456-463 project)

# Relevance

*Objective: Accelerate transformative advancement by creating a robust, common framework*

## Develop methods and core tools to:

- Reduced time to validate new materials, designs, manufacturing processes and use cases
- Access to large amounts of data to enable discovery and deployment
- Provide breadth spanning transportation and stationary storage to support electrified mobility
- Benefit across the storage ecosystem (research to industry and consumers)



Materials  
Development,  
Understanding,  
and  
Manufacturing



Cell design,  
Validation and  
Manufacturing



System  
Integration and  
Deployment

**Common Tools and Data Storage**

## Task milestones

| Milestone  | Due      | Status     |
|--|----------|------------|
| Finalize IP structure and coordination across the team   | 12/31/20 | Complete   |
| Generate synthetic data from Graphite/NMC cells and initiate Deep Learning related to electrochemical signatures   | 3/31/21  | Complete   |
| Predict and validate electrochemical performance of aged cells for at least two different charging conditions using a combination of electrochemical and life models | 6/30/21  | In process |
| Quantify life model accuracy using automated physics-based model generation based on design and experiment duration using either LTO/LMO or graphite/NMC datasets    | 9/30/21  | In process |
| Predict and validate performance and degradation modes within 5% for known duty cycles and 10% for use cases not aligned with training sets                          | 9/30/21  | In process |

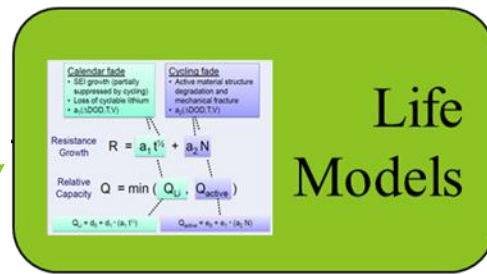


# Approach

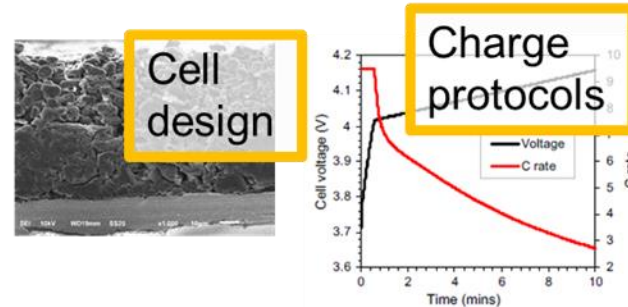
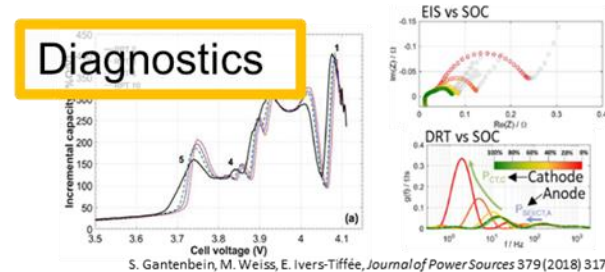
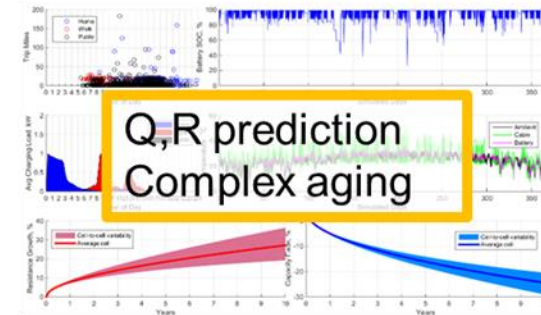
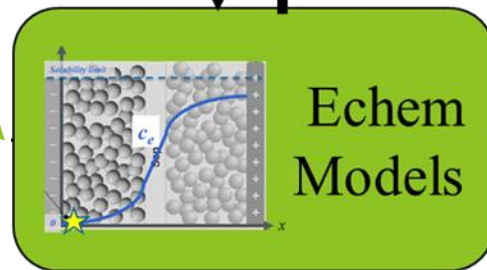
*Accelerating innovation requires  
failure mode classification,  
projection and validation*

*Combination of high-  
quality data  
generation,  
assessment and  
analysis*

Use of robust electrochemical  
analysis with targeted secondary  
characterization for validation



Mechanisms  
Synthetic data  
Observability



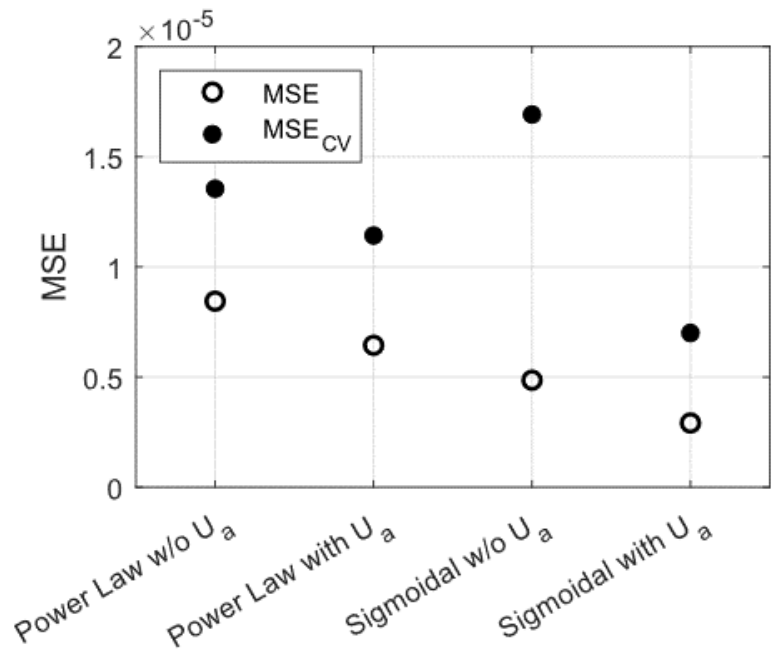
**Early Understanding of Failure Modes to Reduce Development and Deployment Cost**



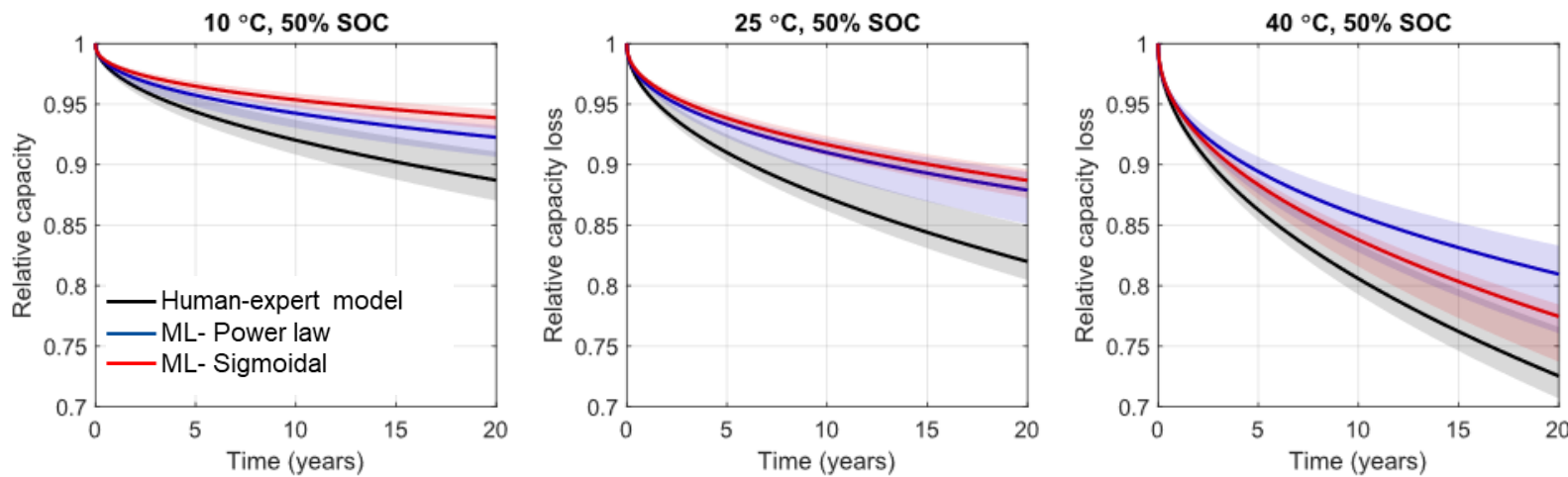


# Generation of algebraic battery life models

- Automatic identification of reduced-order degradation models
  - Bi-level optimization
  - Symbolic regression
  - Cross validation (CV)
- Up to 2x decreased uncertainty using autogenerated models when compared to human model development
- Methods include ability to perform sensitivity analysis and uncertainty quantification



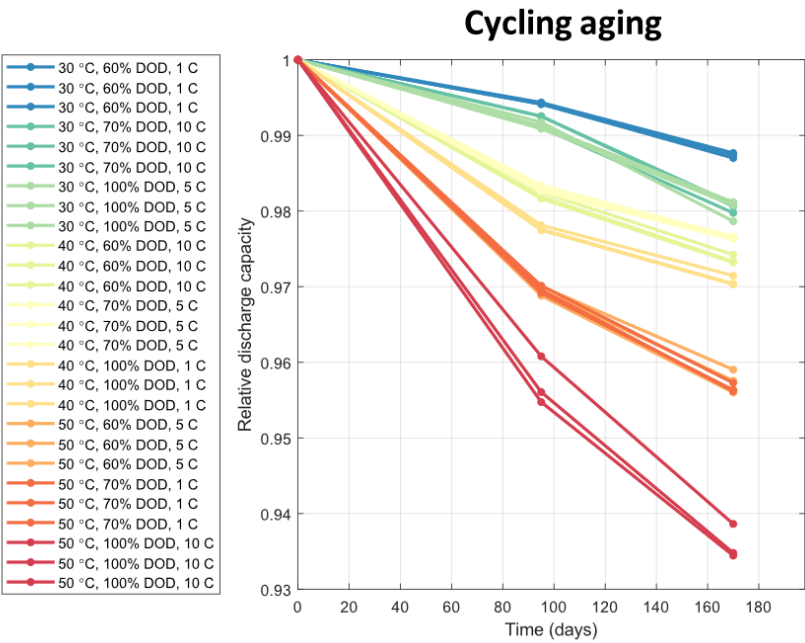
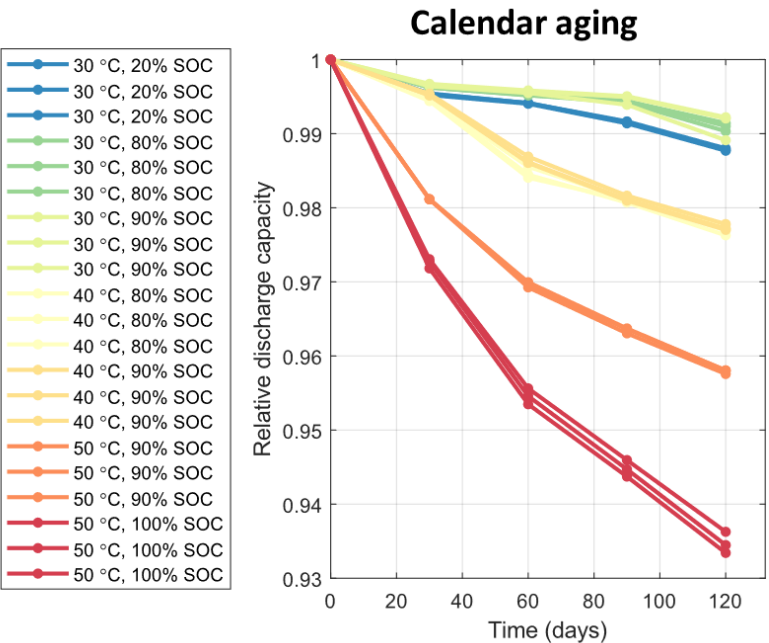
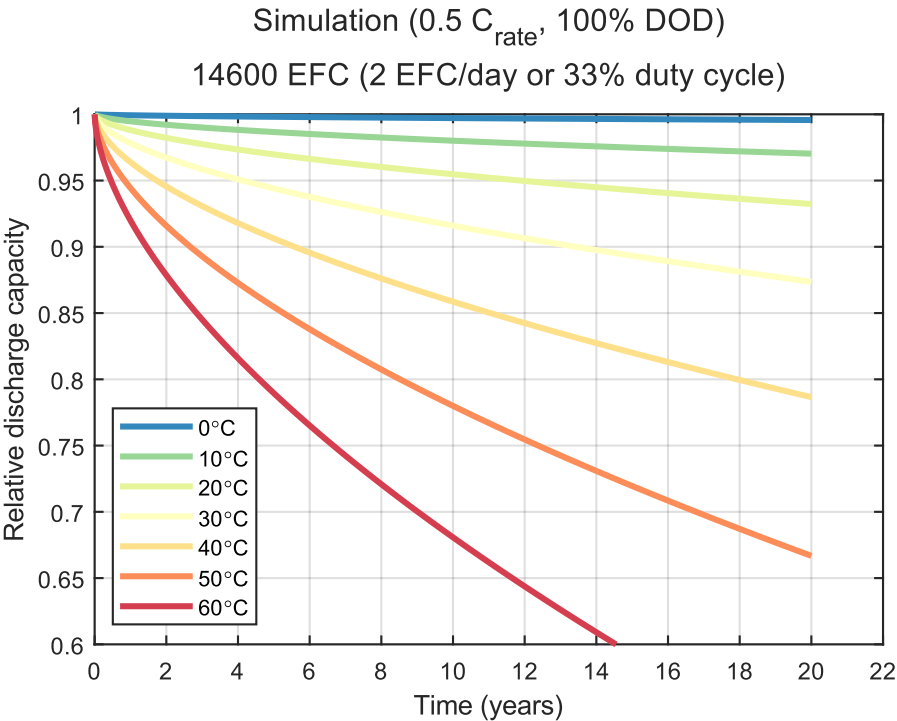
Mean-squared error for auto-generated models



Calendar-life projections w/uncertainty

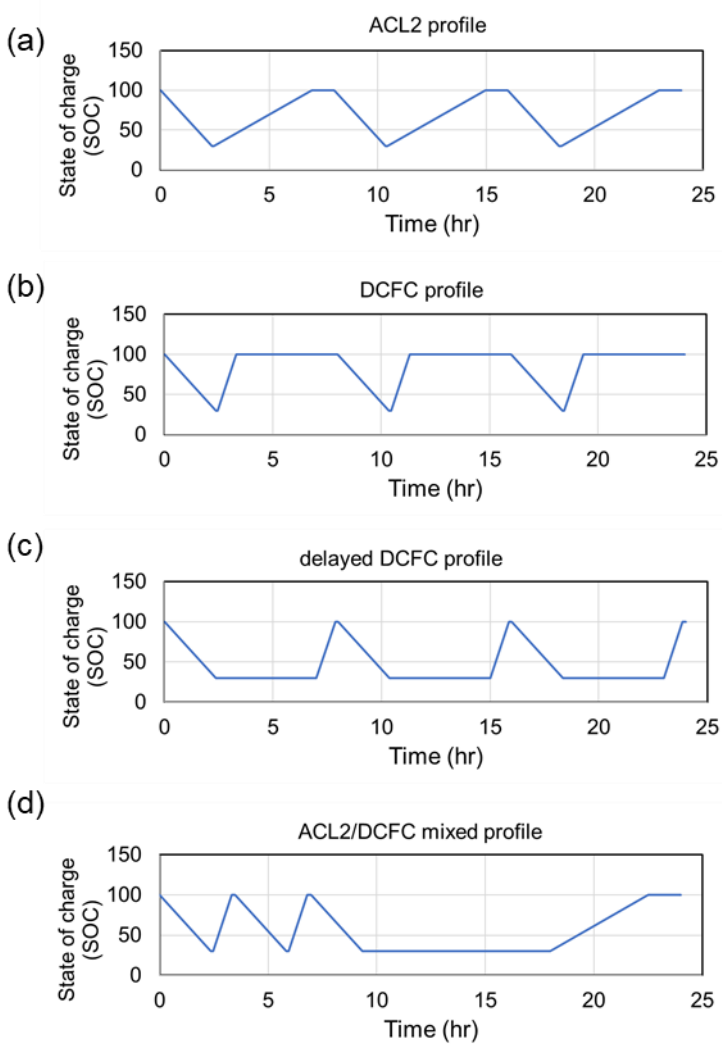
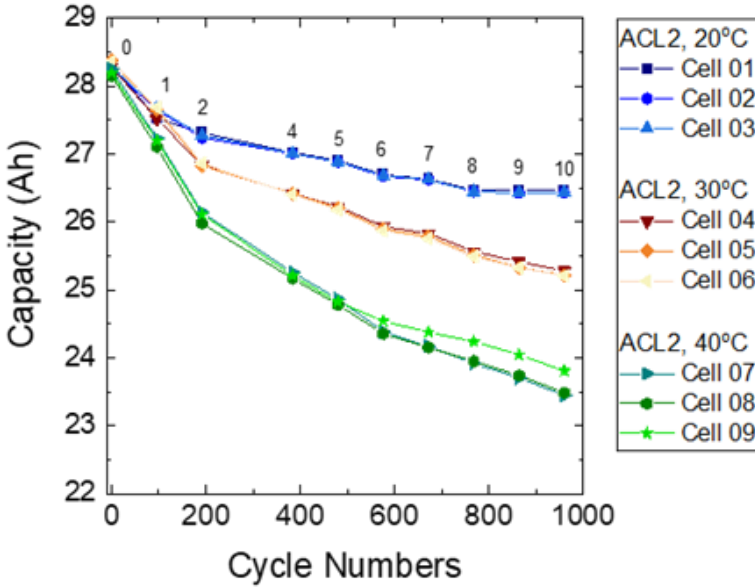
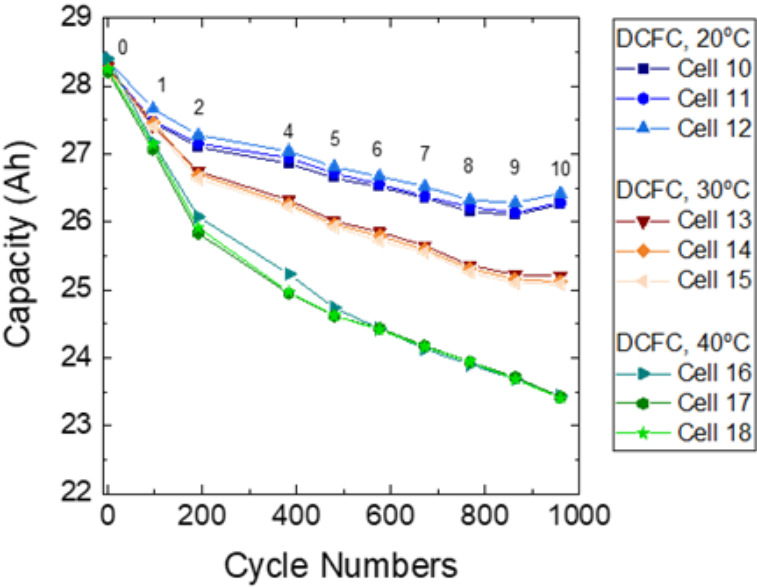
# Model extrapolation to 20 years

- Auto-generated models reduce time needed for predictions.
- Realistic Model predictions for T based on collection of 4-5 months of data.
- Full design of experiments for both calendar and cycle life using LTO/LMO cells



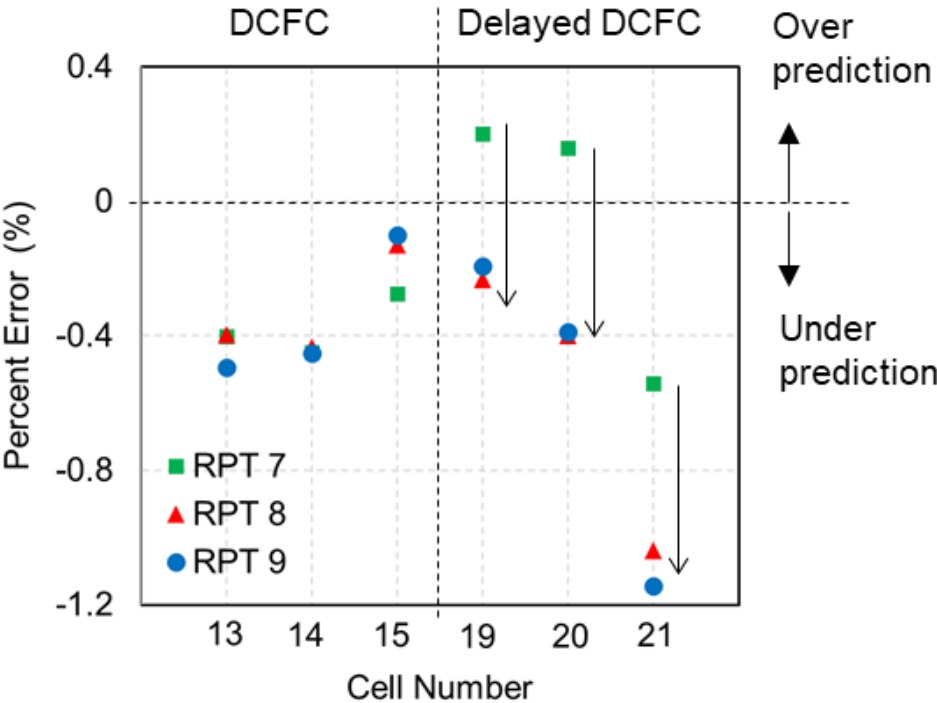
# Understand the degradation from fast charging

- Comparison of AC Level 2 and DCFC
- Mixed use profiles
- Delayed DCFC protocol reduces capacity fade up to 1.3% at RPT9 compared to no delay DCFC
- Nissan Leaf Cells (and aligned pack data)

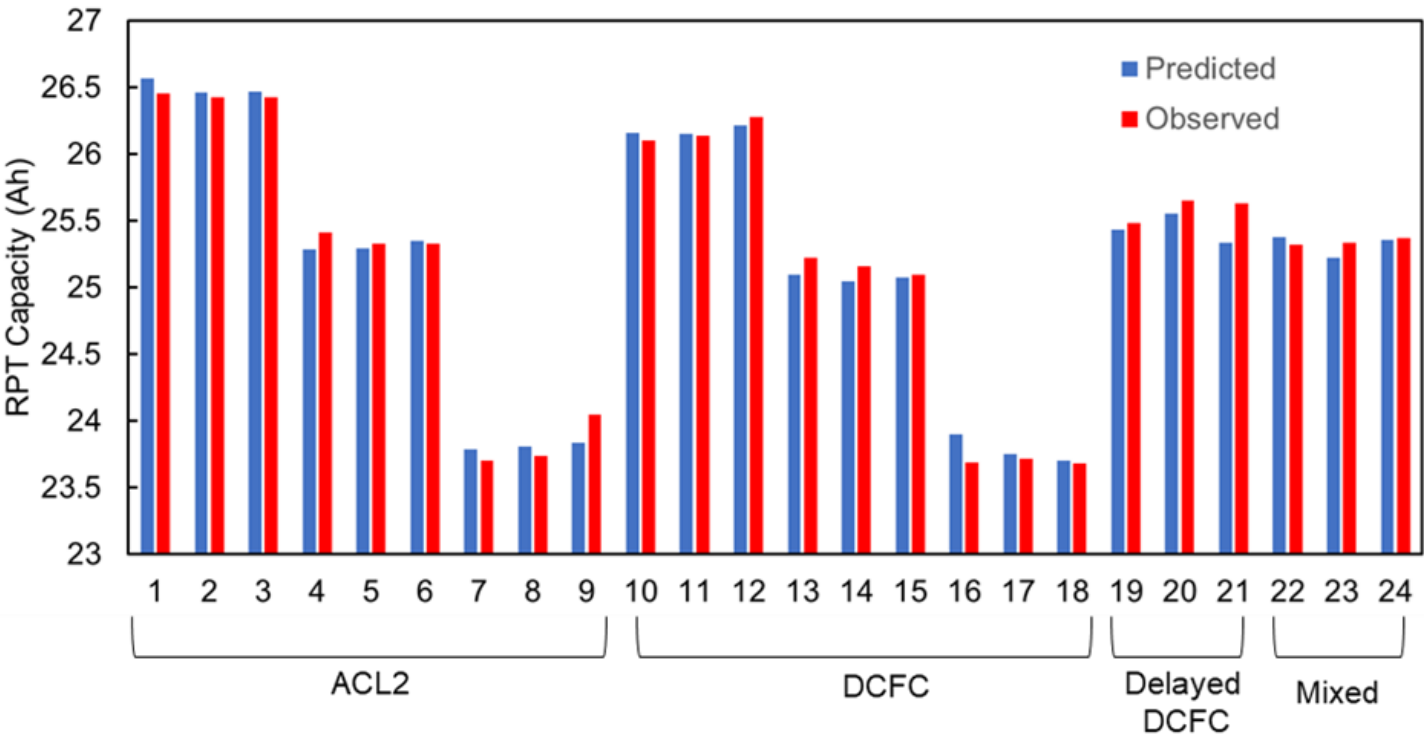


# RPT capacity projection

- Using existing data it is possible to predict capacity at 864 cycles using the first 45 DST cycles - ~5% of data acquired for mixed use conditions
- But, data-driven and even physics-based techniques will fail if calendar considerations aren't evolved
- Use of cycling only data provides low error, but overprediction of fade for mild calendar conditions



**Overprediction of fade for mild calendar conditions**

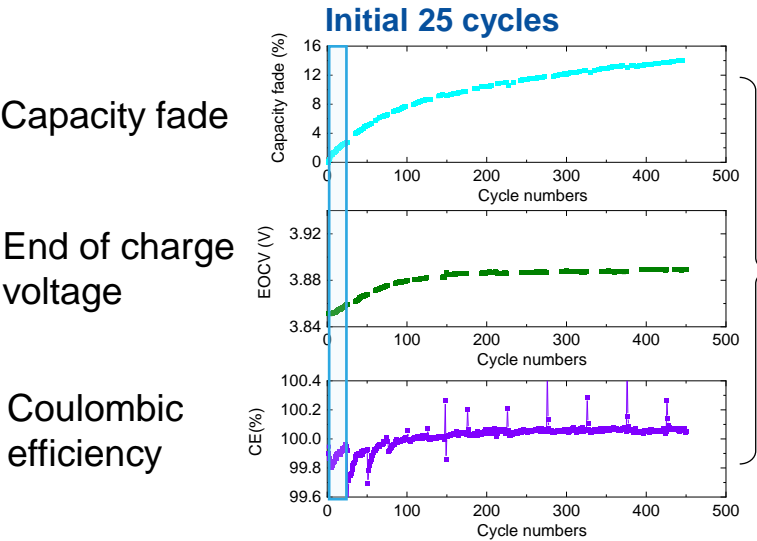


# Constructing an algorithm that separates Li-plating and SEI

Physically meaningful  
electrochemical signatures

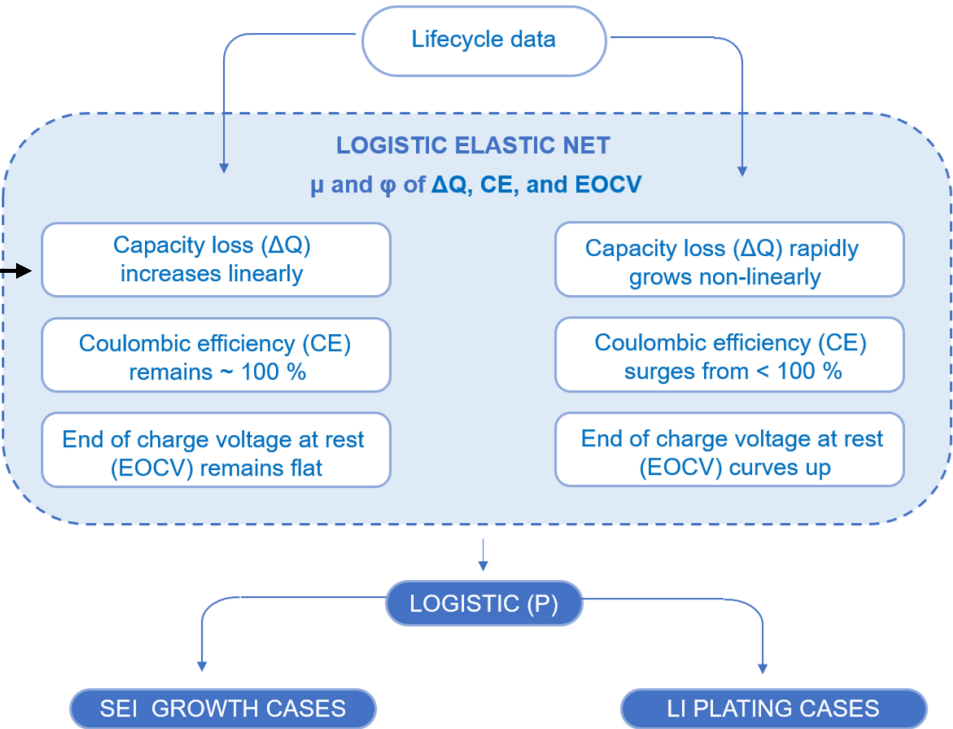
ML classification  
Logistic elastic net

Decision making procedure  
(classification of cells)



Mean ( $\mu$ )  $\rightarrow$  initial state  
Autoregressive ( $\phi$ )  $\rightarrow$  time dependence

Correlation between coherent  
EC signatures

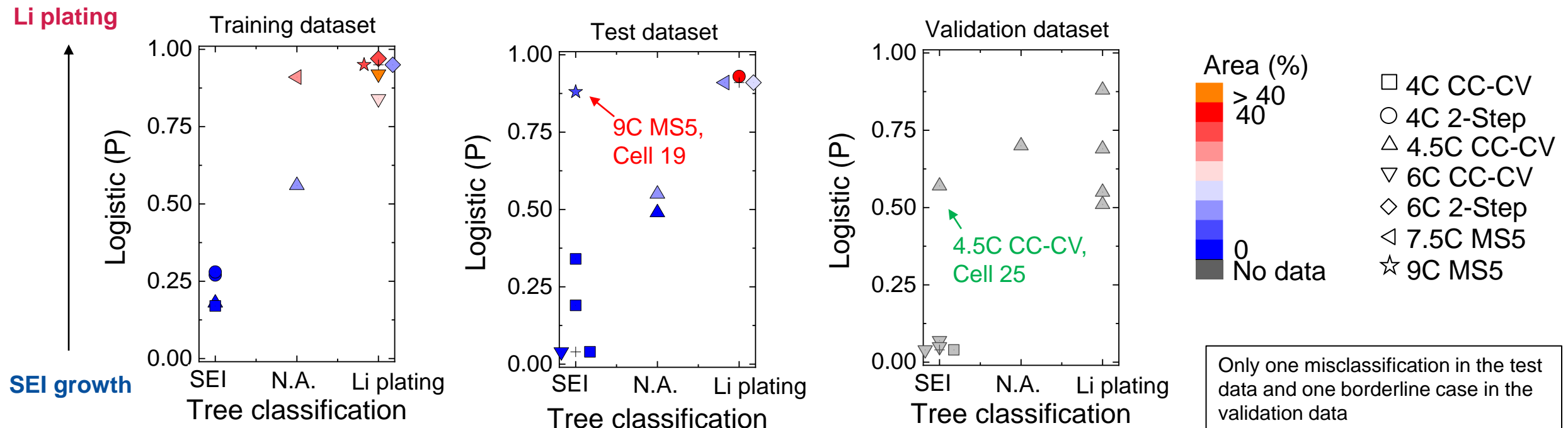


|    | Charging protocols | Cell numbers   |
|----|--------------------|----------------|
| P1 | 4C CC-CV*          | 01, 02, 03     |
|    | 4C 2-step**        | 04, 05, 06     |
|    | 4.5C CC-CV         | 07, 08, 09     |
|    | 6C CC-CV           | 10, 11, 12     |
|    | 6C 2-step          | 13, 14, 15     |
|    | 7.5C MS5***        | 16, 17         |
|    | 9C MS5             | 18, 19         |
|    |                    |                |
| P2 | 4.5C CC-CV         | 20, 21, 22     |
|    |                    | 23, 24, 25     |
| P3 | 4C CC-CV           | 26, 27         |
|    | 6C CC-CV           | 28, 29, 30, 31 |

31 NMC/graphite cells:  
Training: known condition  
Test: known condition  
Validation: unknown condition

# Classifying likelihood of Li plating

- Multiple signatures have coherent response – need to be considered jointly
- Aligning signatures into a decision framework enhances ability to readily encompass in a ML framework
- Analysis can use 25 cycles or less vs 100+ for human evaluation
- **First step toward ability to predict life and failure mode – Tailored cell engineering for cost reduction**



## Remaining challenges and barriers

- Alignment of data quantity, quality and availability
  - Not all data created equal
- Joint prediction of life and performance for both standard and non-typical use cases
  - Based on accelerated cycle and calendar aging
- Continued expansion for other chemistries
- Performance prediction using different order electrochemical models
- Joint use of experimental and synthetic data
- Expanded data needs and coordination of tools for data quality evaluation



## Proposed Future Research

- Continued expansion and inclusion of additional failure modes and prediction schemes
- Expanded synthetic data generation
- Coordinated data sharing across multiple national laboratories and other institutions
- Aligned electrochemical and life models with incorporated failure mode analysis

*Any proposed future work is subject to change based on funding levels*

# Contributors and Collaborators

Andrew Colclasure  
Bor-Rong Chen  
Eric Dufek  
Kandler Smith  
Kevin Gering  
M. Ross Kunz  
Matt Shirk  
Michael Evans  
Nalinrat Guba  
Parameswara Rao

Paul Gasper  
Peter Weddle  
Qiang Wang  
Randy Bewley  
Sangwook Kim  
Tanvir Tanim  
Yugandhar Police  
Zonggen Yi

*Collaborations: Behind-the-meter-storage (bat442)  
and XCEL (bat456-463)*

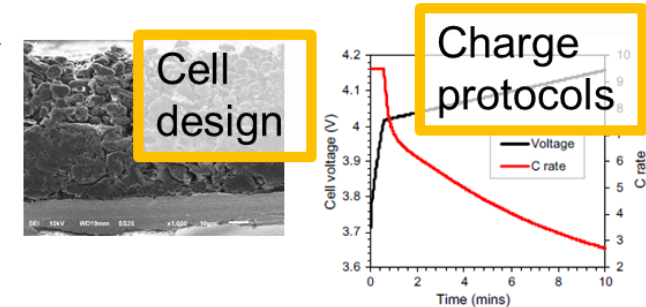
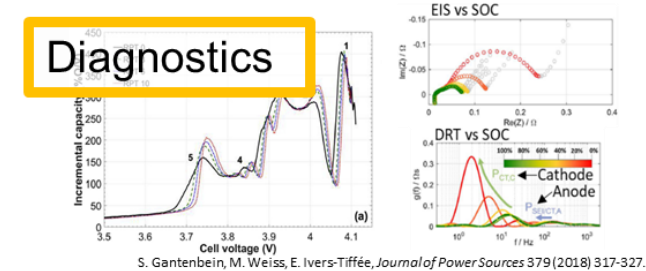
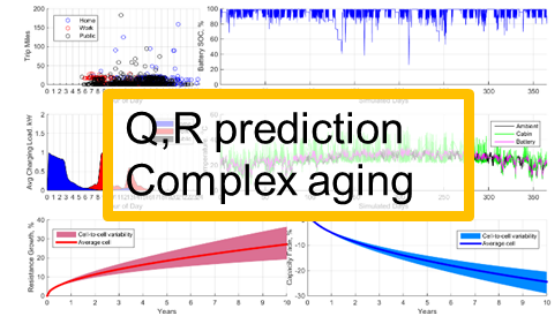
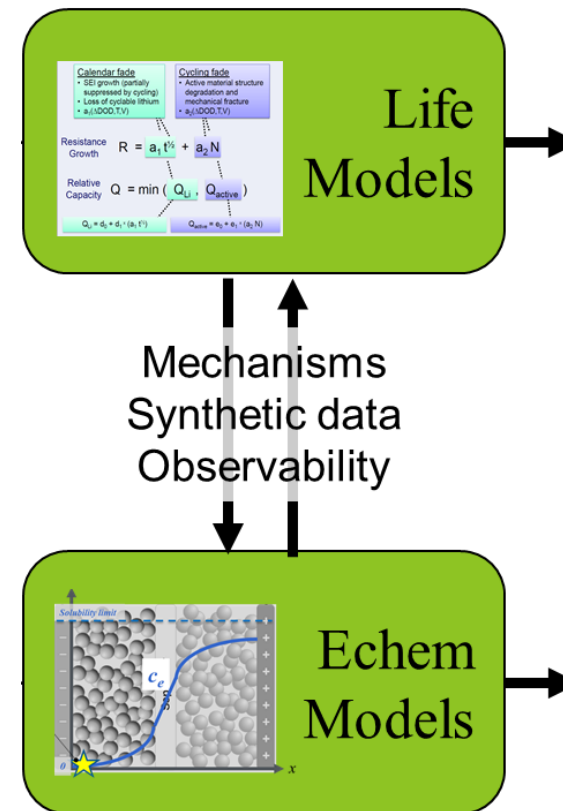
*Collaboration with:*



IDAHO NATIONAL LABORATORY

# Summary

- Autogeneration of life models reduces time for life predictions
- Early life prediction possible using 2 weeks of cycling data
- Methods can be extended to non-training data streams
- Identified EC signatures that physically correlated to SEI or Li plating
- Established an ML classification framework that classifies aging modes
- Using multiple signatures decision can be made early, within the first 25 life cycles



# Acknowledgements

We thank **Simon Thompson, Samuel Gillard, Steven Boyd and David Howell** for VTO programmatic support. **Tony Burrell and Venkat Srinivasan** for coordination on BTMS and XCEL projects

Cells were provided by Cell Analysis, Modeling, and Prototyping (CAMP) facility at Argonne National Laboratory.





## Reviewer Only Slides

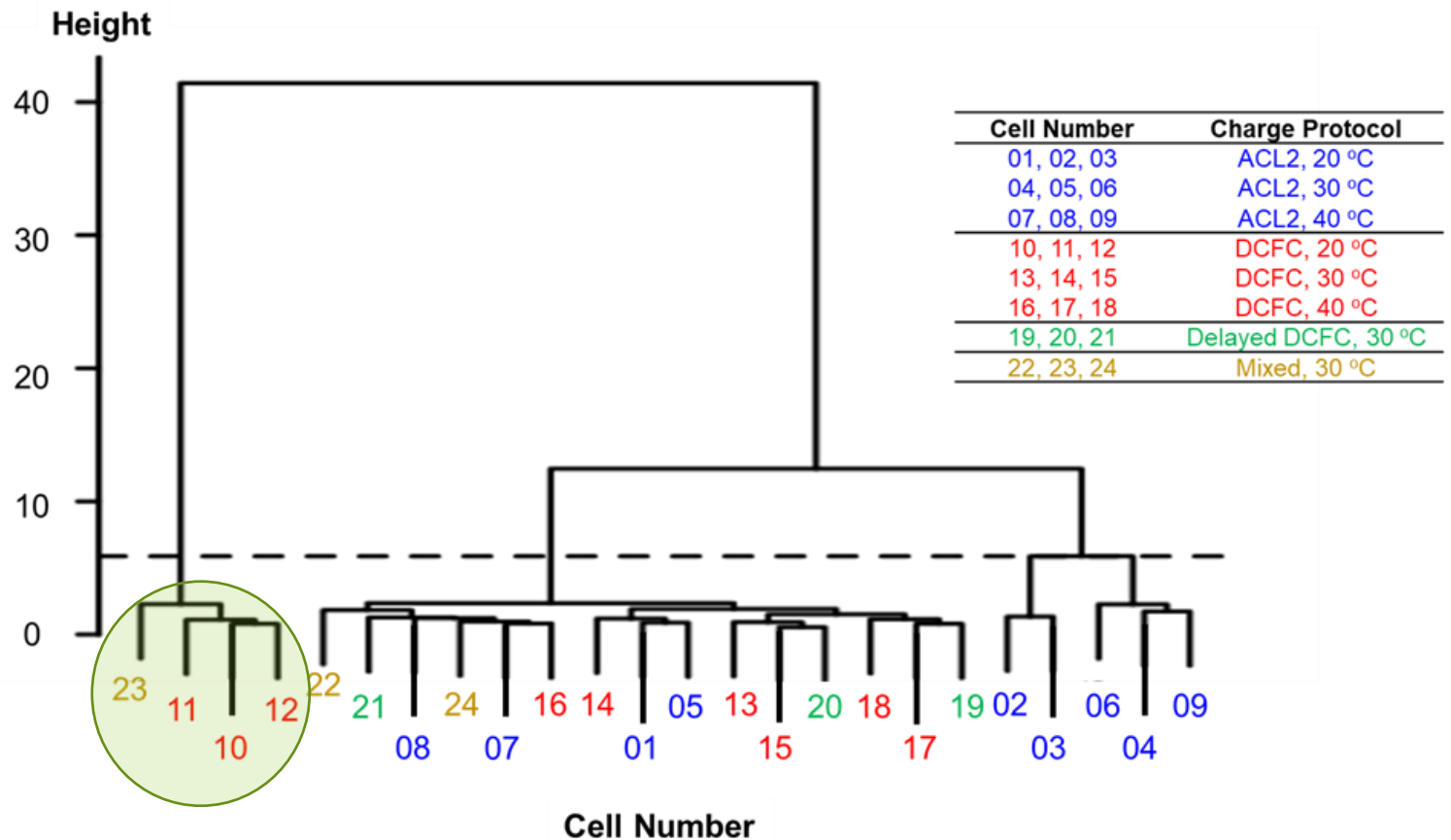


## Publications

- B.R. Chen, M.R. Kunz, T.R. Tanim, E.J. Dufek “A machine learning framework for early detection of lithium plating combining multiple physics-based electrochemical signatures” *Cell Reports Physical Science*, 2(3) 100352 (2021) doi.org/10.1016/j.xcrp.2021.100352
- P. Gasper, K. Gering, E. Dufek, K. Smith “Challenging practices of algebraic battery life models through statistical validation and model identification via machine-learning” *J. Electrochem. Soc.* 168(2), 020502 (2021) doi.org/10.1149/1945-7111/abdde1
- M.R. Kunz, E.J. Dufek, Z. Yi, K.L. Gering, M.G. Shirk, K. Smith, B.R. Chen, Q. Wang, P. Gasper, R.L. Bewley, T.R. Tanim “Early battery performance prediction for mixed use charge profiles using hierarchical machine learning” *Batteries & Supercaps* (2021), in press <https://doi.org/10.1002/batt.202100079>
- L. Ward, S. Babinec, E.J. Dufek, D.A. Howey, V. Viswanathan et al “Principles of the Battery Data Genome” *submitted*
- S. Kim, Z. Yi, B.R. Chen, T.R. Tanim, E. Dufek “A Deep Learning Modeling Framework for Early Classification and Quantification of Lithium-ion Aging Modes” *submitted*

# Robust Analysis

- Capture and analyze what effects noise and outliers have on the prediction
  - Necessary in linking to Physics Based Models
  - Assess if more samples are needed
  - Understand variation trends
- Applied ML algorithms:
  - Time series outlier detection
  - Prototypical Clustering





# Supervised forecast algorithm examples

**ARIMA** auto-regressive integrated moving average

## auto-regressive

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

forecast the variable of interest (y) at time point t using a linear combination ( $\phi$ ) of past values ( $y_{t-1} \dots y_{t-p}$ ) of the variable. However, battery degradation is usually **non-linear**.

## moving average

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

uses past forecast errors ( $\varepsilon$ ) in a regression-like model.

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

**Q** – capacity during constant power portion of charge (kinetic)

**V (EOC)** – after rest to link with thermodynamic change

**T** – skin T tracked, followed environmental chamber

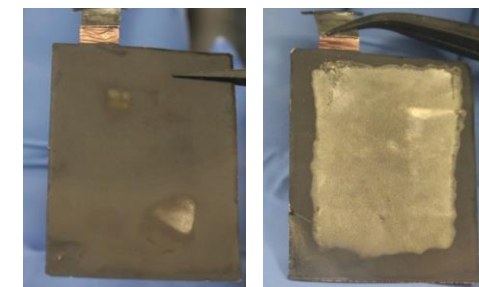
| Input: | $x$ : experimental response for $Q_{cp}$ , EOCV and T, $y$ : RPT capacity at end of life |
|--------|--|
| 1.     | for Each predictor by cycle do   |
| 2.     | Treat outliers within time series  |
| 3.     | Application of ARIMA   |
| 4.     | Reform predictors as X   |
| 5.     | Hierarchical clustering on each cell within X  |
| 6.     | Ridge regression on RPT using DST  |
| 7.     | Generalized Adaptive LASSO with RR coefficients  |

# 31 fast charging NMC532/Gr pouch cells

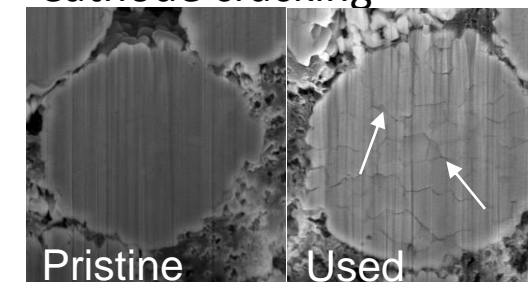
| Cell design group | Charging protocols | Cell numbers   | Anode loading            | Anode porosity |
|-------------------|--------------------|----------------|--------------------------|----------------|
| P1                | 4C CC-CV*          | 01, 02, 03     | 9.94 mg.cm <sup>-2</sup> | 34.5%          |
|                   | 4C 2-step**        | 04, 05, 06     |                          |                |
|                   | 4.5C CC-CV         | 07, 08, 09     |                          |                |
|                   | 6C CC-CV           | 10, 11, 12     |                          |                |
|                   | 6C 2-step          | 13, 14, 15     |                          |                |
|                   | 7.5C MS5***        | 16, 17         |                          |                |
|                   | 9C MS5             | 18, 19         |                          |                |
| P2                | 4.5C CC-CV         | 20, 21, 22     | 9.94 mg.cm <sup>-2</sup> | 22.3%          |
|                   |                    | 23, 24, 25     | 9.94 mg.cm <sup>-2</sup> | 49.1%          |
| P3                | 4C CC-CV           | 26, 27         | 6.38 mg.cm <sup>-2</sup> | 37.4%          |
|                   | 6C CC-CV           | 28, 29, 30, 31 |                          |                |

## Major aging phenomena

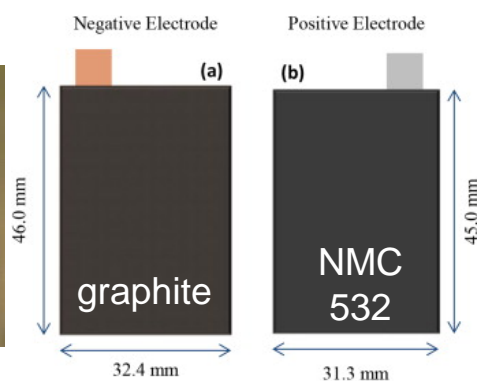
### Li plating



### Cathode cracking



## Cell architecture



## Charging protocols (time = 10 mins)

