INL/CON-22-67069-Revision-0

Bat554: Machine Learning for Accelerated Life Prediction and Cell Design

June 2022

Changing the World's Energy Future

Eric J Dufek

INL/CON-22-67069-Revision-0

Bat554: Machine Learning for Accelerated Life Prediction and Cell Design

Eric J Dufek

June 2022

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

Bat554: Machine Learning for Accelerated Life Prediction and Cell Design

Eric Dufek

Idaho National Laboratory

June 21-25, 2021

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

INL-CON-22-67069

Idaho National Laboratory

Overview

Timeline

- Start: October 1, 2020
- End: September 30, 2022
- Percent Complete: 80%

Budget

- Funding for FY22 – $1.35M

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, and Extreme Fast Charge and Cell Evaluation of Lithium-ion Batteries (XCEL)

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
Generate first reduced order synthetic data aligned with commercial cells	12/31/21	Complete
Post data on batterydata.energy.gov	3/31/22	Complete
Develop tools to predict life and energy based on cell design and limited characterization data	6/30/22	In process
Acquire and start test of cells from life and energy design matrix	9/30/22	In process
Use combined synthetic data and early laboratory assessment to classify and quantify failure modes for aged commercial cells.	9/30/22	In process

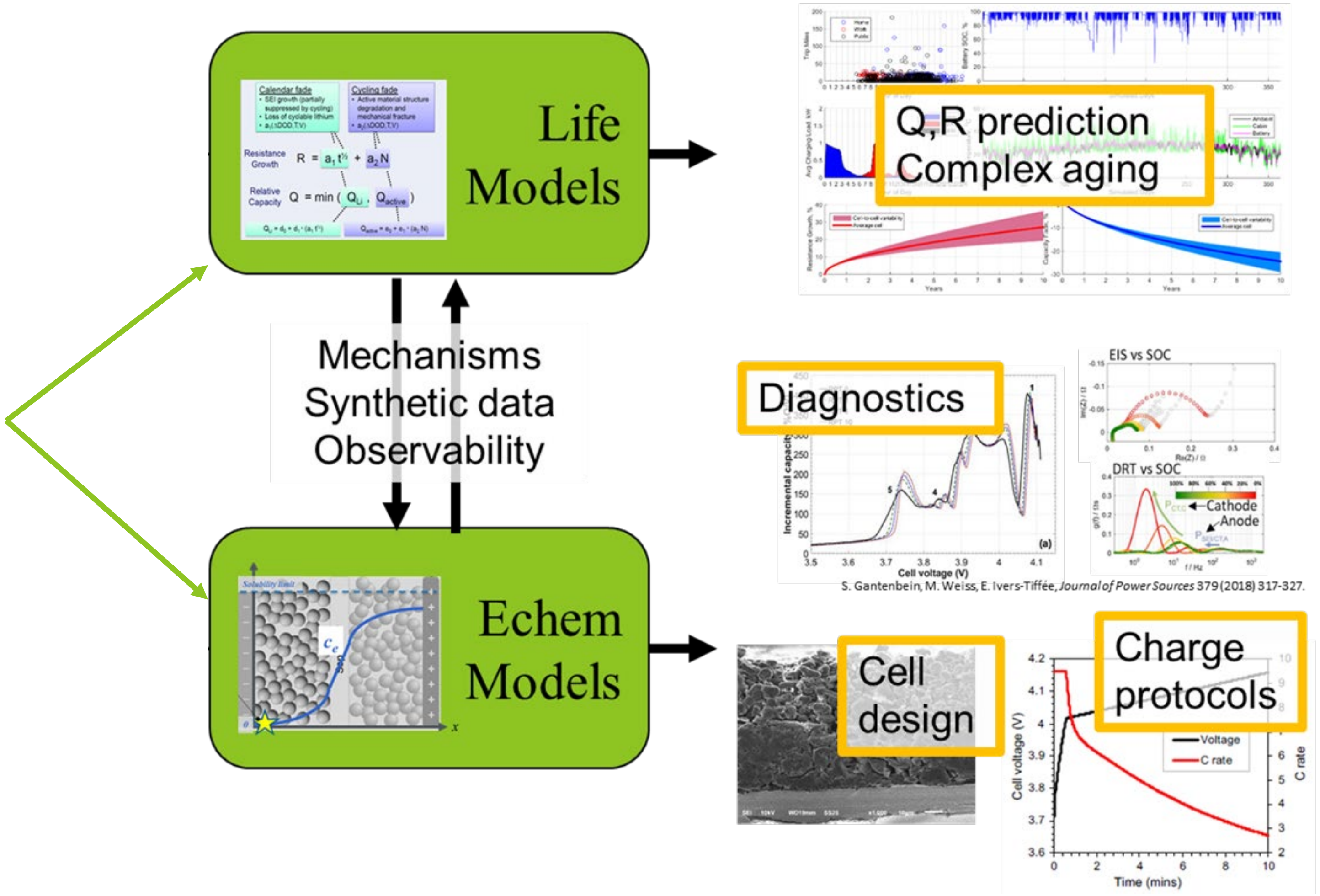
Approach

Accelerating innovation requires failure mode classification, projection and validation
Combination of high-quality data generation, assessment and analysis

Hierarchy of ML tools to use across many complex datasets

Automation of different analyses and model generation

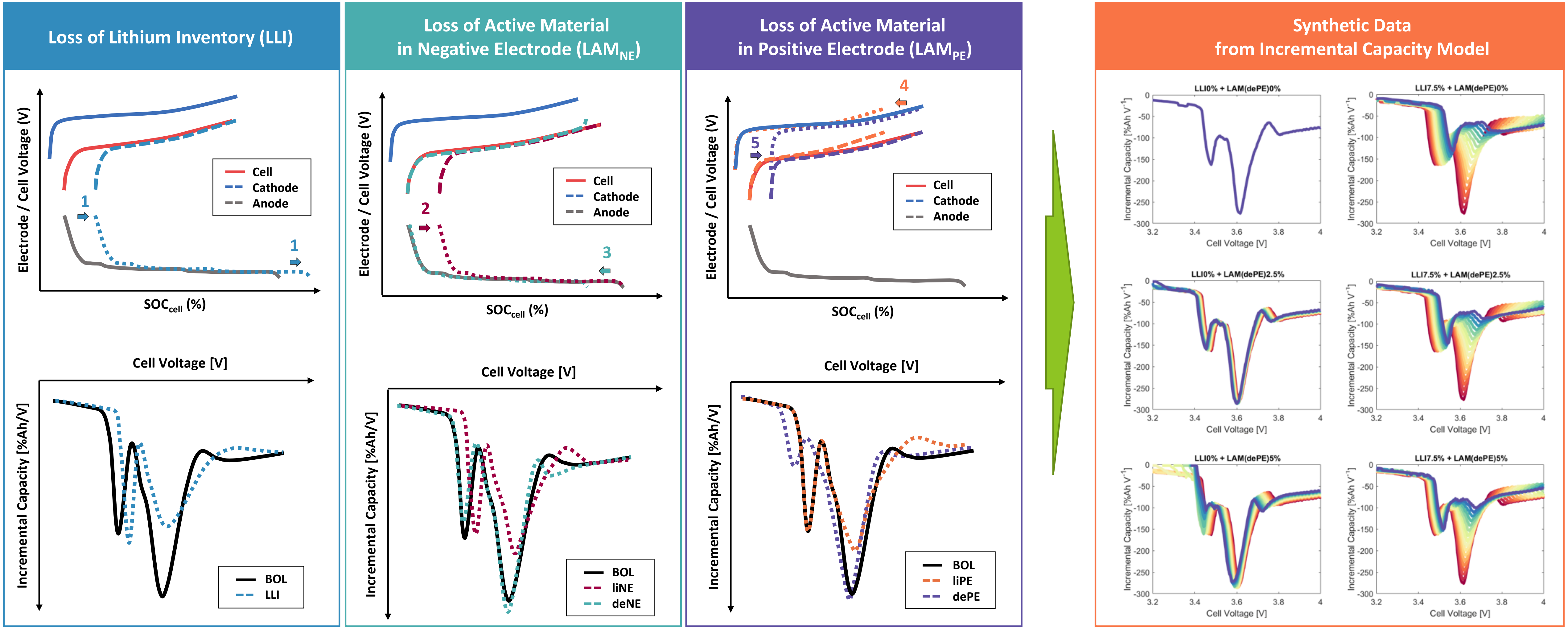
Establish batterydata.energy.gov



Early Understanding of Failure Modes to Reduce Development and Deployment Cost

Kim et al *Energy Storage Materials* (2021)
https://doi.org/10.1016/j.ensm.2021.07.016

Synthetic Data from Incremental Capacity (IC) model

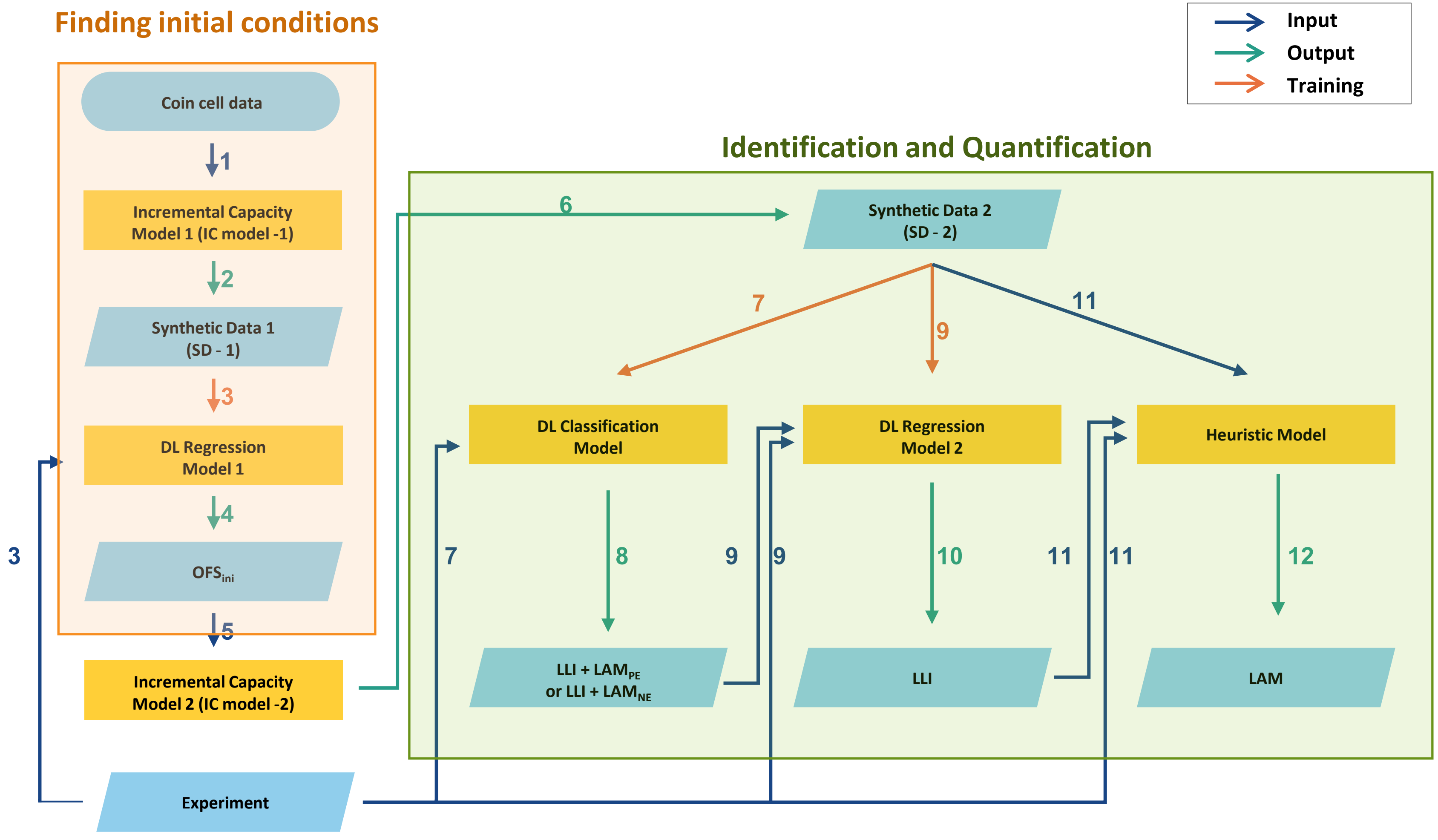


Synthetic data generated by incremental capacity (IC) model

Kim et al *Energy Storage Materials* (Impact Factor 17.78, 2021) https://doi.org/10.1016/j.ensm.2021.07.016

Overview of the DL Modeling Framework

- ❑ **Incremental Capacity (IC) Analysis**
  - Can sweep design and chemistry of batteries
  - Requires extensive experimental data and subject matter experts (SMEs). Uniqueness to solutions is a key issue.
- ❑ **Deep Learning (DL)**
  - Efficient and does not require SME due to automatic classification
  - Requires extensive experimental data for training
- ❑ **Develop an efficient and systematic synthetic data framework to work together with DL algorithms.**



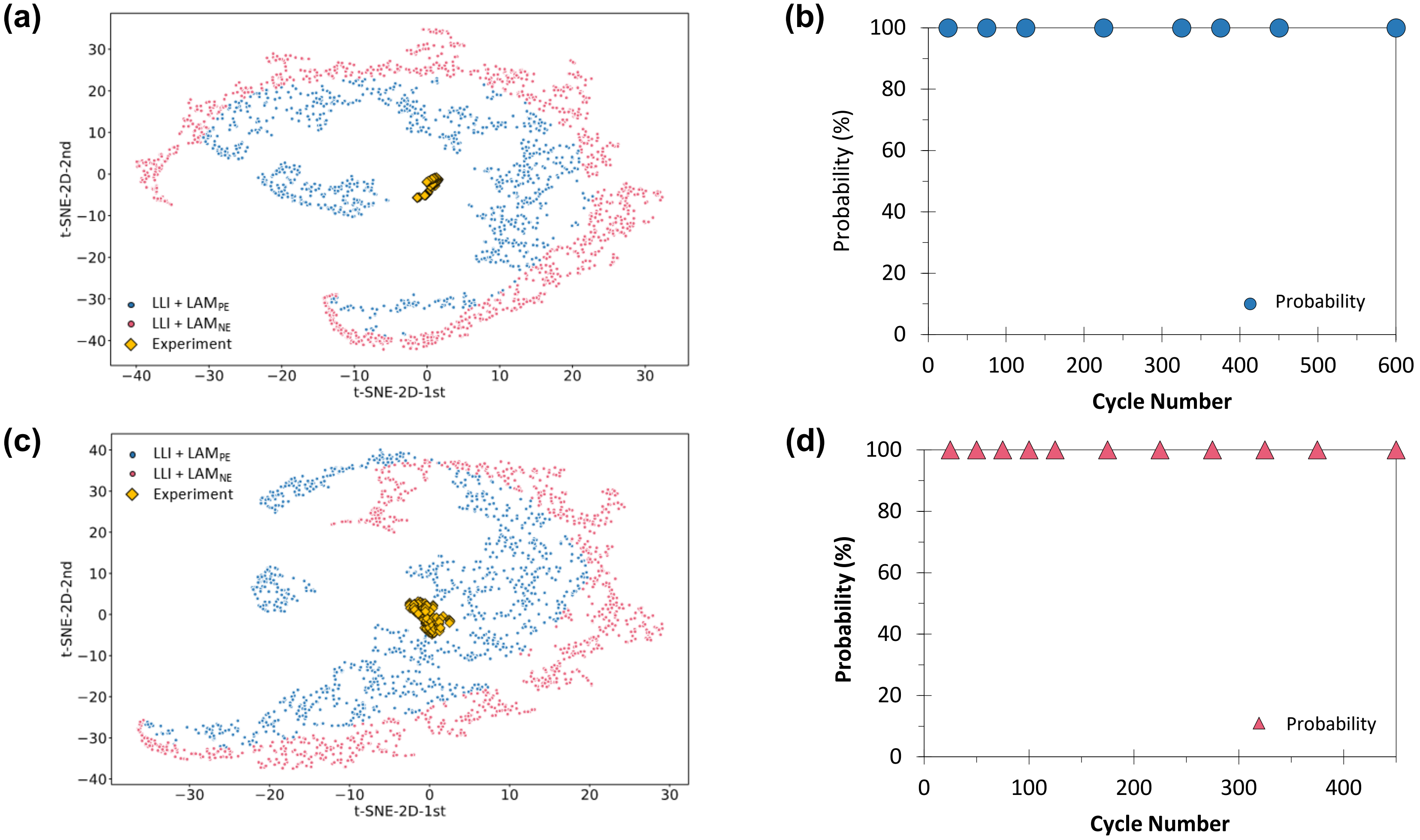
2 sets of IC model-based synthetic data & 3 deep learning & Heuristic Models

Kim et al *Energy Storage Materials* (2021)
https://doi.org/10.1016/j.ensm.2021.07.016

Deep Learning-based Early Classification of Aging Modes

13 Low loading cells
1C: 6cells
6C: 3 cells
9C: 4 cells

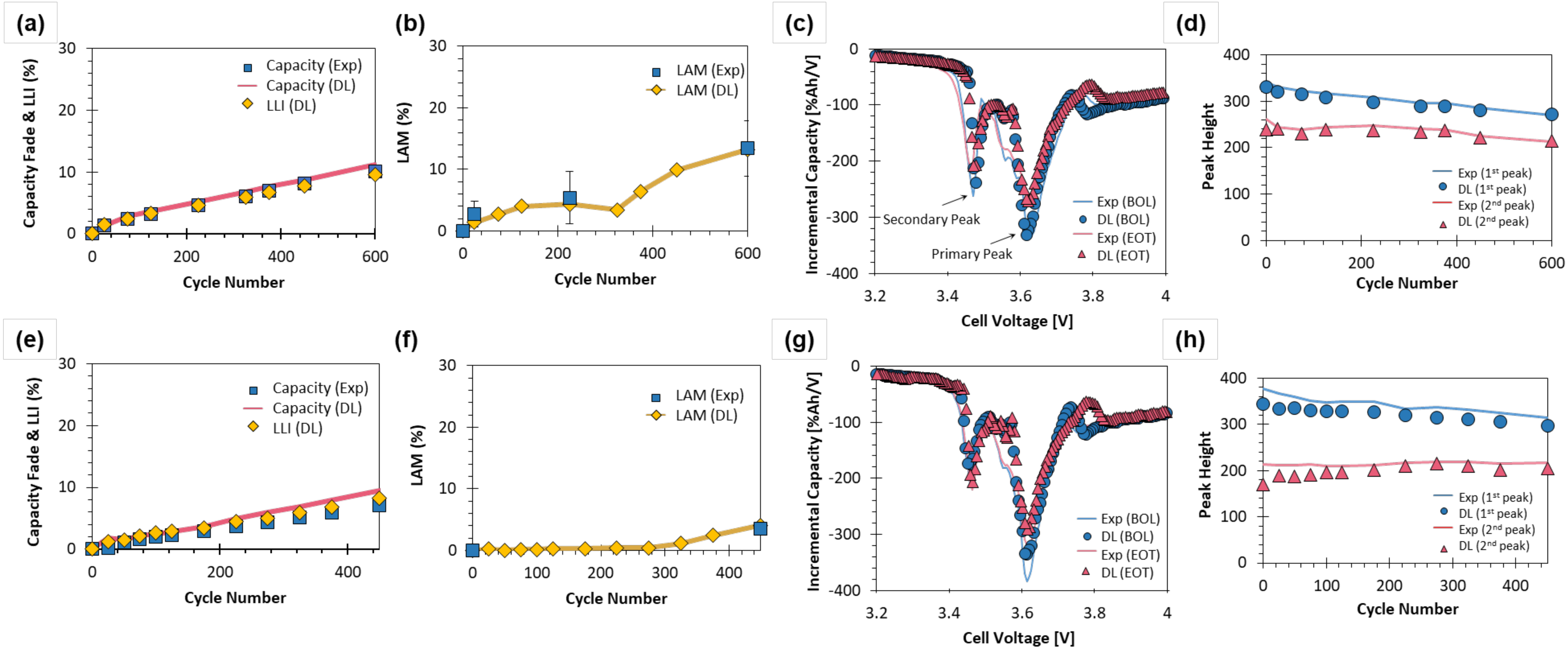
9 Moderate loading cells
1C: 3 cells
4C: 3 cells
6C: 3 cells



Successfully differentiates the dominant aging mode combination

Kim et al *Energy Storage Materials* (2021)
https://doi.org/10.1016/j.ensm.2021.07.016

Validation of Aging Constituents



DL framework successfully validated with experimental data

Performance and Failure Mode Prediction Frameworks

Needs of Experimental Data →

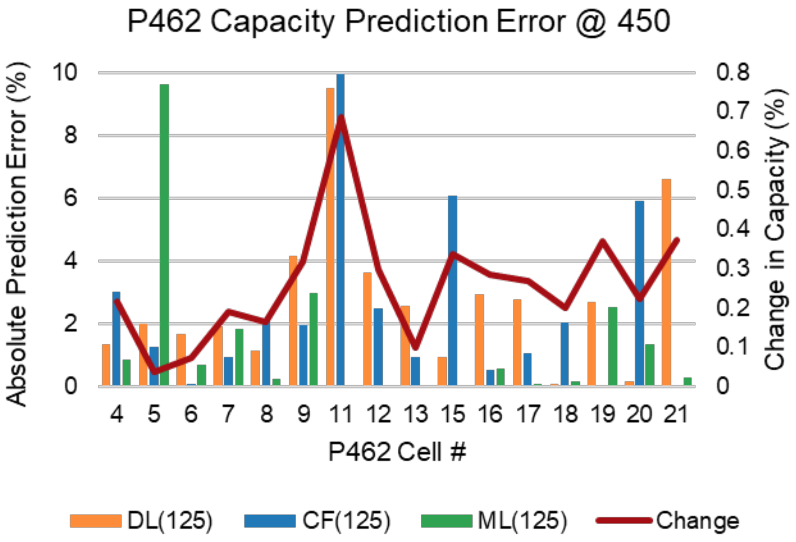
	Deep Learning (DL) Prediction	Curve Fitting (CF) Prediction	Machine Learning (ML) Prediction
Method	Simulation + DL	Curve Fitting	ML (Random Forest)
Input	Early cycle RPT data & SRE simulation data	Early cycle RPT data	Training data includes experimental data with late life information
Output	Capacity fade & LLI prediction	Capacity fade & LLI prediction	Capacity Fade & LLI & LAM prediction
Distinct Features	A created model can work for all cells in the same group	Each cell needs a specific set of model parameters	All cells in different groups are used together

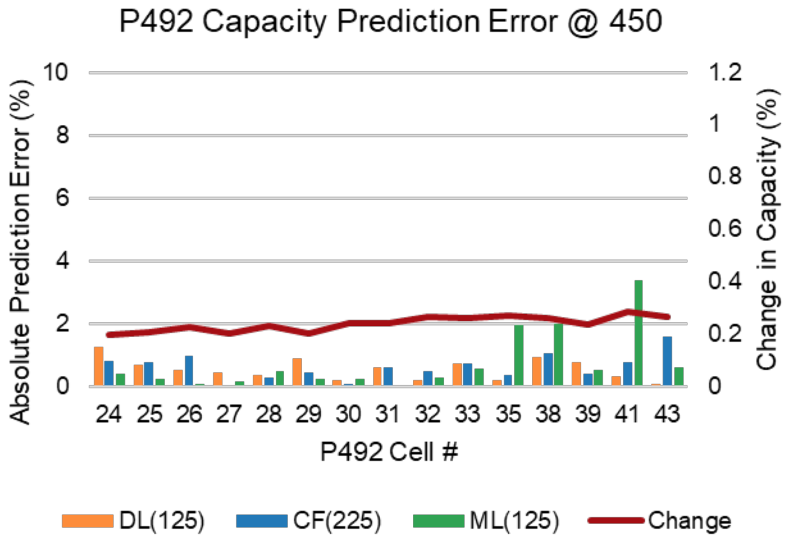
Comparison of different prediction framework for cells undergoing fast charge

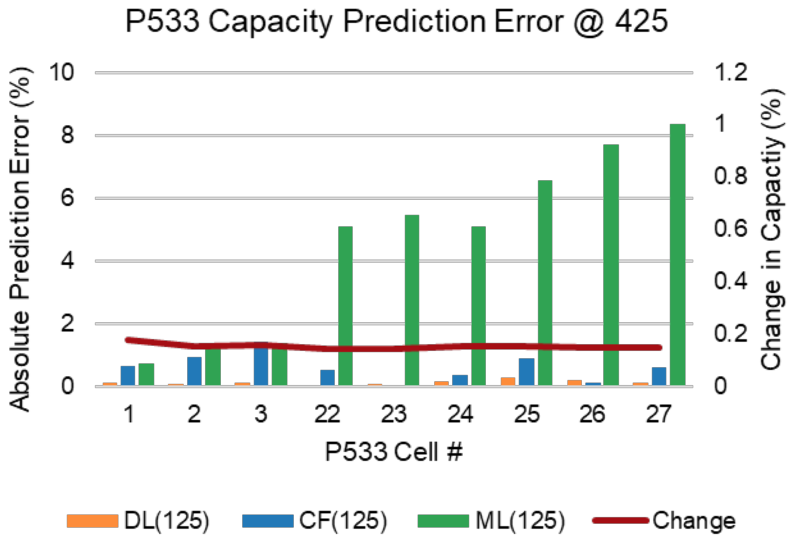
Kim. et al, submitted

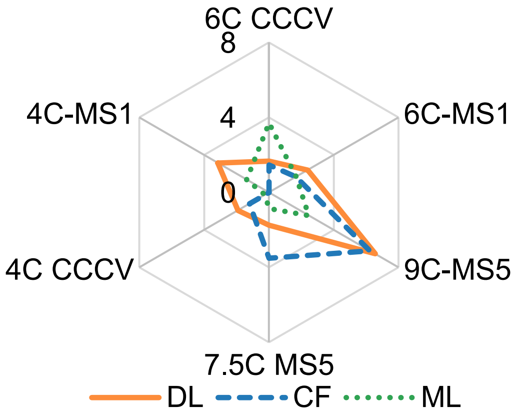
Graphite/NMC
Gen 2 electrolyte
P462 low loading (2.0 mAh/cm^2)
P492&533 moderate loading (3.0 mAh/cm^2)

Capacity Prediction

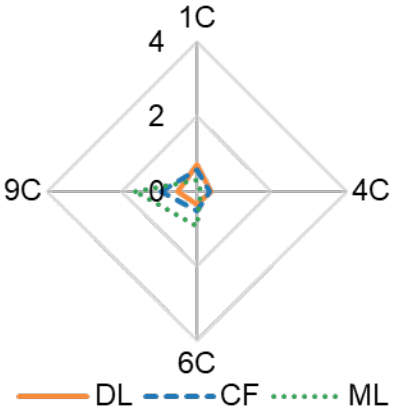




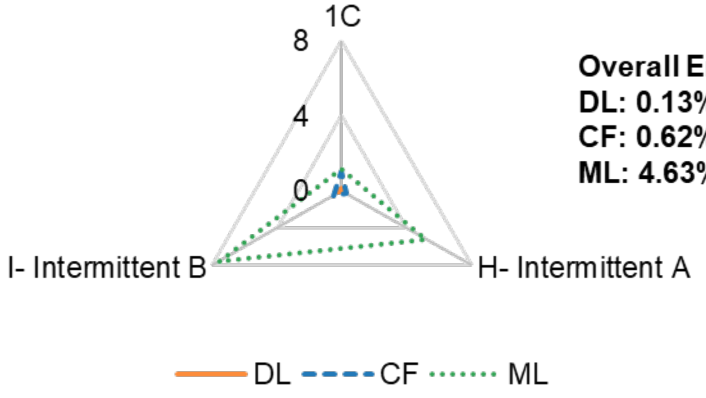




Overall Error
DL: 2.91%
CF: 2.81%
ML: 1.72%



Overall Error
DL: 0.49%
CF: 0.60%
ML: 0.74%



Overall Error
DL: 0.13%
CF: 0.62%
ML: 4.63%

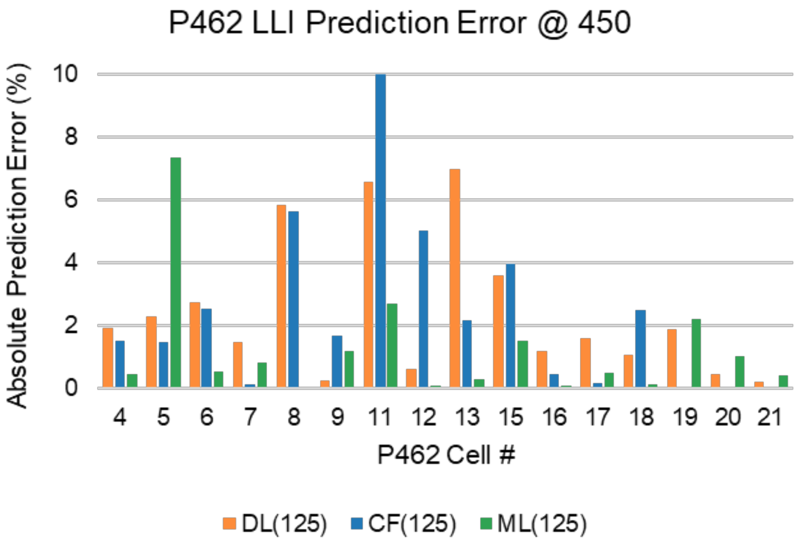
Capacity is successfully predicted based on limited cycling data with less than 4.63% error

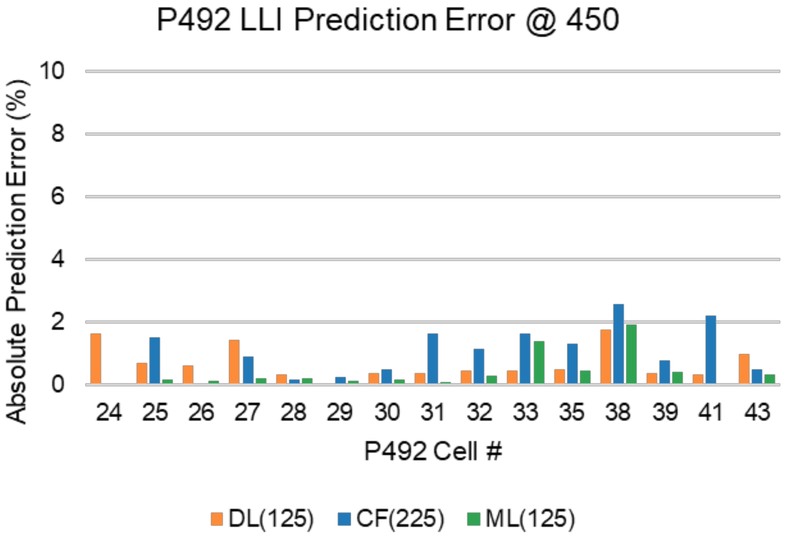
* CF method in P492 requires 225 cycling data due to limited RPT data

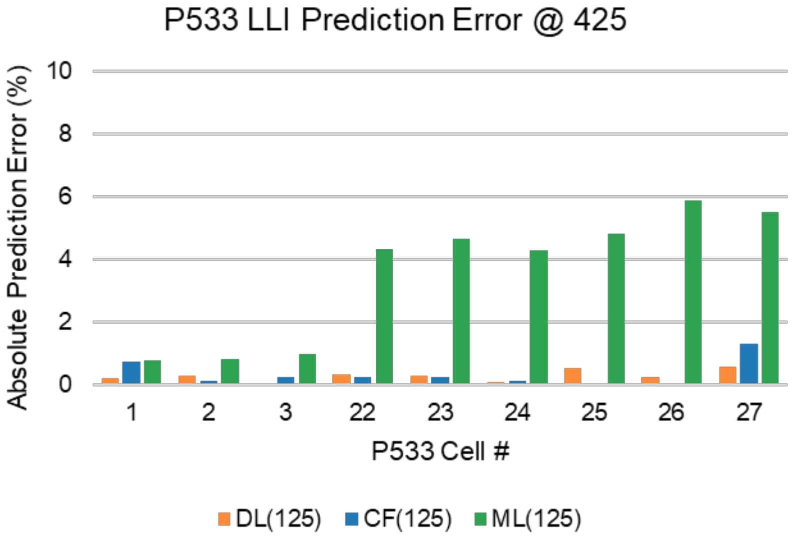
Kim. et al, submitted

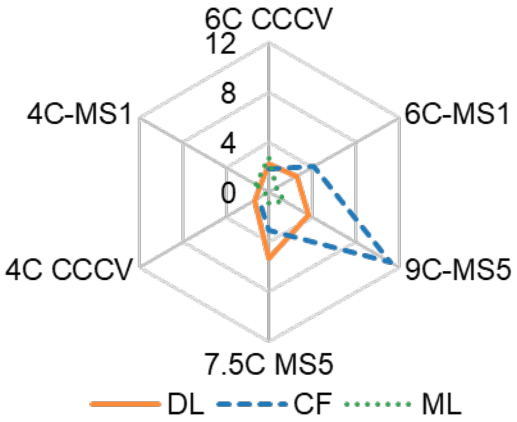
Loss of Lithium Inventory (LLI) Prediction

Graphite/NMC
Gen 2 electrolyte
P462 low loading (2.0 mAh/cm^2)
P492&533 moderate loading (3.0 mAh/cm^2)

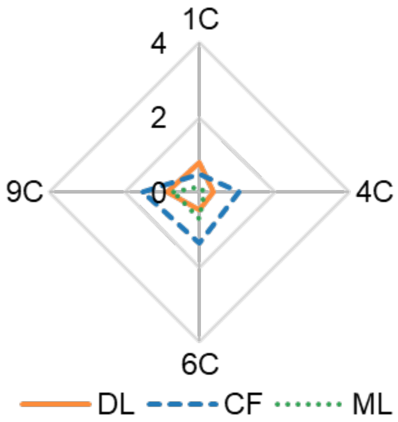




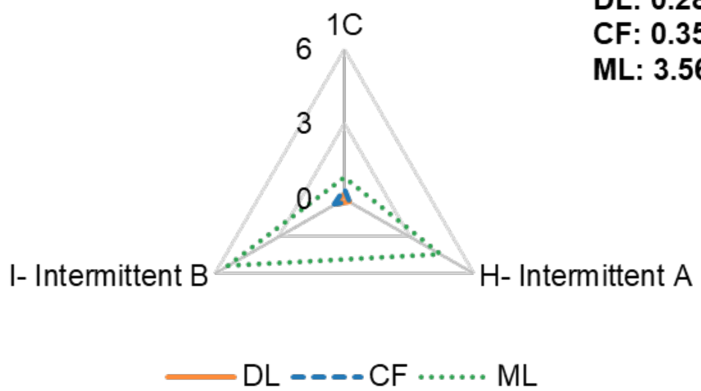




Overall Error
DL: 2.65%
CF: 4.28%
ML: 1. 23%



Overall Error
DL: 0.62%
CF: 1.10%
ML: 0.41%



Overall Error
DL: 0.28%
CF: 0.35%
ML: 3.56%

LLI is successfully predicted based on 125 cycling data with less than 4.28% error*

* CF method in P492 requires 225 cycling data due to limited RPT data

Machine-learned calendar life model (Gasper, JES, 2021)

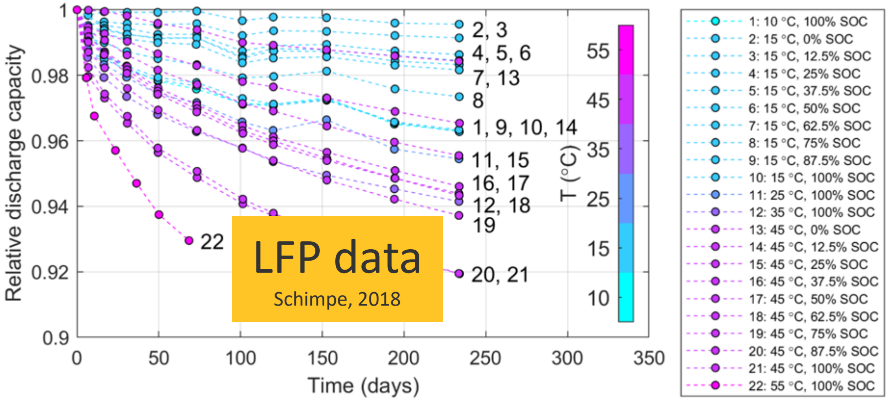


Figure 2: Relative discharge capacity versus time during calendar aging of LFP/graphite cells from Schimpe et. al.[18]. Cells at high temperature and high SOC degrade much more rapidly than cells at low temperature and low SOC. Several data series at differing conditions result in similar degradation, indicating that multiple temperature and SOC effects impact degradation.

Local models… w/ various rate laws

- Physically informed, "ArrTfl" $\quad \boldsymbol{\beta_1(\gamma, T, U_a)} = \gamma_0 \exp\left(\gamma_1 \frac{\boldsymbol{1}}{\boldsymbol{T}}\right) exp\left(\gamma_2 \frac{\boldsymbol{U_a}}{\boldsymbol{T}}\right)$
- Modified/empirical, "ArrTfl$_{mod}$" $\quad \boldsymbol{\beta_1(\gamma, T, U_a)} = \gamma_0 \exp\left(\gamma_1 \frac{\boldsymbol{1}}{\boldsymbol{T}}\right)\left(\boxed{\gamma_2} + exp\left(\gamma_3 \frac{\boldsymbol{U_a}}{\boldsymbol{T}}\right)\right)$
  Schimpe, 2018
- Machine learned (symbolic regression)

$$\beta_1 = \gamma_0 \cdot exp(\gamma_1 T^2) exp(\gamma_2 \sqrt{U_a}/T^2) exp(\gamma_3 T^2/\sqrt{U_a}) exp(\gamma_4/(U_a^2 T^2)) exp(\gamma_5/(U_a^3 T^2))$$

Model ID	Description	Model equation	# of global params (α)	# of local params (β)	Total # params
1	Power law ($t^{0.5}$)	$q = 1 - \beta_1 t^{0.5}$	0	1	16
2	Power law	$q = \alpha_0 - \beta_1 t^{\alpha_1}$	2	1	18
3	Power law	$q = \alpha_0 - \beta_1 t^{\beta_2}$	1	2	33
4	Power law	$q = \beta_0 - \beta_1 t^{\alpha_1}$	1	2	33
5	Power law	$q = \beta_0 - \beta_1 t^{\beta_2}$	0	3	48
6	Stretched exponential	$q = \alpha_0 - \beta_1(1 - 1/exp((\alpha_1 t)^{\alpha_2}))$	3	1	19
7	Stretched exponential	$q = \alpha_0 - \beta_1(1 - 1/exp((\beta_2 t)^{\alpha_1}))$	2	2	34
8	Stretched exponential	$q = \alpha_0 - \beta_1(1 - 1/exp((\alpha_1 t)^{\beta_2}))$	2	2	34
9	Stretched exponential	$q = \alpha_0 - \beta_1(1 - 1/exp((\beta_2 t)^{\beta_3}))$	1	3	49
10	Stretched exponential	$q = \beta_0 - \beta_1(1 - 1/exp((\beta_2 t)^{\beta_3}))$	0	4	64
11	Sigmoidal	$q = \alpha_0 - 2\beta_1\left(\frac{1}{2} - \frac{1}{1 + exp((\alpha_2 t)^{\alpha_3})}\right)$	3	1	19
12	Sigmoidal	$q = \alpha_0 - 2\beta_1\left(\frac{1}{2} - \frac{1}{1 + exp((\beta_2 t)^{\alpha_2})}\right)$	2	2	34
13	Sigmoidal	$q = \alpha_0 - \boxed{2\beta_1}\left(\frac{1}{2} - \frac{1}{1 + exp((\alpha_2 t)^{\beta_2})}\right)$	2	2	34
14	Sigmoidal	$q = \alpha_0 - 2\beta_1\left(\frac{1}{2} - \frac{1}{1 + exp((\beta_2 t)^{\beta_3})}\right)$	1	3	49
15	Sigmoidal	$q = \beta_0 - 2\beta_1\left(\frac{1}{2} - \frac{1}{1 + exp((\beta_2 t)^{\beta_3})}\right)$	0	4	64

Table 4: Model structures studied in this work.

Model selection using cross validation

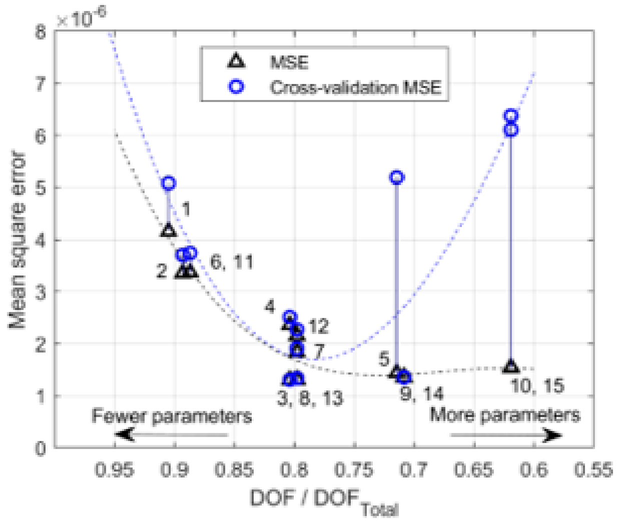
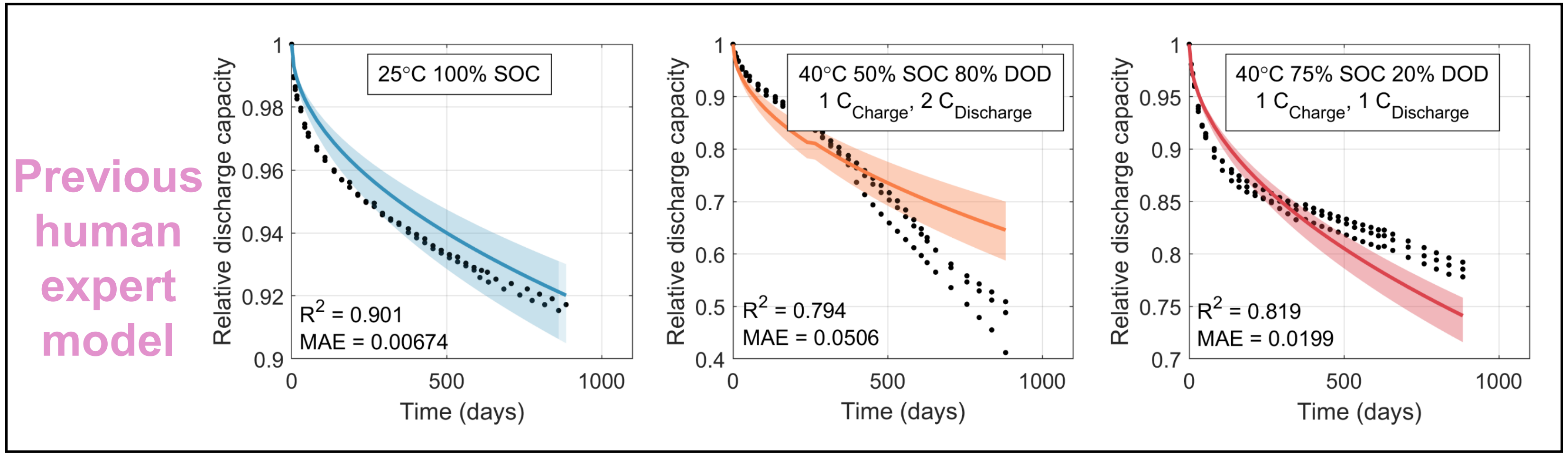
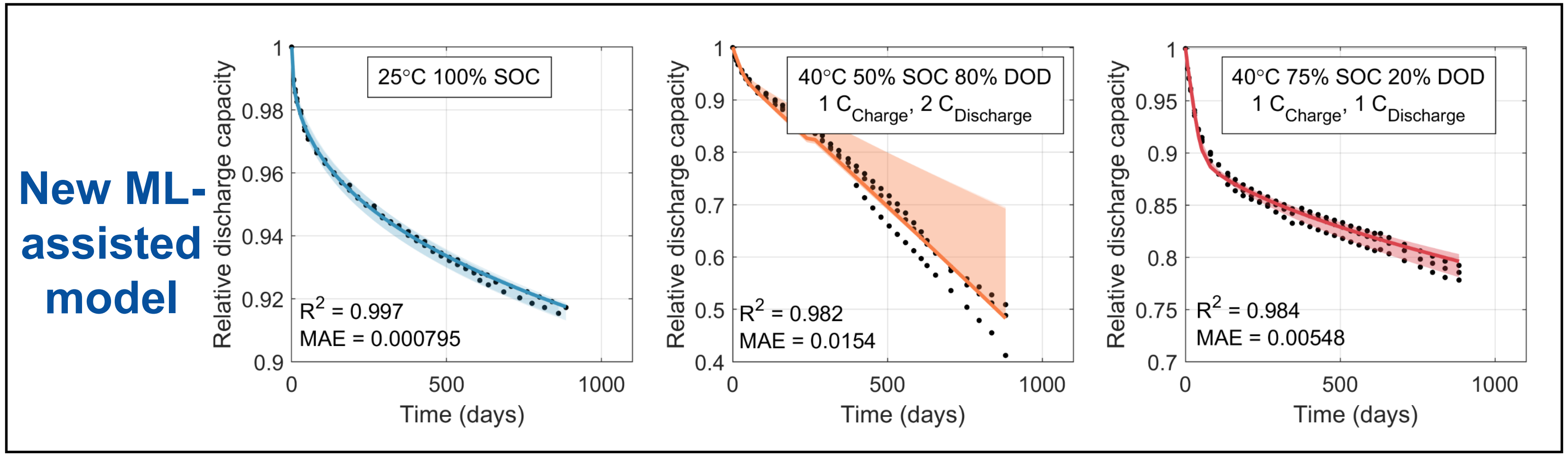


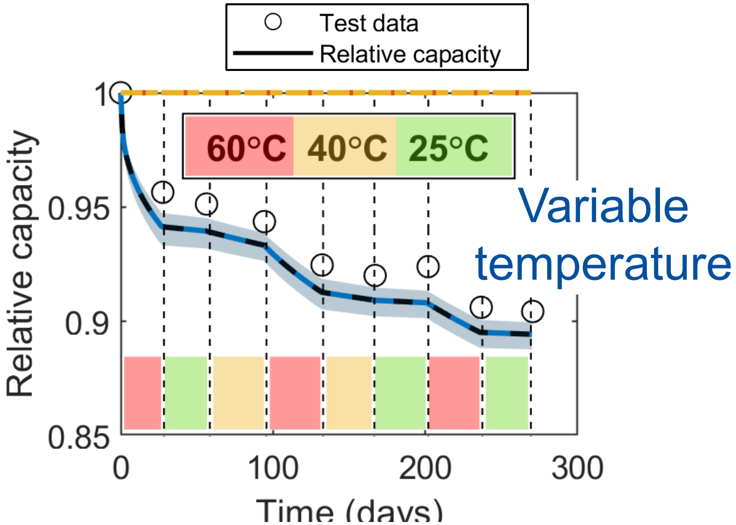
Figure 3: MSE and MSEcv after bi-level optimization of models when trained on all RPT data. Polynomial trendlines have been added to each data series to guide the eye.

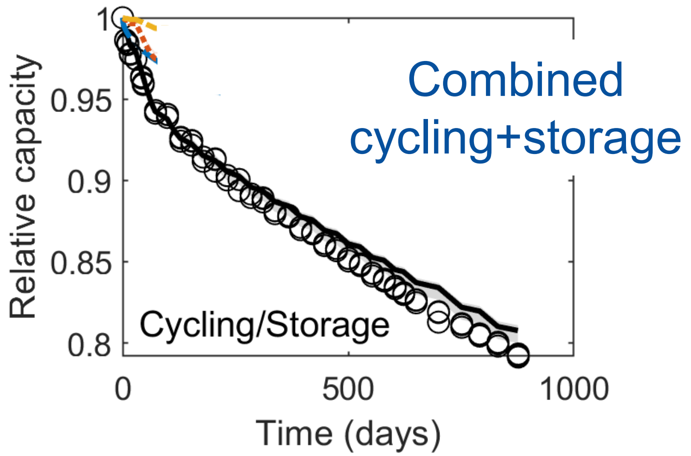
Combined calendar- and cycle-life model for LFP/Gr





Independent validation





New ML-assisted life model has 50% less error than previous human expert model

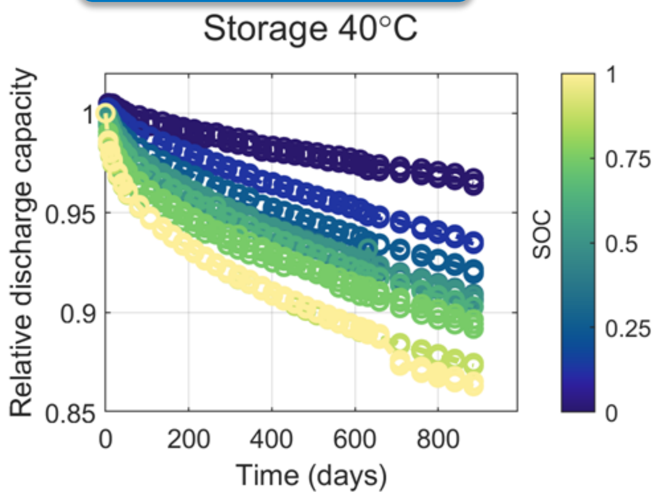
P. Gasper et al., "Machine-learning assisted identification of accurate battery lifetime models with uncertainty," *submitted*.

AI-Batt Tool

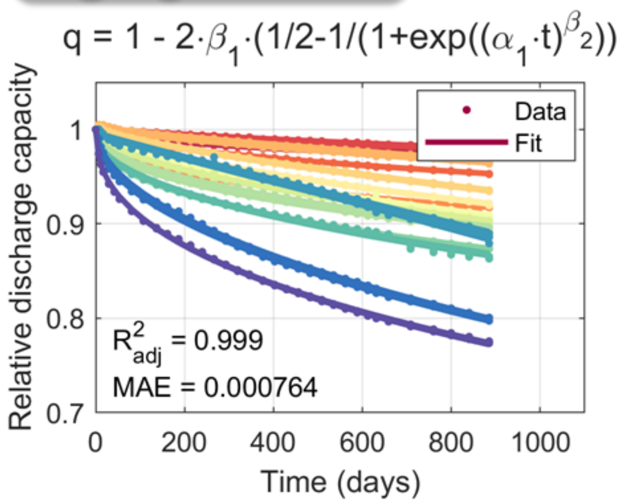
Reduced-order lifetime modeling

Lifetime = f(T,SOC,DOD,C-rate)

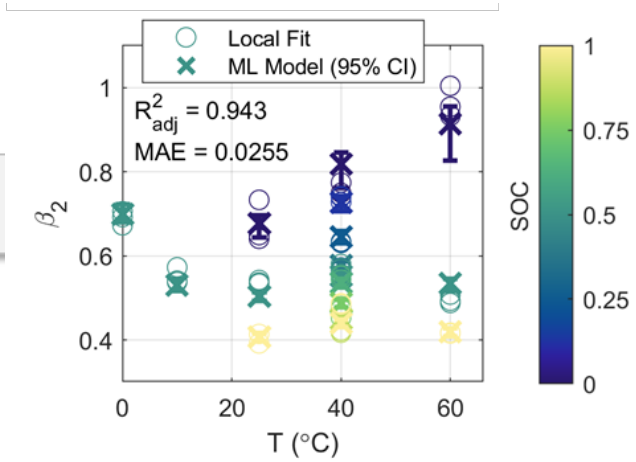
Visualization



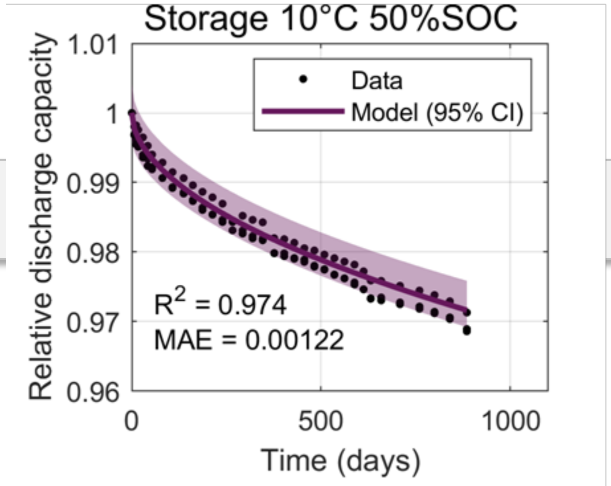
Deconvolute aging modes



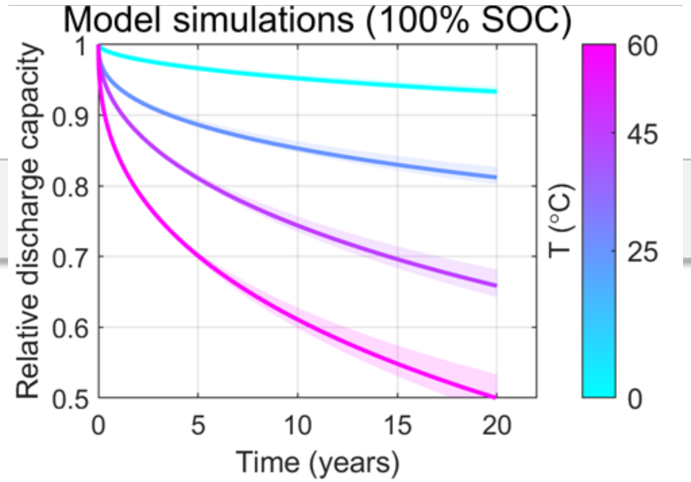
ML modeling of aging rate



Model vs data w/ uncertainty



Extrapolation



Lifetime Models of Various Li-ion Technologies

Gr/LFP

Journal of The Electrochemical Society, 2021 **168** 020502

Challenging Practices of Algebraic Battery Life Models through Statistical Validation and Model Identification via Machine-Learning

Paul Gasper,[1,z] Kevin Gering,[2] Eric Dufek,[2] and Kandler Smith[1]

[1]Energy Conversion and Storage Systems Center, National Renewable Energy Laboratory, Golden, CO, 80401, United States of America
[2]Energy Storage & Advanced Vehicles Department, Idaho National Laboratory, Idaho Falls, ID, 83415, United States of America

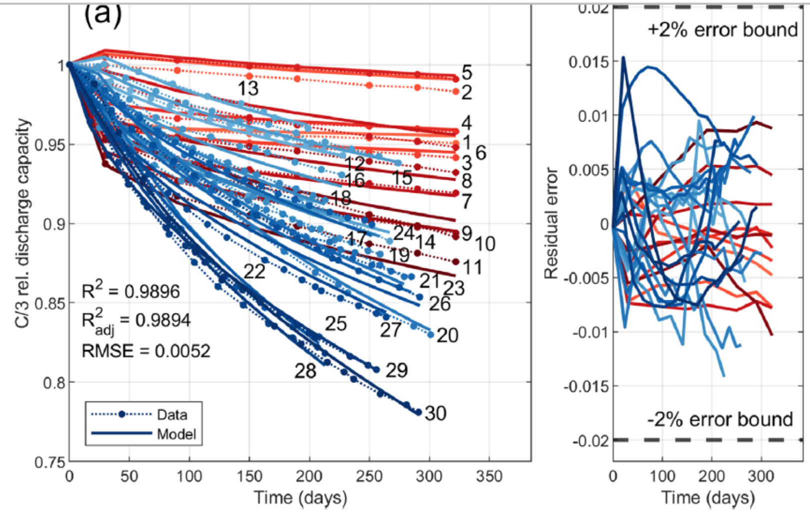
Gr/NMC622

Journal of The Electrochemical Society, 2021 **168** 100530

Lithium-Ion Battery Life Model with Electrode Cracking and Early-Life Break-in Processes

Kandler Smith,[1,z] Paul Gasper,[1] Andrew M. Colclasure,[1] Yuta Shimonishi,[2] and Shuhei Yoshida[2]

[1]National Renewable Energy Laboratory, Golden, Colorado, United States of America
[2]DENSO Corporation, Kariya, Aichi 448-866, Japan



- Results have supported DOE techno-economics, application, and systems studies (XCEL, BTMS, Grid…)

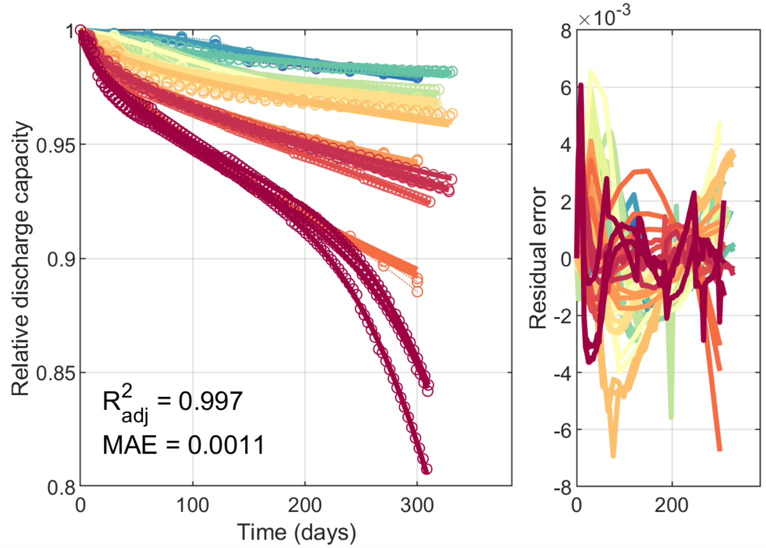
LTO/LMO

$$q = 1 - q_{Loss,Cal} - q_{Loss,Cyc} + q_{Gain}$$

$$q_{Loss,Cal} = 2 \cdot \boldsymbol{q_1} \cdot \left(\frac{1}{2} - \frac{1}{1 + exp\left((q_2 t)^{q_3}\right)}\right)$$

$$q_{Loss,Cyc} = (\boldsymbol{q_7} \cdot EFC)^{q_8}$$

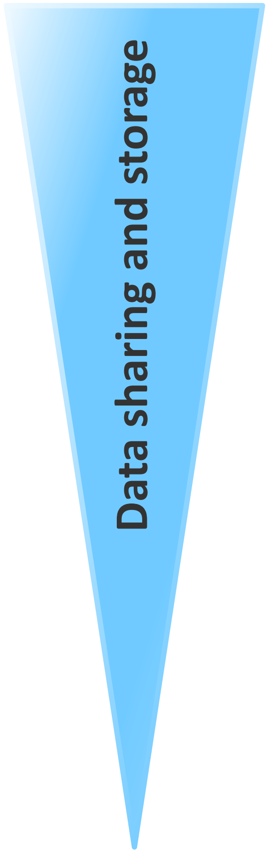
$$q_{Gain} = 2 \cdot q_4 \cdot \left(\frac{1}{2} - \frac{1}{1 + exp\left((\boldsymbol{q_5} t)^{q_6}\right)}\right)$$



Batterydata.energy.gov – Vision

Leverage existing and new activities to benefit development from low thru high TRL

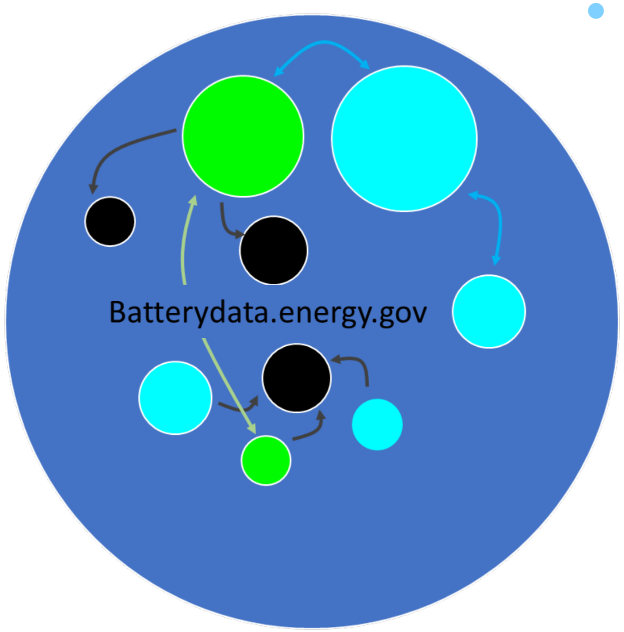
High TRL



Low TRL

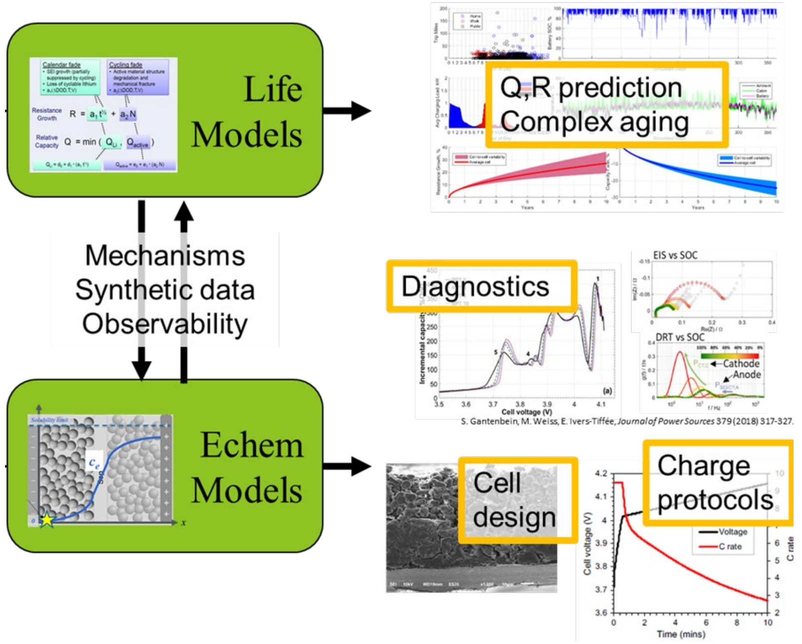
Data capture and storage across organizations and programs

- Proprietary
- Public



- = batterydata.energy.gov
- = Totally Open
- = Partially Open
- = Totally Closed
- = Fully open Bidirectional information flow
- = Partially open Bidirectional information flow
- = Unidirectional information flow

Common analysis tools to enable rapid analysis, reporting, and transfer learning

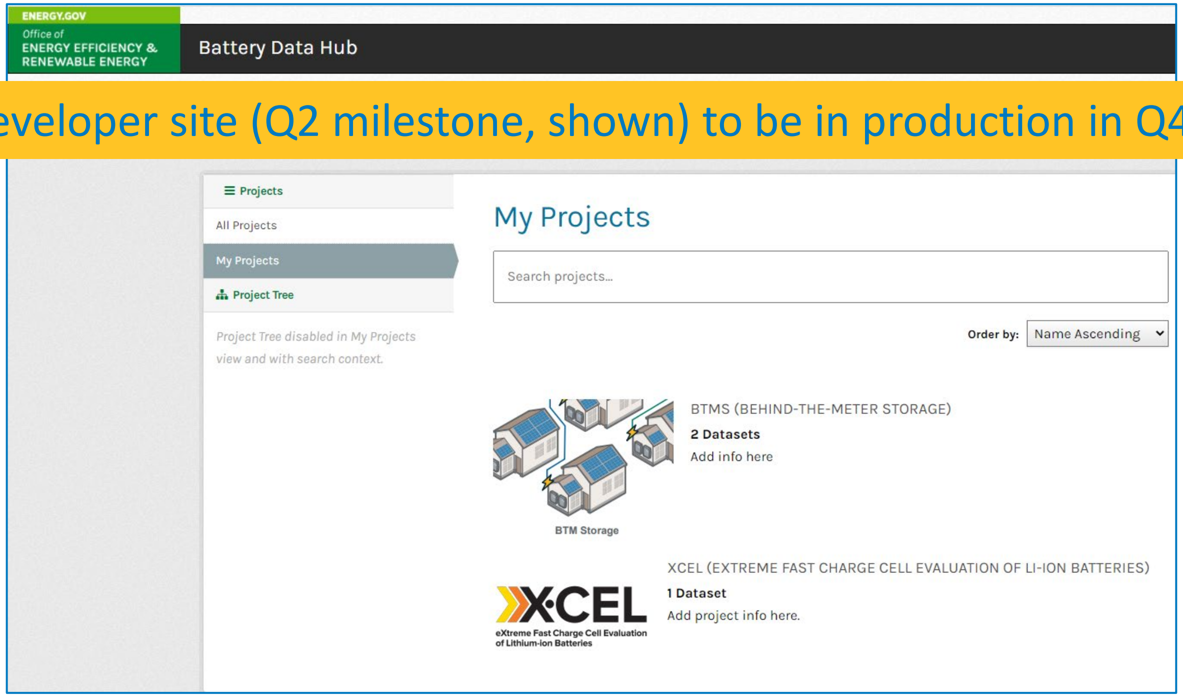


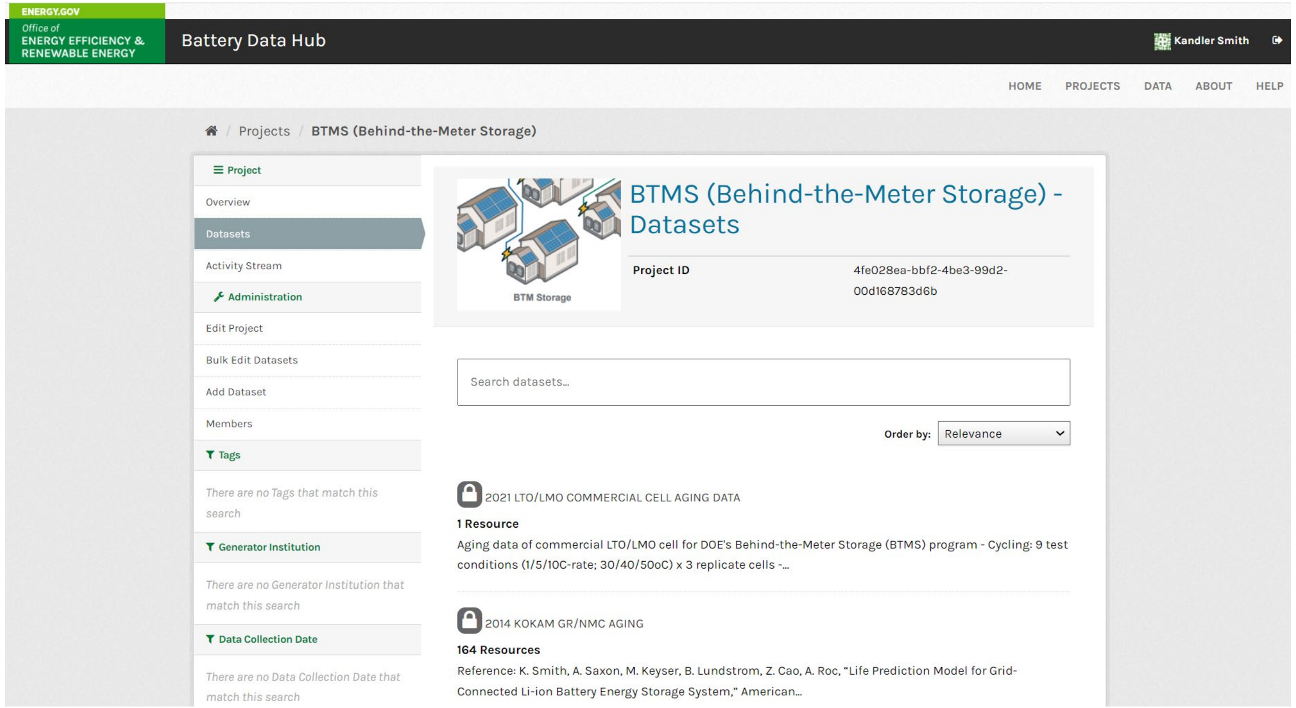
Batterydata.energy.gov is being built on pre-existing Energy Materials Network (EMN) Platform

Other EMN sites: ElectroCat, HydroGen, HyMARC, ChemCatBio, DuraMAT

- Support collaborative science through secure sharing of data among project team members.
    - ‘Moderate’ datahub with two-factor authentication
    - Suitable for DOE and other U.S. government R&D programs
    - Make selected datasets publicly available when ready (e.g. upon publication).
- Search across all data for which you have permission using defined metadata.
- Access advanced data analysis tools.

Developer site (Q2 milestone, shown) to be in production in Q4.





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 and cell design
- 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

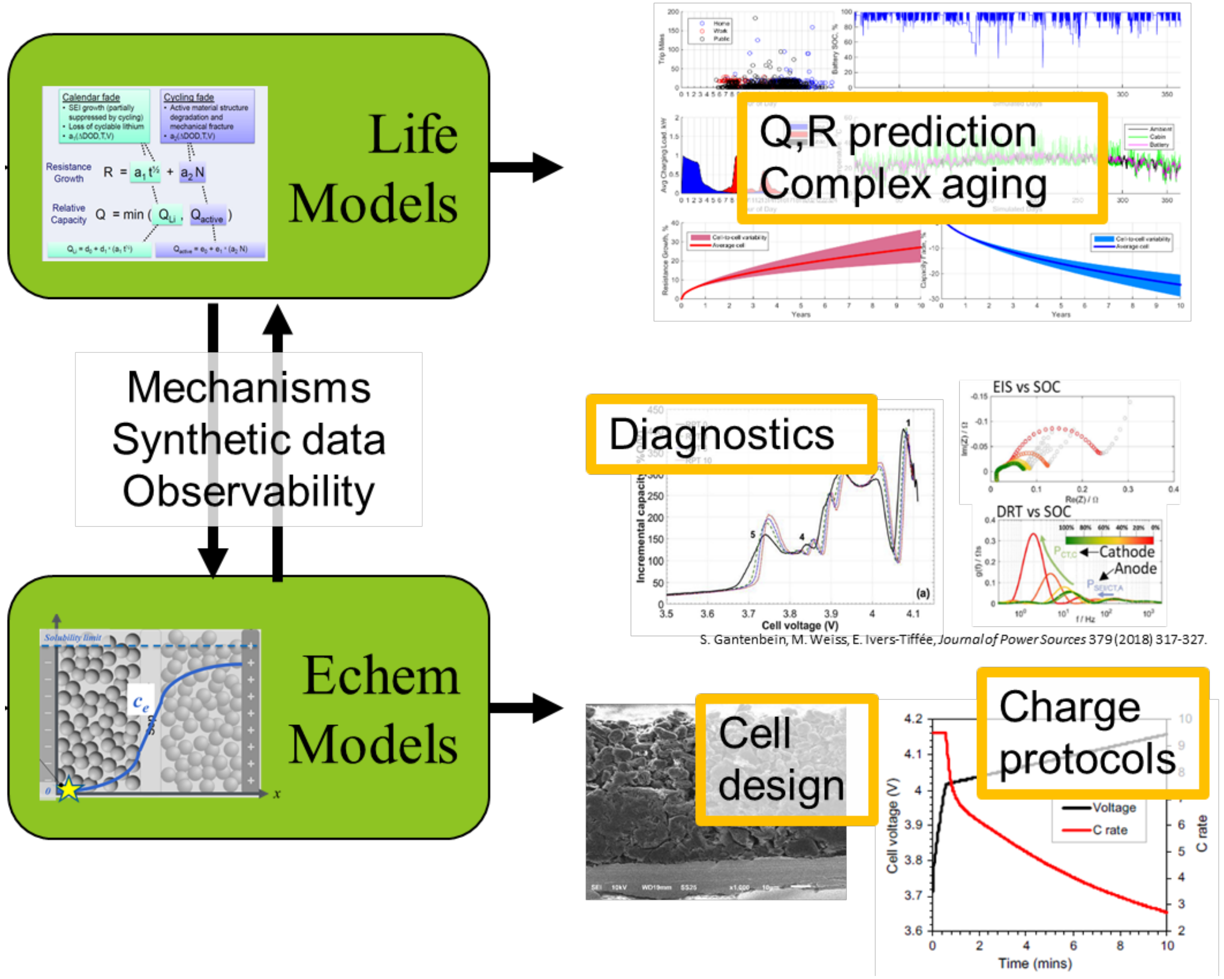
Any proposed future work is subject to change based on funding levels

Response to previous years reviews

- The reviewer noted that the approach involves machine learning to accelerate the cycle life prediction of batteries including capturing degradation mechanisms through logistic regressions. The reviewer commented that the milestone lists "initiate Deep Learning related to electrochemical signatures", while current models are mostly shallow learning models. The reviewer indicated that there may not be a need to do deep learning, but that it would be good to check and show that shallow learning models work well enough.
  - *The work described in the presentation includes a comparison of deep learning, traditional fitting and a ML model. Each of the methods has advantages as shown above*
- The reviewer acknowledged that the proposed project plan is sound. The reviewer remarked that the process for collaboration of data will be a key development activity that will be broadly useful, and that setting the right practices in place would be crucial.
  - *The team has worked to establish batterydata.energy.gov to enhance the ability to share data. The team has also been involved in multiple collaborations and outreach events including through MRS tutorials and perspective articles to help establish practices across the community*

Summary

- Auto-generation 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
- Generated IC-based synthetic data and integrated into a deep learning model
- Established an ML classification framework that classifies aging modes
- Compared multiple methods to identify both failure mode and performance
- Established and uploaded data to batterydata.energy.gov



Contributors, Collaborators and Acknowledgements

Eric Dufek	Kandler Smith
Tanvir Tanim	Andrew Colclasure
Kevin Gering	Nalinrat Guba
Matthew Shirk	Rachel Hurst
Jack Deppe	Peter Weddle
Bor-Rong Chen	Paul Gasper
Cody Walker	Emily Harrell
Sangwook Kim	Malik Hassanaly
Ross Kunz	Ryan King
Zonggen Yi	

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 Data

Reviewer Only Slides

Publications

- 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 hierarchal machine learning" *Batteries & Supercaps* (2021), 4(7), 1186. 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" *https://arxiv.org/abs/2109.07278*
- 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" *Energy Storage Materials* (2022), 45, 1002. https://doi.org/10.1016/j.ensm.2021.07.016
- K. Smith, P. Gasper, A.M. Colclasure, Y. Shimonishi, S. Yoshida "Lithium-ion Battery Life Model with Electrode Cracking and Early-Life Break-in Processes" *J. Electrochem. Soc* (2021), 168, 100530. https://doi.org/10.1149/1945-7111/ac2ebd
- S. Kim, Z. Yi, M.R. Kunz, E.J. Dufek, T.R. Tanim, B.R. Chen, K.L. Gering, "Accelerated Battery Life Predictions through Synergistic Combination of Physics-based models and Machine Learning" *submitted.*
- E.J. Dufek, T.R. Tanim, B.R. Chen, S. Kim "Battery calendar aging: Enhancing uniformity through adoption of common practices" *submitted.*
- P.M. Attia, A. Bills, F.B. Planella, P. Dechent, G. dos Reis, M. Dubarry, P. Gasper, R. Gichrist, S. Greenbank, D. Howey, O. Liu, E. Khoo, Y. Preger, A. Soni, S. Sripad, A.G. Stefanopoulou, V. Sulzer, ""Knees" in lithium-ion battery aging trajectories," *accepted*.
- P. Gasper, K. Gering, E. Dufek, K. Smith, "Challenging Practices of Algebraic Battery Life Models through Statistical Validation and Model Identification via Machine-Learning," *Journal of the Electrochemical Society* 168, 020502 (2021)
- P. Gasper, N. Colath, H.C. Hesse, A. Jossen, K. Smith, "Machine-learning assisted identification of accurate battery lifetime models with uncertainty," *submitted*.
- P. Gasper, A. Schiek, K. Smith, Y. Shimonishi, S. Yoshida, "Predicting battery capacity from impedance at varying temperature and state-of-charge using machine learning," *submitted*.
- C. Xu., P. Behrens, P. Gasper, K. Smith, M. Hu, A. Tukker, B. Steubing, "Electric vehicle batteries alone could satisfy short-term grid storage demand as early as 2030." *submitted*.

Kim et al *Energy Storage Materials* (Impact Factor 17.78, 2021) https://doi.org/10.1016/j.ensm.2021.07.016

Deep Learning Models

Synthetic data

Synthetic Data (SD-1)

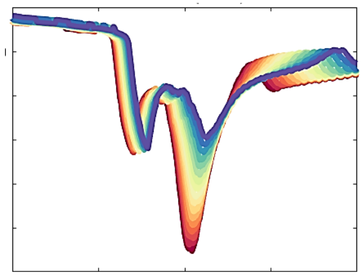
6000 initial conditions (~100MB)
OFS_{ini}: 0 - 15%

Synthetic Data (SD-2)

20,000 aging modes (~1GB)
LLI: 0 - 50%
$LAM_{PE\ or\ NE}$: 0 - 50%

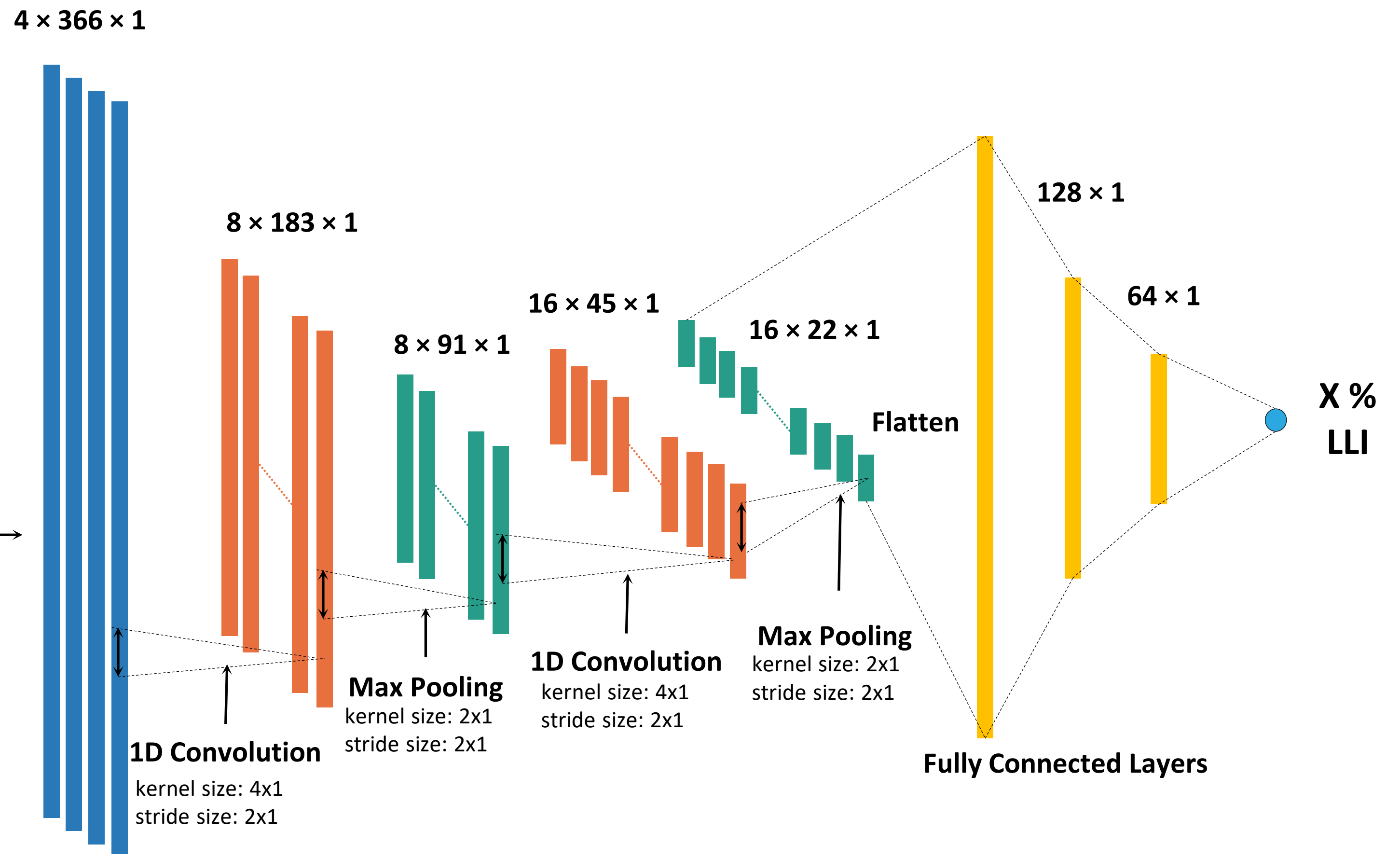
Generation of 26,000 synthetic datasets provides data that would take years to experimentally capture

Synthetic Data (SD-2)



Preprocessing

- IC curve
- Capacity fade
- 1st peak height
- 2nd peak height



1D Convolutional Neural Network to classify and quantify aging modes

Curve Fitting (CF) Prediction

Find M_n, a_n, and b_n → **Predict M_{12}, a_{12}, b_{12} @ 450 cycle** → **Prediction @ 450 cycle**

SRE_n: F=2*M_n*(0.5-1./(1+exp((a_n*cycle)^b_n)));

Data	Cycle	Capacity
1	0	0
2	25	12.78135
3	50	17.179
4	75	18.8563
5	100	19.61334
6	125	20.02941
7	175	20.49761
8	225	20.97087
9	275	21.36115
10	325	21.73976
11	375	22.03629
12	450	22.47789

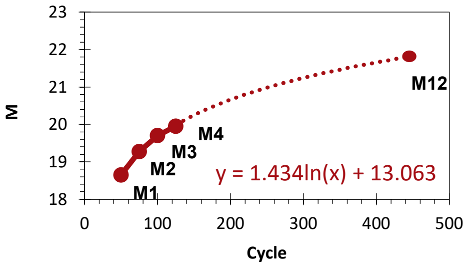
M_1, a_1, b_1 (3points)

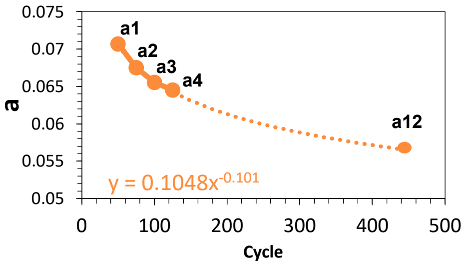
M_2, a_2, b_2 (4points)

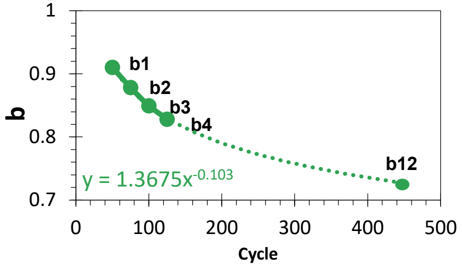
M_3, a_3, b_3 (5points)

M4, a4, b4 (6points)

M12, a12, b12
(To be estimated)







$SRE_{n=12}$: F=2*$M_{n=12}$*(0.5-1./(1+exp(($a_{n=12}$*cycle)^$b_{n=12}$)))