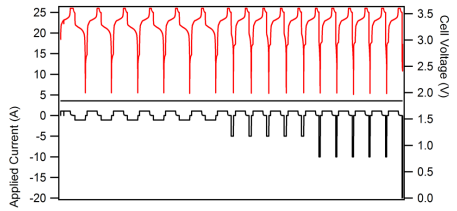
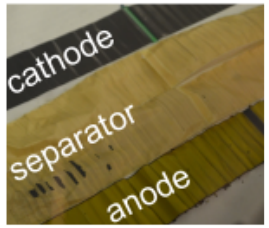
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

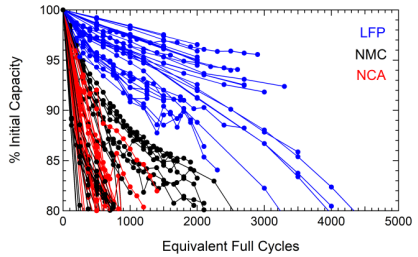
SAND2021-6330C

Sandia National Laboratories

BatteryArchive.org – Insights From a Public Repository for Visualization, Analysis, and Comparison of Battery Data Across Institutions







PRESENTED BY

Valerio De Angelis, Yuliya Preger

2021 ECS Spring

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Challenges in comparing battery performance data

Little raw data are publicly available and extracting information from figures is time consuming

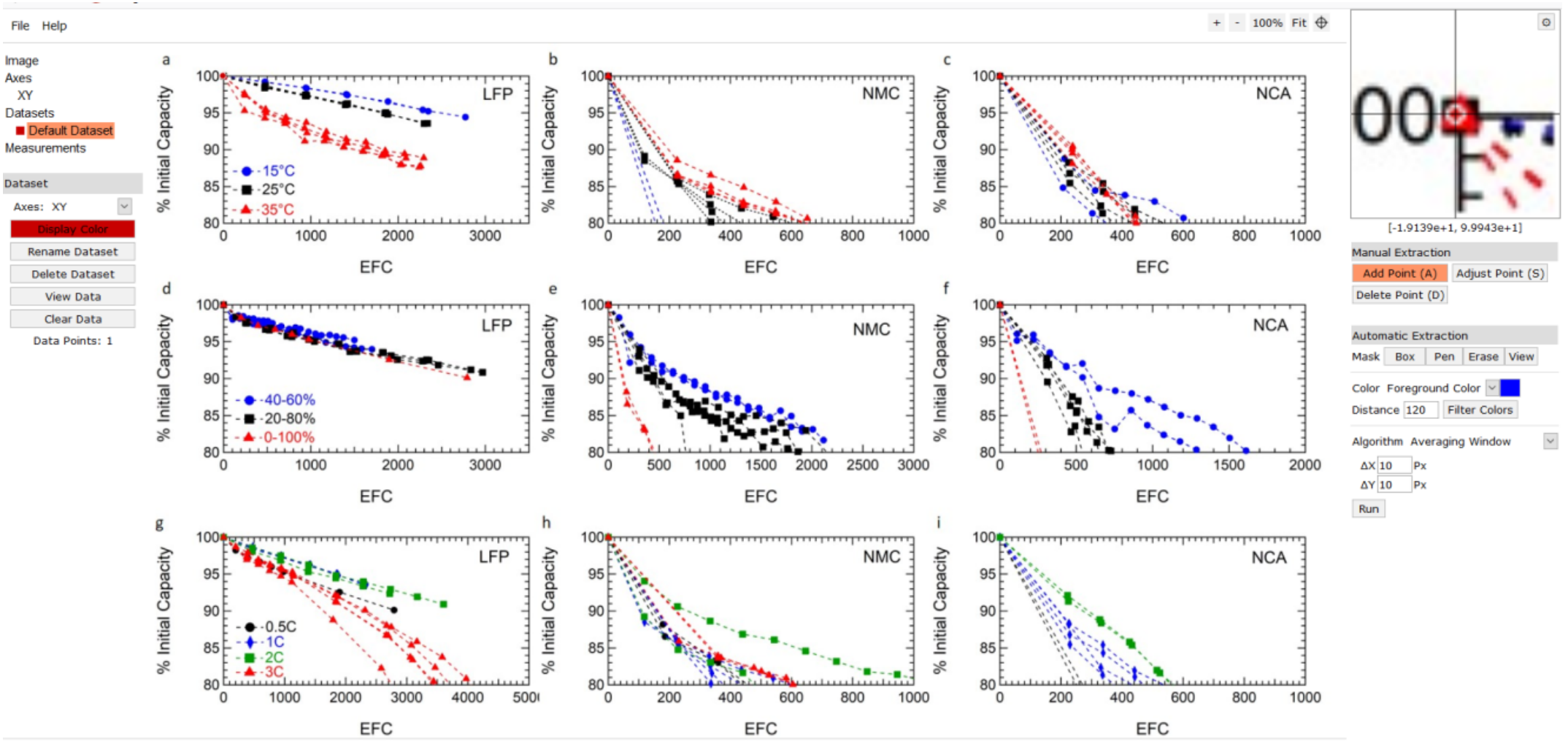
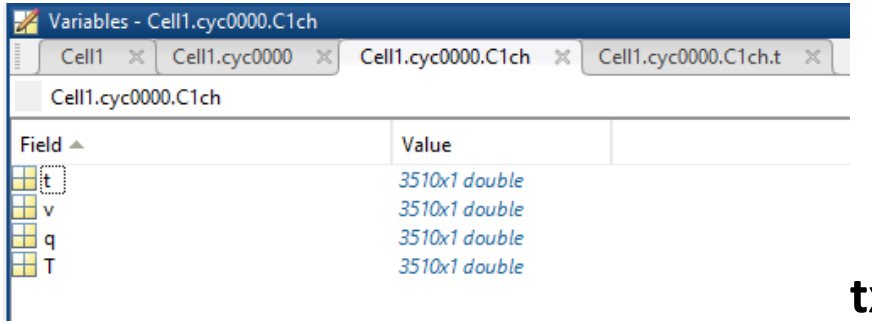


Figure: Preger et al. *J. Electrochem. Soc*. **2020**, *167*, 12; Data extraction: WebPlotDigitizer

Challenges in comparing battery performance data

Even when raw data are available, they are not standardized

Matlab



txt

File Edit Format View Help

Time	Status code	Status	category	Status	color	Pgm code	
0.000000	8	3	3	0	1	2	2
0.940317	8	3	3	0	1	2	2
1.954083	8	3	3	0	1	2	2
2.950567	8	3	3	0	1	2	2
3.945600	8	3	3	0	1	2	2
4.940633	8	3	3	0	1	2	2
5.937117	8	3	3	0	1	2	2
6.952317	8	3	3	0	1	2	2
7.948800	8	3	3	0	1	2	2
8.942400	8	3	3	0	1	2	2
9.938883	8	3	3	0	1	2	2
10.952633	8	3	3	0	1	2	2
11.944800	8	3	3	0	1	2	2
12.952800	8	3	3	0	1	2	2
13.940633	8	3	3	0	1	2	2

Excel

Cycle_Ind	Start_Tim	End_Time	Test_Time	Min_Curre	Max_Curr	Min_Volta
1	02:10.0	22:45.3	15645.31	-0.55	0.55	1.998
2	22:45.3	40:11.8	31091.75	-0.55	0.55	1.995
3	40:11.8	57:06.2	46506.17	-0.55	0.55	1.997
4	57:06.2	26:31.6	69871.59	-1.1	0.55	1.995
5	28:31.6	34:59.9	81179.92	-1.1	0.55	1.993
6	37:00.0	43:14.7	92474.68	-1.1	0.55	1.995
7	45:14.7	51:21.0	103761	-1.1	0.55	1.998
8	53:21.1	59:21.1	115041.1	-1.1	0.55	1.995
9	01:21.2	07:15.5	126315.5	-1.1	0.55	1.995
10	09:15.6	15:06.6	137586.6	-1.1	0.55	1.996
11	17:06.6	22:52.5	148852.5	-1.1	0.55	1.998
12	24:52.5	30:37.9	160117.9	-1.1	0.55	1.994

*columns different for different testers

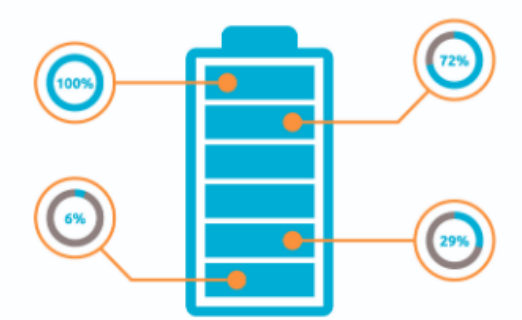
This makes it difficult to compare results from different studies and do larger-scale analyses

First multi-institution battery cycling database

BatteryArchive.org

A repository for easy visualization, analysis, and comparison of battery data across institutions

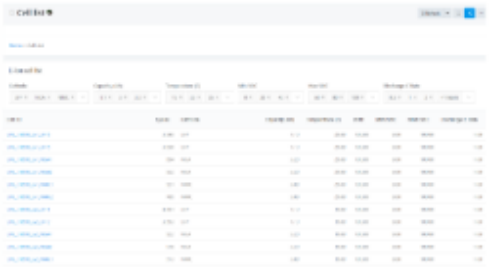
View Data



Features

1

Filter battery data



Query and filter for specific experimental conditions.

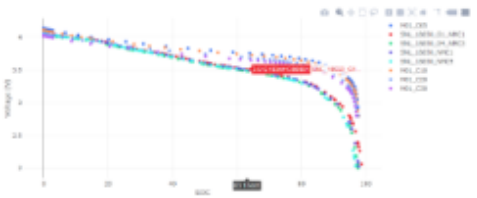
2

Visualize and compare data

Display battery data, including voltage curves and capacity fade.

3

Compare data with models



Apply performance and degradation models to battery data.

Site development background

Target: simple enough for casual users to compare battery performance, but still useful for advanced modeling users

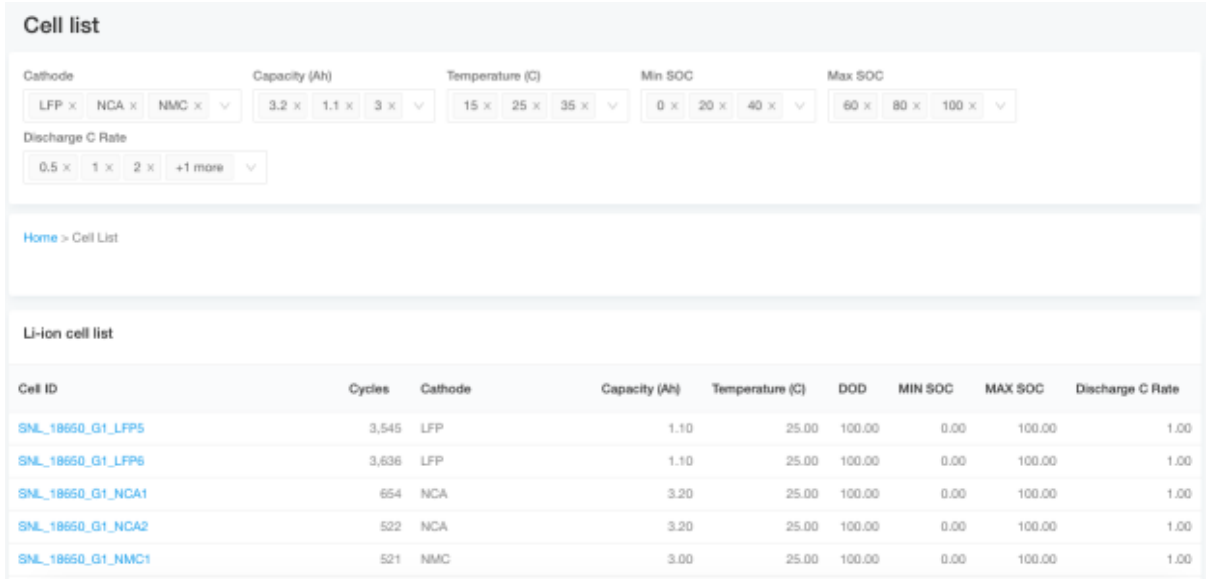
Key Features

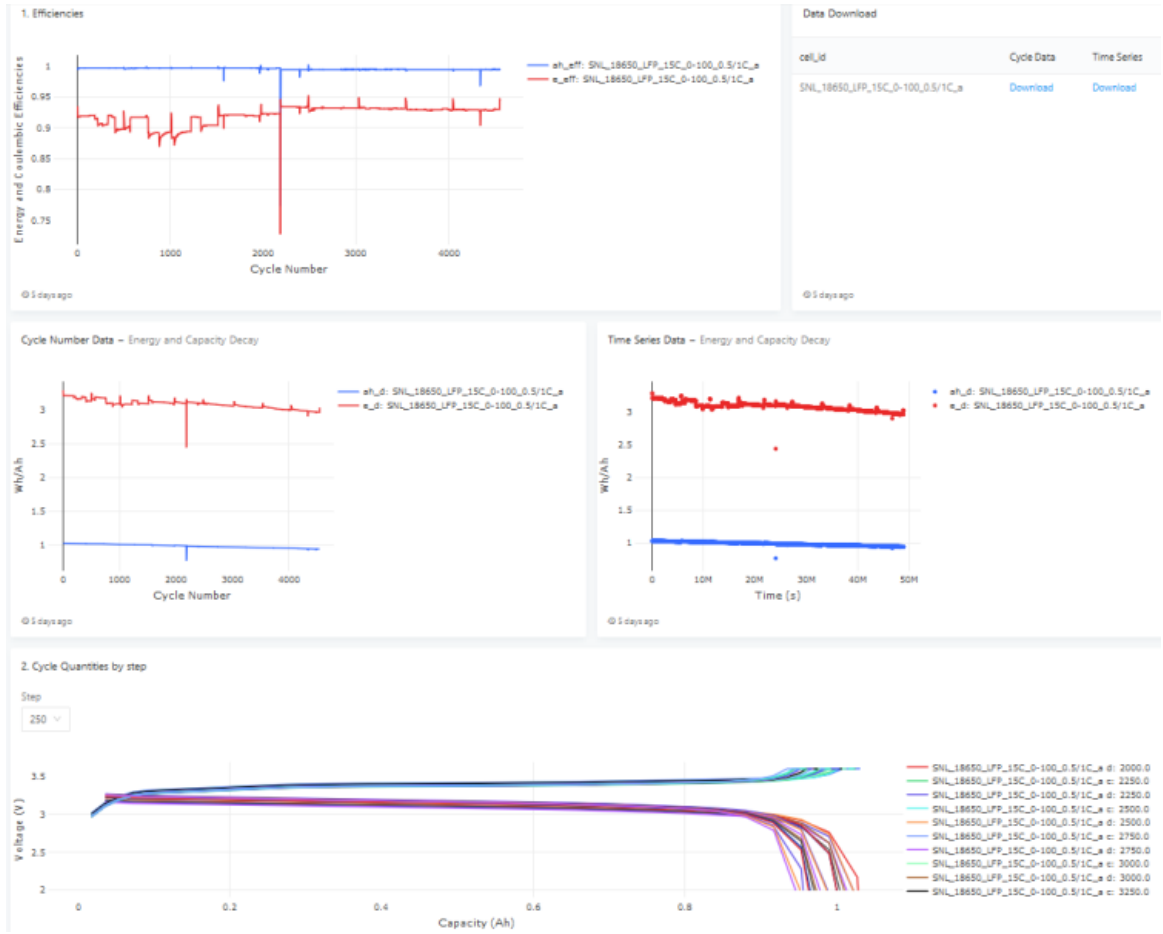
- Open access repository of battery data based on open-source software
- Import data from excel, battery testers, online repositories (Google Sheets), and databases
- Display the data in a responsive web interface and allow user download
- Offer modeling and comparison functionality on the web app and through Jupyter Notebook
- Adopt a standard minimum set of metadata, but plan for additional optional metadata
- Complementary to existing archival approaches (e.g. have data in zip folder on your group's website, data.mendeley, etc., but also combine here for comparison/manipulation)

First multi-institution battery cycling database

Search by metadata related to cell + cycling conditions

Efficiencies, capacity and energy decay, and voltage curves automatically plotted for selected cells





All data translated to same format to enable easy comparison

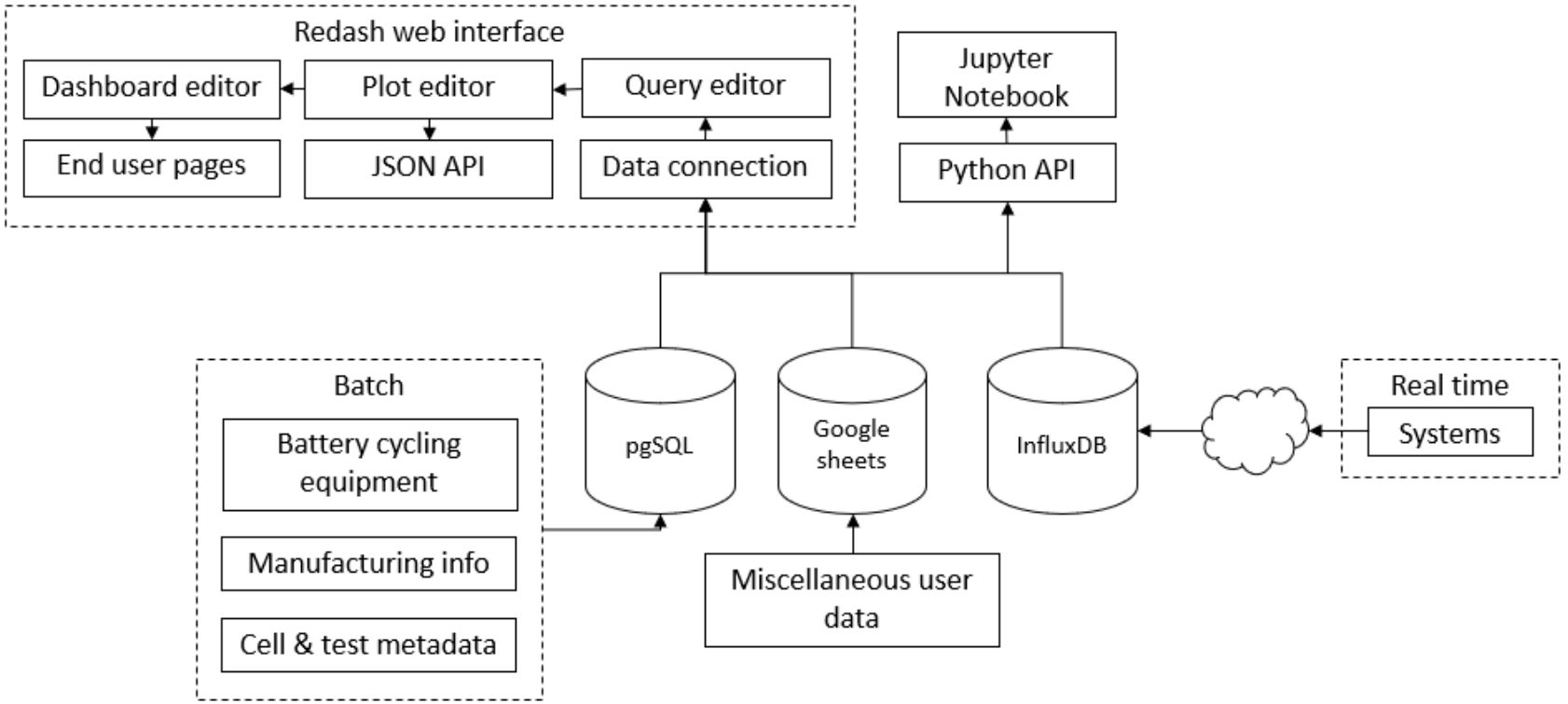
Current datasets

Over 2000 site users, many return visits, from over 40 countries, academia and industry

Release of open-source platform underpinning Battery Archive

- Public site is based on the Battery Lifecycle Framework (BLC) - an open-source platform that provides tools to visualize, analyze, and share battery data through the technology development cycle

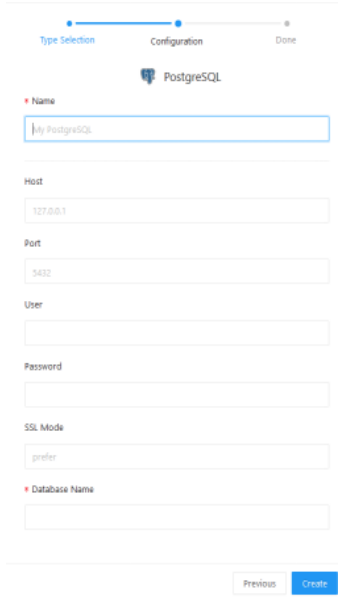
- BLC has four components: (1) data importers, (2) one or more databases, (3) a front-end for querying the data and creating visualizations, (4) an application programming interface to process the data



Preprint: https://ecsarxiv.org/h7c24; Github: https://github.com/battery-lcf/battery-archive-sandbox

Managing data connections and visualizations with Redash

1) Create new data source

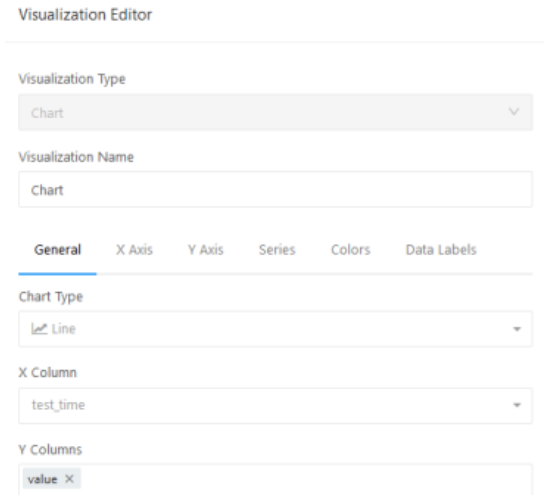


2) Build data queries

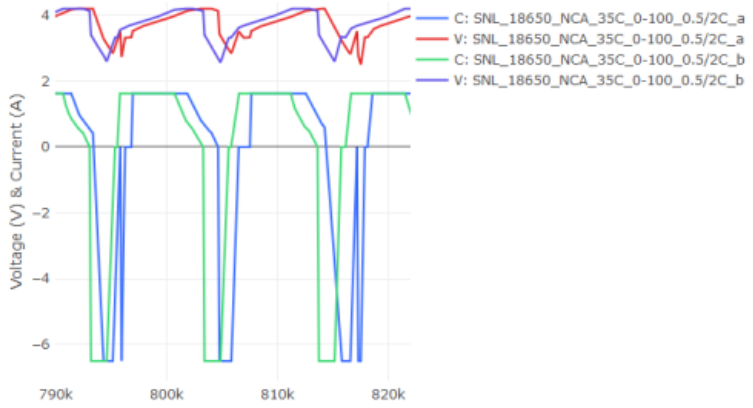
```
1  SELECT
2     key || ': ' || r.cell_id as series,
3     r.cycle_index,
4     r.test_time,
5     value
6  FROM (SELECT timeseries_data.cell_id, cycle_index, test_time, json_build_object('V', v, 'C', i) AS line
7  FROM timeseries_data TABLESAMPLE BERNOULLI ({{sample_vc}})
8  where cell_id IN ({{cell_id}}) and
9      cycle_index>= {{min}} and
10     cycle_index <= {{max}}) as r
11 JOIN LATERAL json_each_text(r.line) ON (key ~ '[V,C]')
12 order by r.cell_id,r.test_time,  key
13
```

{{ }} Save* ▶ Execute

3) Build visualizations



4) Manipulate plots



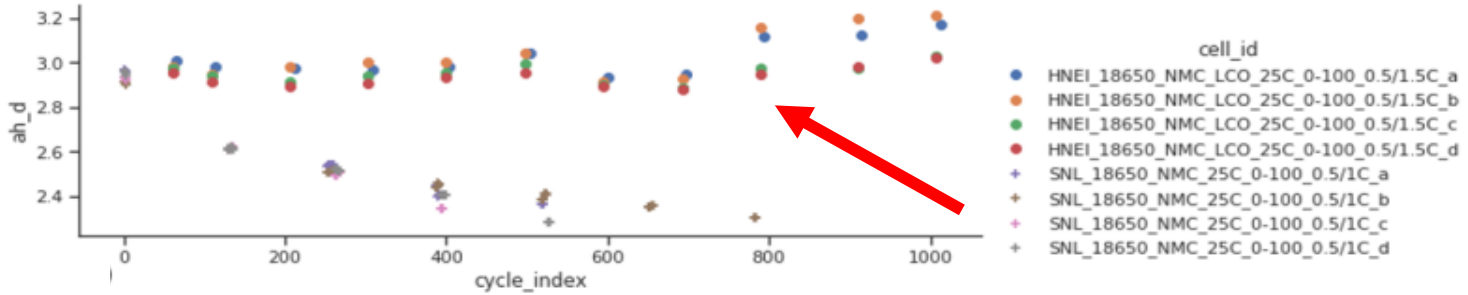
Linking to Jupyter Notebook for data processing

Case study: pulling out capacity fade based on reference performance tests (RPTs) in different studies

Linking to Jupyter Notebook for data processing

Case study: pulling out capacity fade based on reference performance tests (RPTs) in different studies

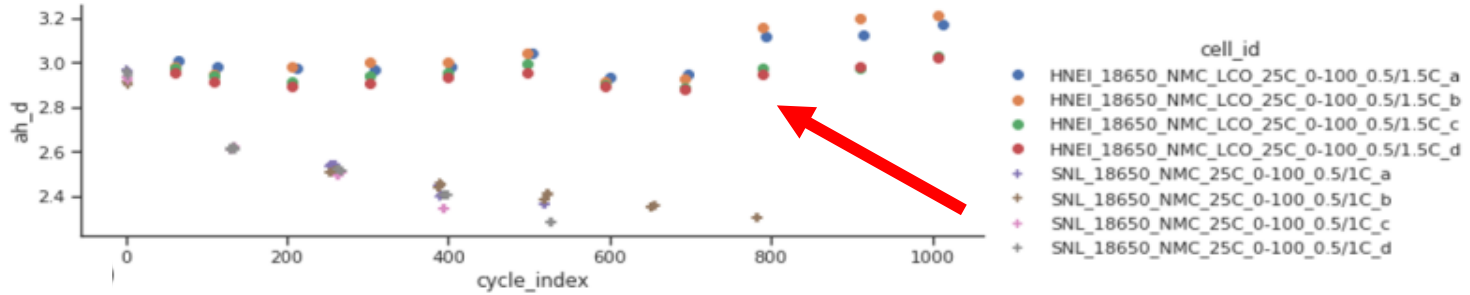
Discharge capacity overestimated in one dataset uploaded ‘as is’ to Battery Archive



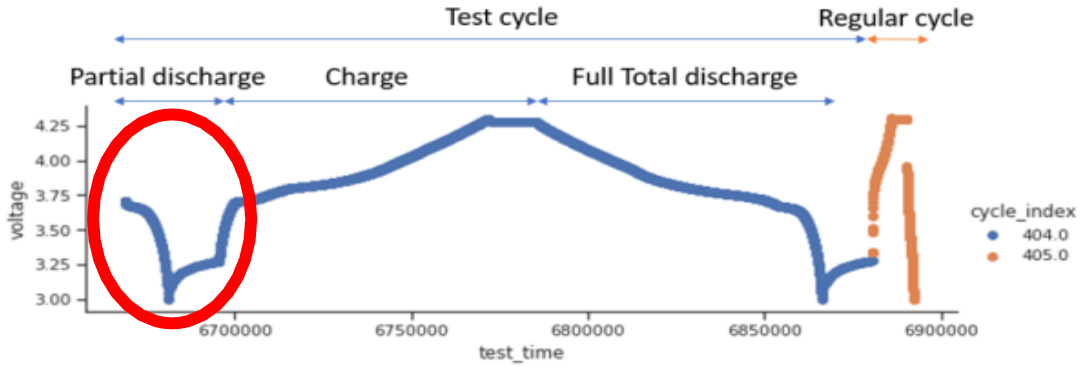
Linking to Jupyter Notebook for data processing

Case study: pulling out capacity fade based on reference performance tests (RPTs) in different studies

Discharge capacity overestimated in one dataset uploaded ‘as is’ to Battery Archive



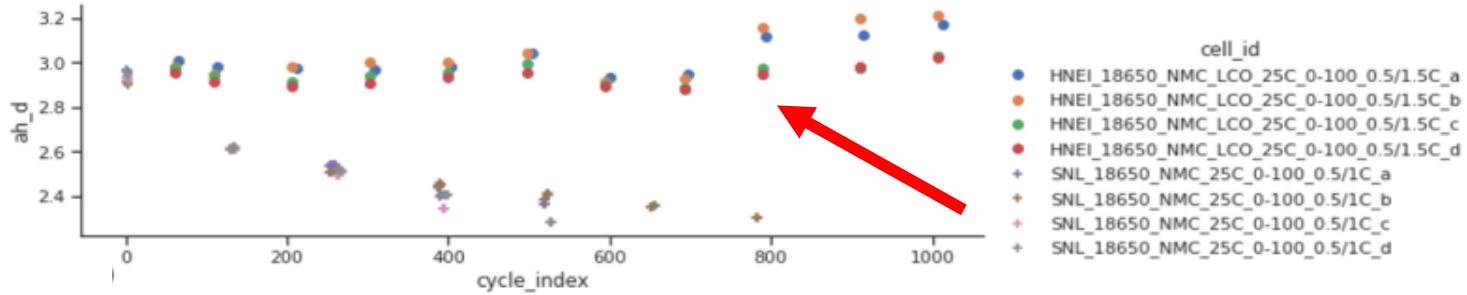
RPT of dataset included a partial discharge



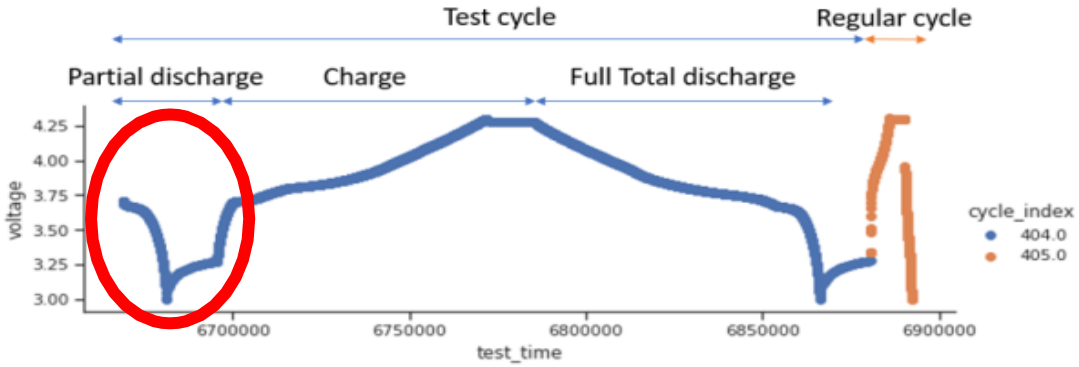
Linking to Jupyter Notebook for data processing

Case study: pulling out capacity fade based on reference performance tests (RPTs) in different studies

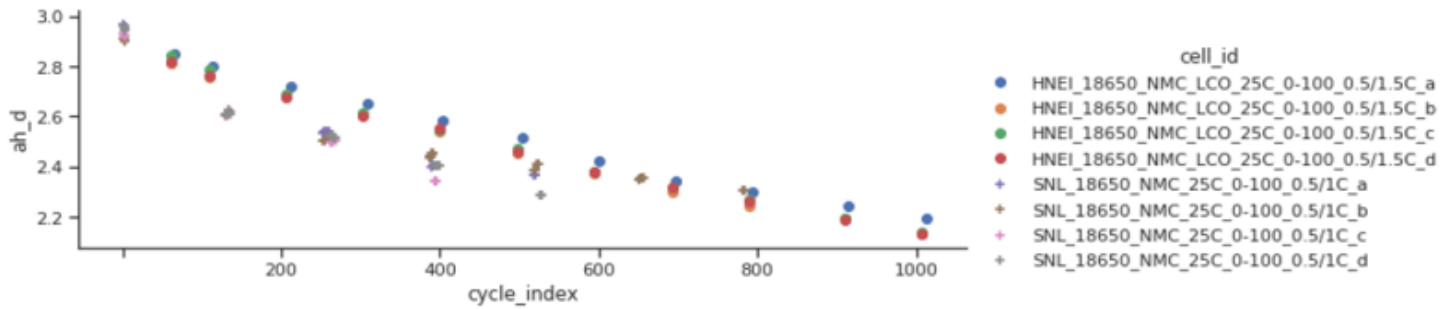
Discharge capacity overestimated in one dataset uploaded ‘as is’ to Battery Archive



RPT of dataset included a partial discharge in addition to full cycle

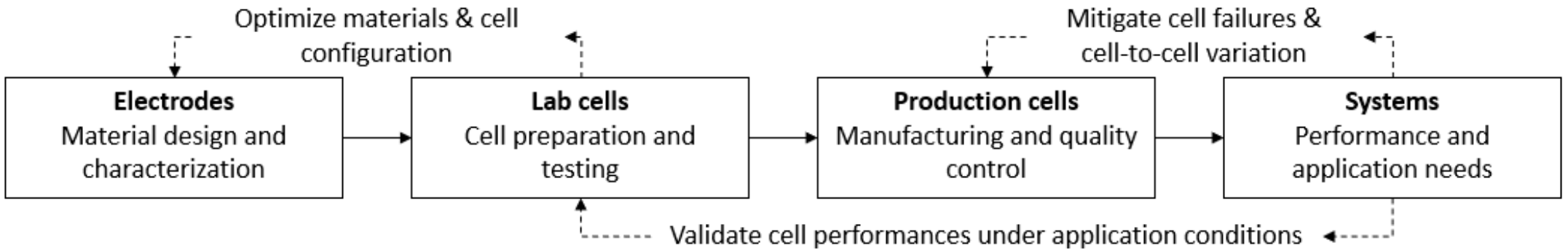


Partial discharge filtered out and correct discharge capacity reported



Beyond single cell degradation – linking data across a battery's lifecycle

- BLC permits linking of multiple data sources, queries, and dashboards to track data from materials to systems
- Linking data across steps key to accelerating the battery development process



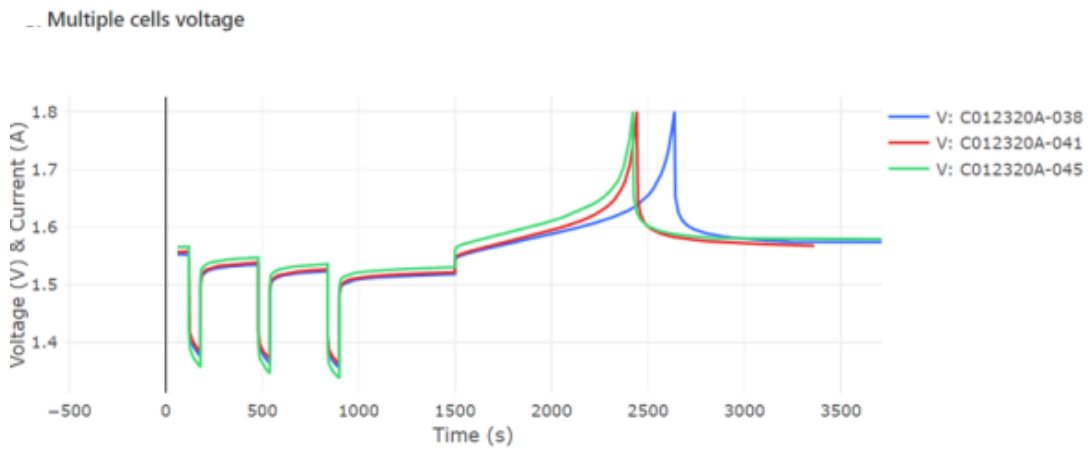
Linking data from multiple data streams

Case study: linking data from lab cells to fielded systems

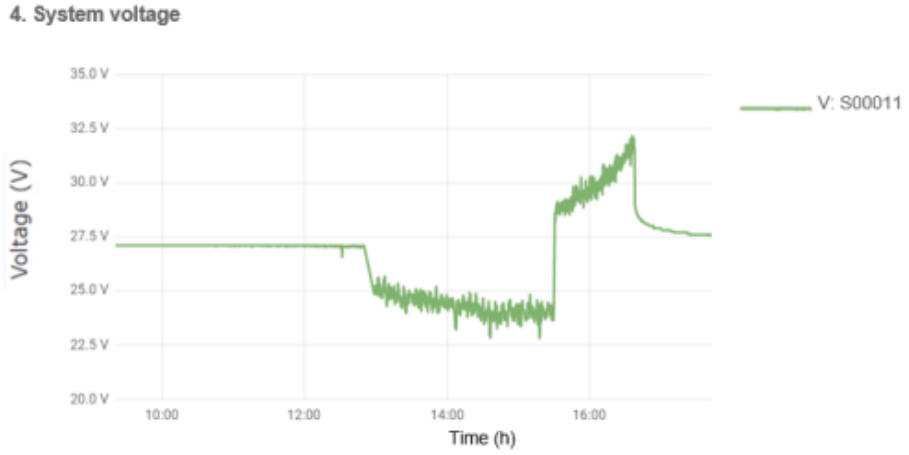
1) Manufacturing data for 3 cells imported into pgSQL database from a FileMaker database

cell_id	study	Anode Batch ID	Cathode Batch ID	Anode Weight (kg)	Cathode Weight (kg)
C012320A-038	3-pulse+1-cycle-Battery 1-SBT	A - 011520	C - 110819	0.98	1.80
C012320A-041	3-pulse+1-cycle-Battery 1-SBT	A - 011520	C - 110819	0.98	1.80
C012320A-045	3-pulse+1-cycle-Battery 1-SBT	A - 012120	C - 110819	0.96	1.86

2) Cell-level pulse testing data imported from Arbin testers (MS Access) and a PEC tester (Oracle Database)



3) InfluxDB data from a fielded system



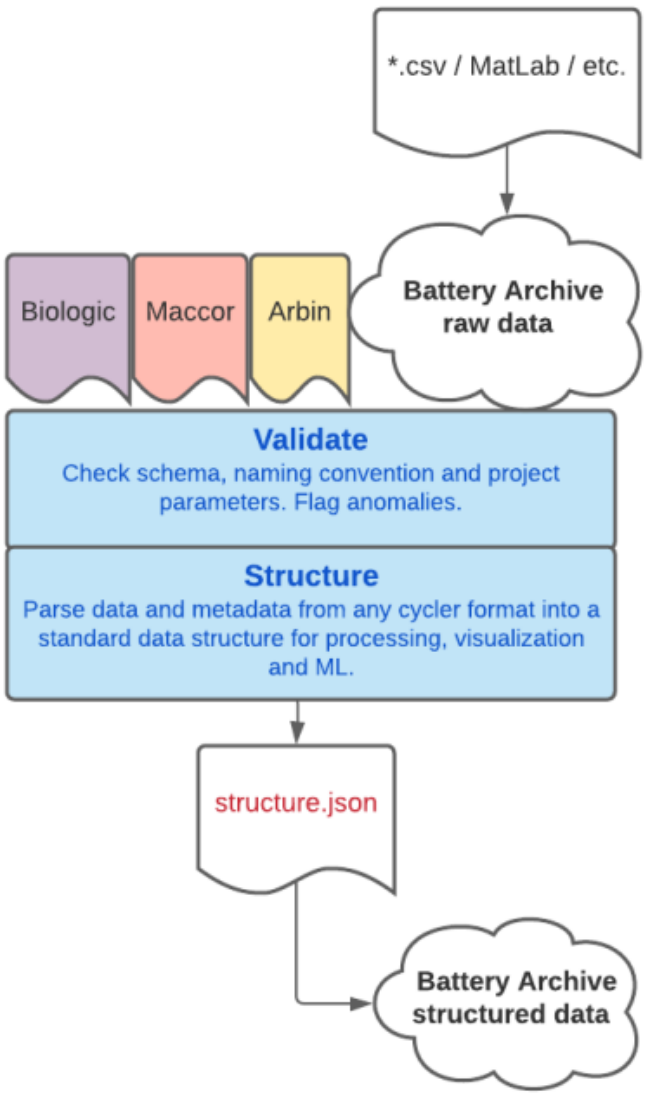
Beyond Battery Archive – linking to BEEP to automate data analysis

- BEEP was developed by TRI, Stanford, and MIT as an open-source software package for **B**attery **E**valuation and **E**arly **P**rediction
  - Includes tools for many battery testers
  - Organizes raw cycling data and featurizes them to serve as input for analytics, including machine learning
    - For example, extract charge, discharge, rest segments from any cycling data

- Linking Battery Archive and BEEP will enable:
  - Importing all the data formats that BEEP supports
  - Visualizing private data and public data from Battery Archive on a single dashboard
  - Developing models using data from multiple sources

Publication: https://doi.org/10.1016/j.softx.2020.100506
Github: https://github.com/TRI-AMDD/beep

Beyond Battery Archive – linking to BEEP to automate data analysis



BEEP

Public and private data are treated the same way

The results of processing the data using BEEP can be visualized. For example, only plot charge segments.

Acknowledgments

Funded by the U.S. Department of Energy, Office of Electricity, Energy Storage program. Dr. Imre Gyuk, Program Director.

Thank you to the CUNY Energy Institute for supporting the site.

Core site development: Valerio De Angelis (SNL)
Early framework testers at CUNY Energy Institute: Sanjoy Banerjee, Jinchao Huang, Andreas Savva, Gautam Yadav
Front end: Sam Roberts-Baca (SNL)
Database: Mark Spoonamore (SNL)
Early site feedback: Matthieu Dubarry (HNEI)

BEEP integration: Patrick Herring

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Public site: www.BatteryArchive.org
Battery Lifecycle Framework Preprint: https://ecsarxiv.org/h7c24
Github: https://github.com/battery-lcf/battery-archive-sandbox

For questions about this presentation: ypreger@sandia.gov