CACTOS

Energy,
but better

21.11.2025
Otaniemi, Espoo

Aalto University
School of Science

SoC and SoH Estimation of LFP Cells Using Operational Data

Master’s Thesis presentation by Olli Ruokojoki

Outline

1. Why Model SoC?
2. Laboratory Testing vs Operational Data
3. Model Architecture: EKF + ECM
4. Model Creation
5. SoC Model Results
6. SoH Estimation
7. SoH Model Results
8. Key Takeaways

Why estimate the State of Charge (SoC)?

- Optimal battery operations
- Bad estimates lead to:
  - → Unexpected SoC behavior
  - → Inefficient operation
  - → Decreased revenue
  - → Energy imbalance
- **Can’t be measured directly**
  - Inferred from: U, I, T

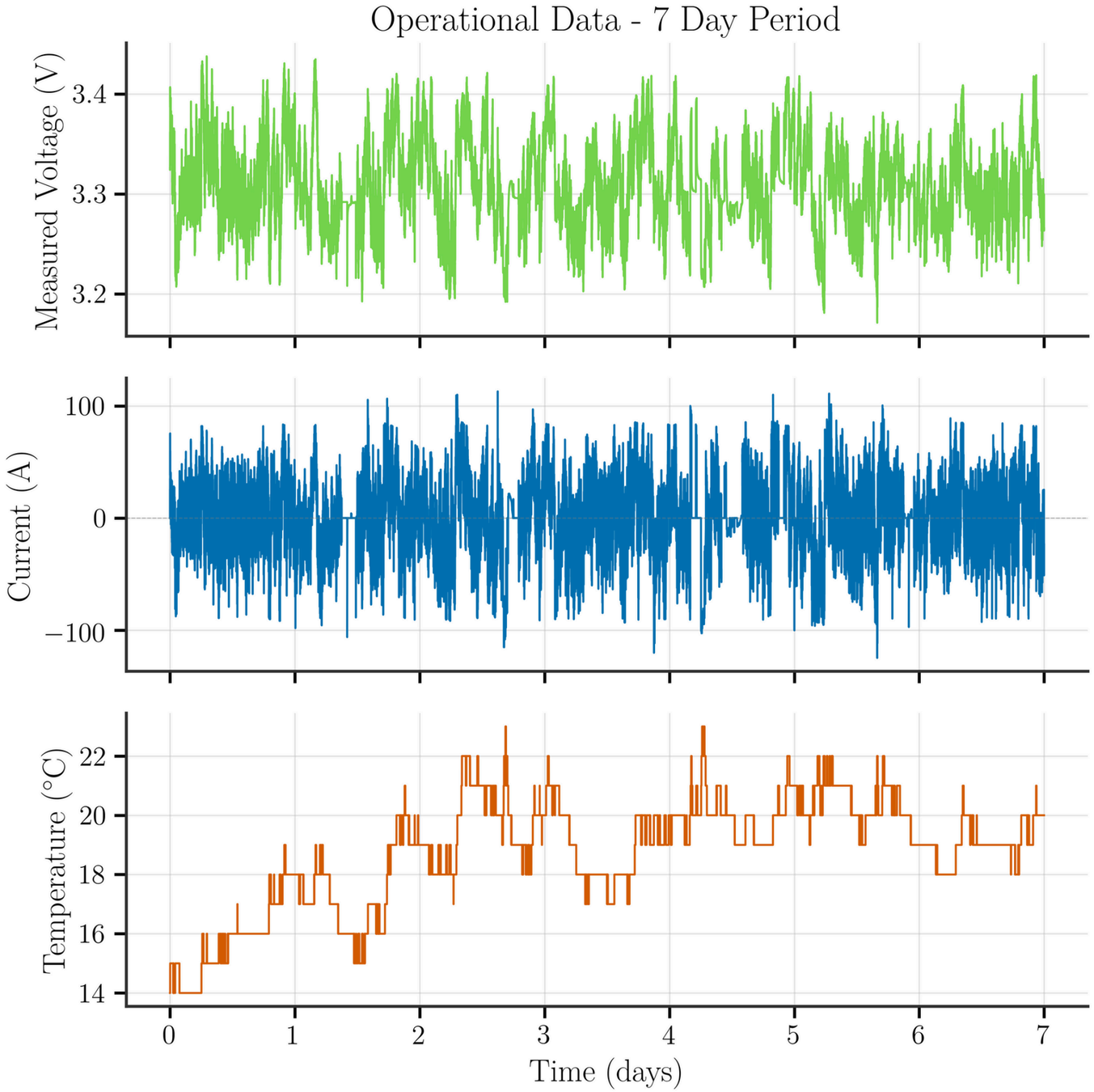
Model Creation:
The Standard Way

- Precise measurement equipment
- Controlled temperature chambers
- Weeks or months of testing
- Highly controlled environments
- **Not representative of real use**

Image generated by Grok Imagine

What If We Used Operational Data Instead?

- **Actually fit for application**
- No extra equipment
- **Difficult**
  - Chaotic and messy
  - No controlled test sequences
  - No ground truth
  - Novelty

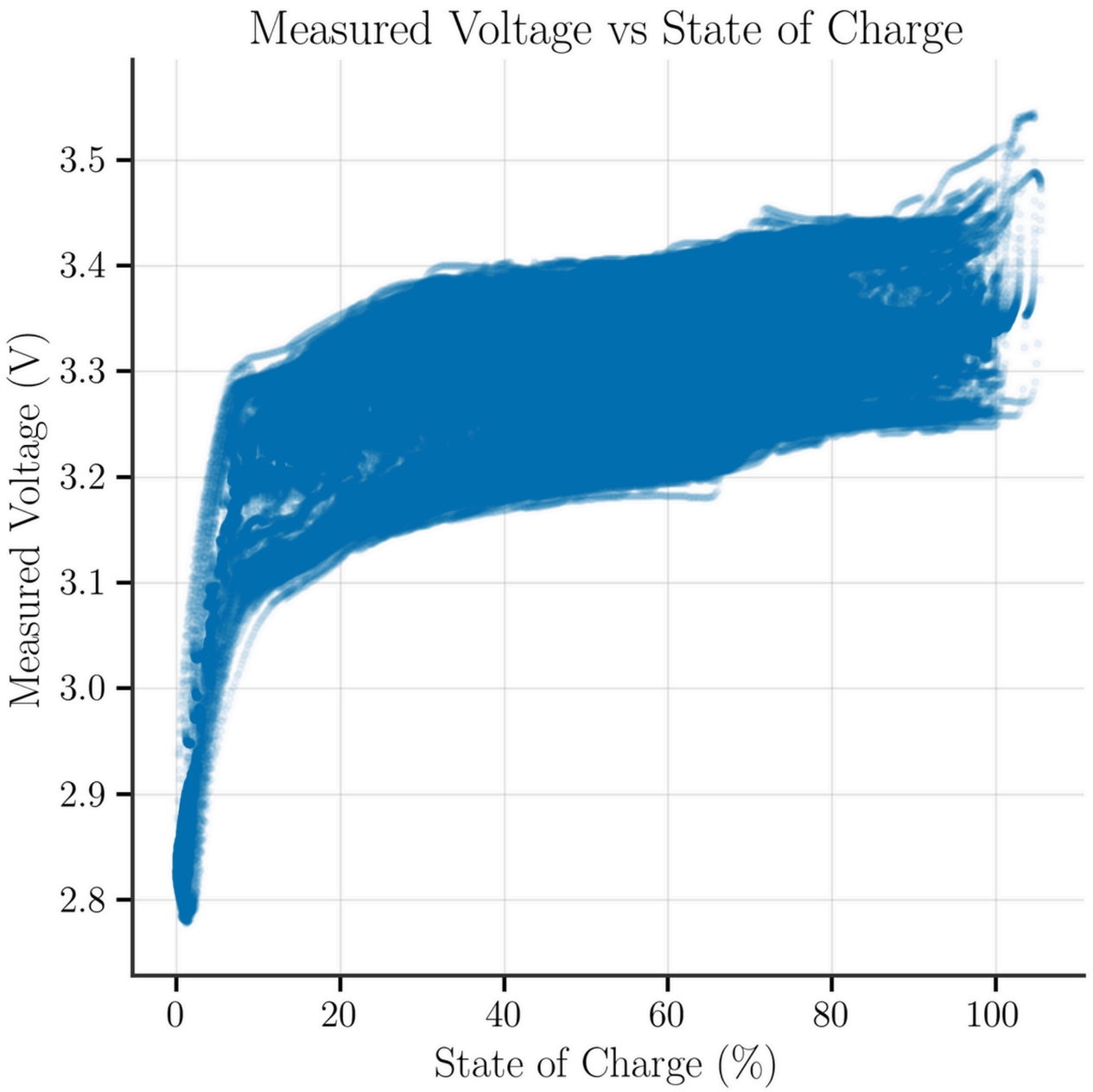


How does an SoC Model work?

1. Current integration (*drift*)
2. Voltage measurements (*noise*)

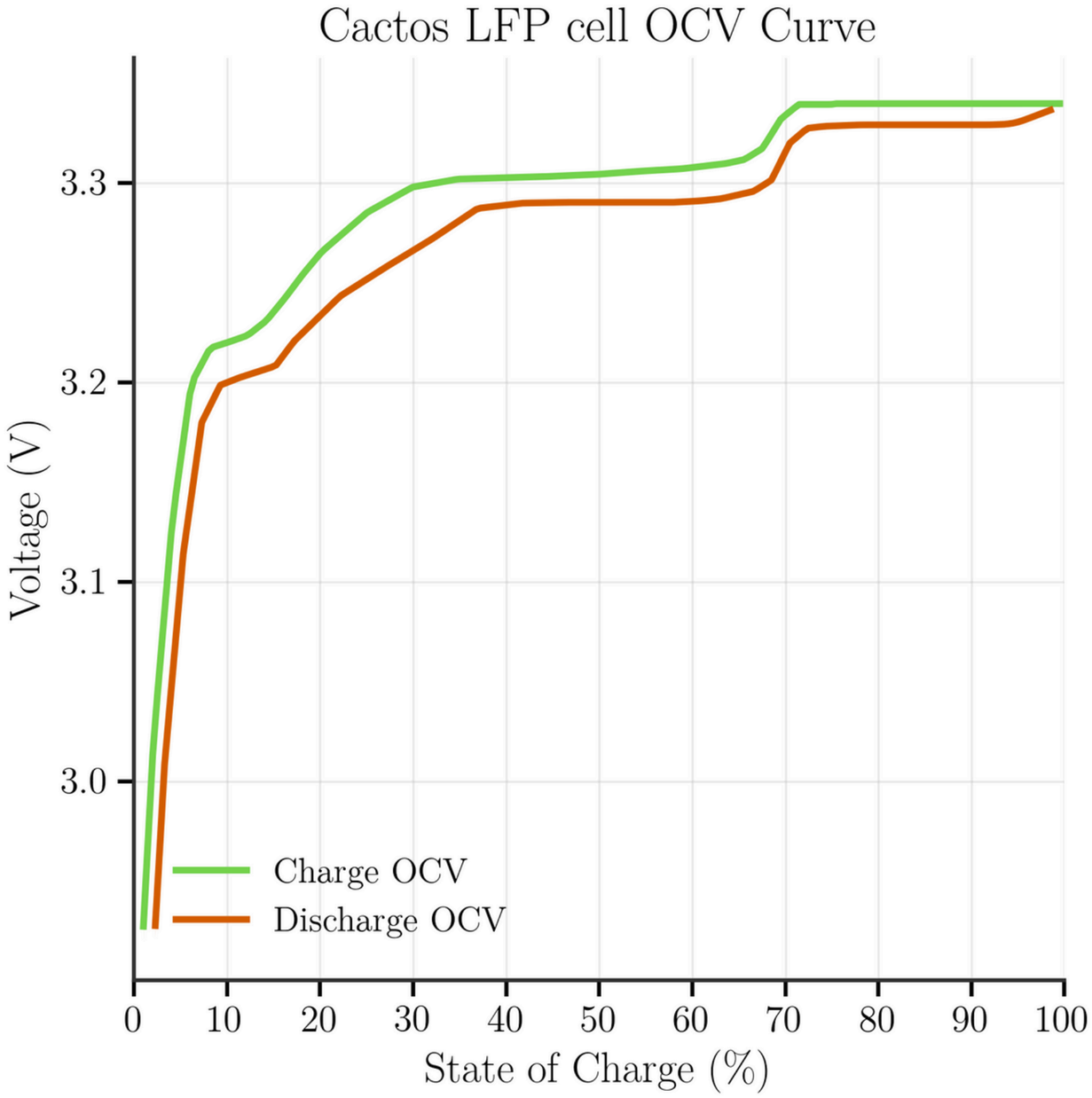
Sensor fusion:

- **Extended Kalman Filter (*balance*)**

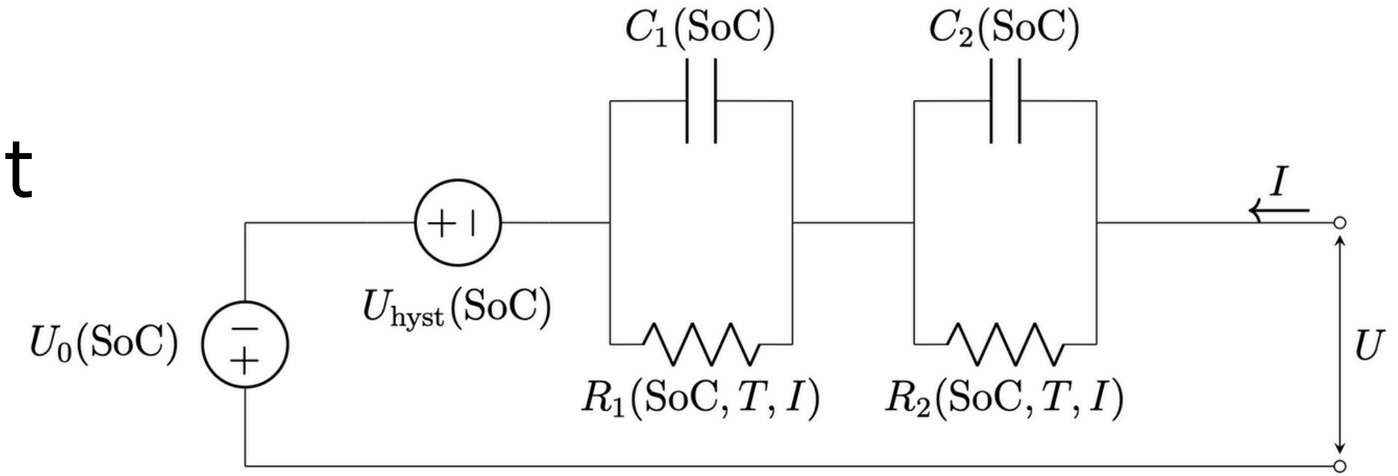


The Extended Kalman Filter (EKF)

1. **Prediction: Process model**
   - Current integration
2. **Update: Measurement model**
   - Voltage → SoC
   - **Requires** a battery model
3. **Correction**



The Equivalent Circuit Model (ECM)



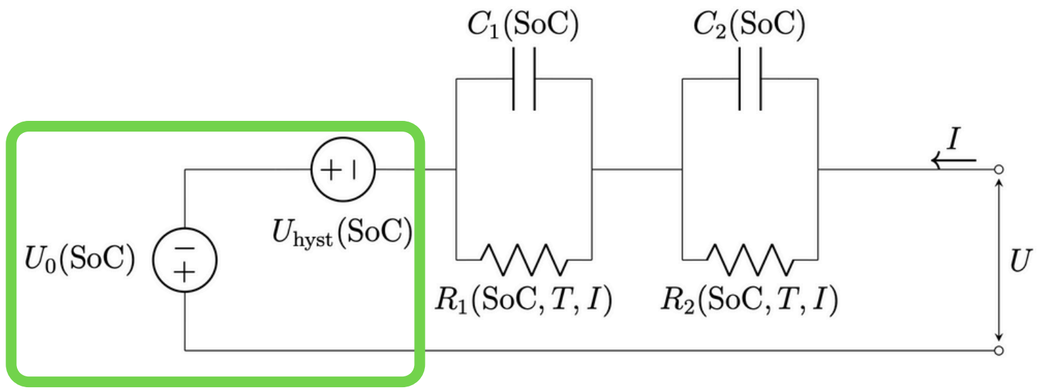
ECM parameter dependencies

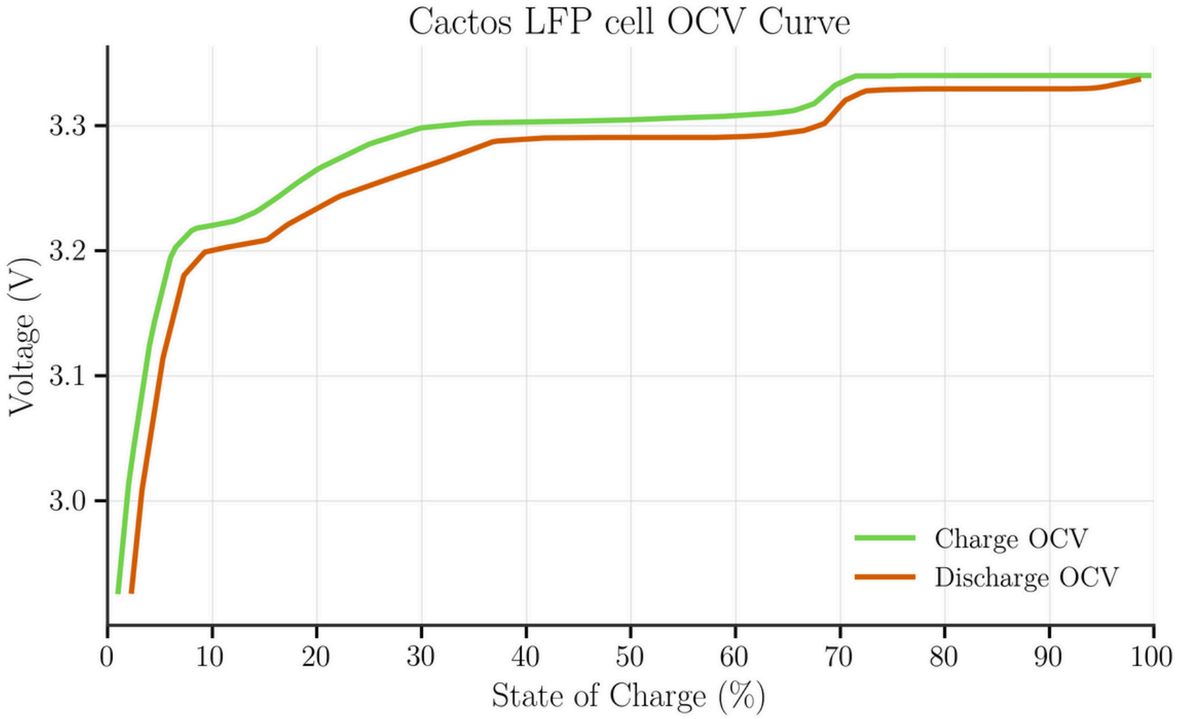
	SoC	Temperature T	Current rate $\lvert I \rvert$	Current sign	SoH / aging
U_0	x	[88]			[50, 89]
R_0	[55, 70, 88, 90, 91]	[70, 90–92]	[91, 92]	[39, 70, 91, 93]	[50, 53, 55]
R_1	[50, 55, 70, 88, 90, 91]	[50, 70, 90–92]	[91, 92]	[70, 91, 93]	[53, 55, 88, 94]
C_1	[88, 90, 91]	[90–92]	[91, 92]	[91]	[94]
R_2	[50, 55, 70, 88, 90, 91]	[50, 70, 90–92]	[91, 92]	[70, 91, 93]	[53, 55, 94]
C_2	[88, 90, 91]	[90–92]		[91]	[94]
U_{hyst}	x	[39, 70, 81, 90]			[81, 95, 96]

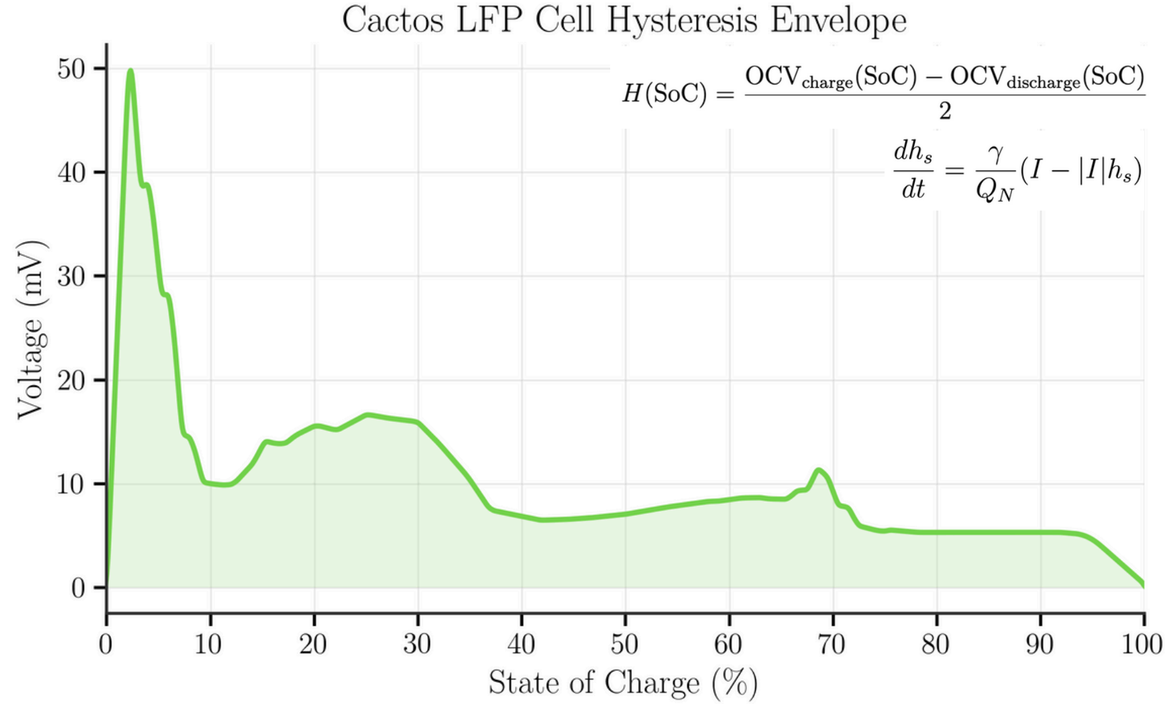
Table 1 of the Master's Thesis

The Open-Circuit Voltage (OCV) curve

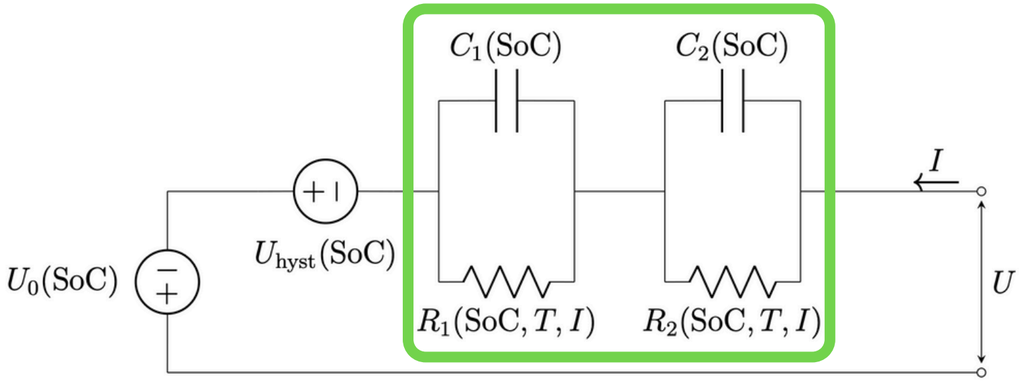
- The “rest” voltage
- **Maps** OCV to SoC
- Most important component of the ECM
- **Tested** with a 0–100–0% SoC run







Iterative Model Creation



START: INITIAL SOC ESTIMATES

PRE-PROCESSING: TEST SECTIONS

FIT 1: TIME-DEPENDENT PARAMETERS

FIT 2: RESISTANCE TEMPERATURE AND CURRENT DEPENDENCIES

FIT RESIDUAL: RESISTANCE SOC DEPENDENCY

CACTOS FIT 1

$$U_k = \alpha_k U_k^- + (1-\alpha_k) U_{\mathrm{ss},k}, \quad k \in \{1,2\}$$

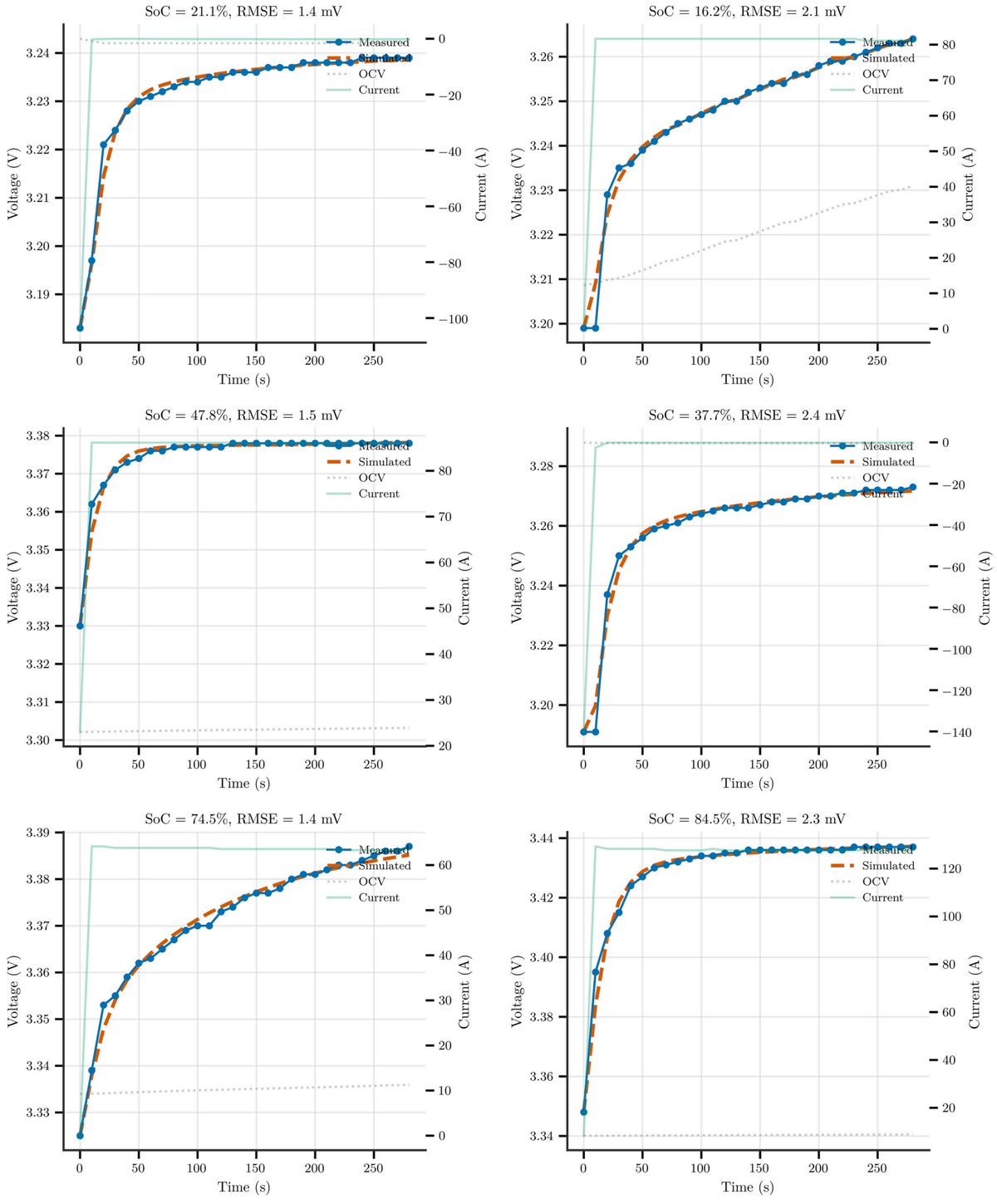
$$\alpha_k = \exp\left(-\frac{\Delta t}{\tau_k}\right)$$

Time dependent parameters

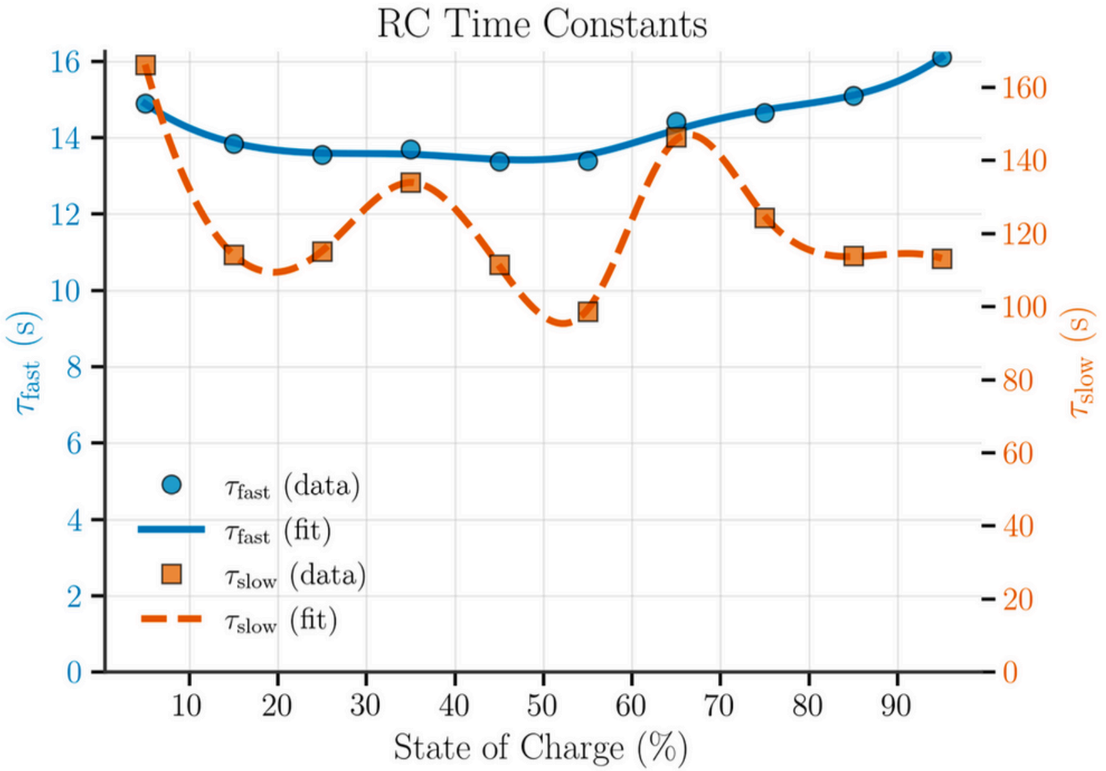
- Assume: exponential decay
- Find: stable current step sections
  - Sharp current change
  - Stable current prior and after
- Fit: time constants and R1/R2
- Resistance values not needed

$$\frac{dU_k}{dt} = -\frac{U_k}{\tau_k} + \frac{R_k}{\tau_k} I(t), \qquad k \in \{1,2\}$$

PRE-PROCESSING: TEST SECTIONS → **FIT 1**: TIME-DEPENDENT PARAMETERS



Time dependent parameters

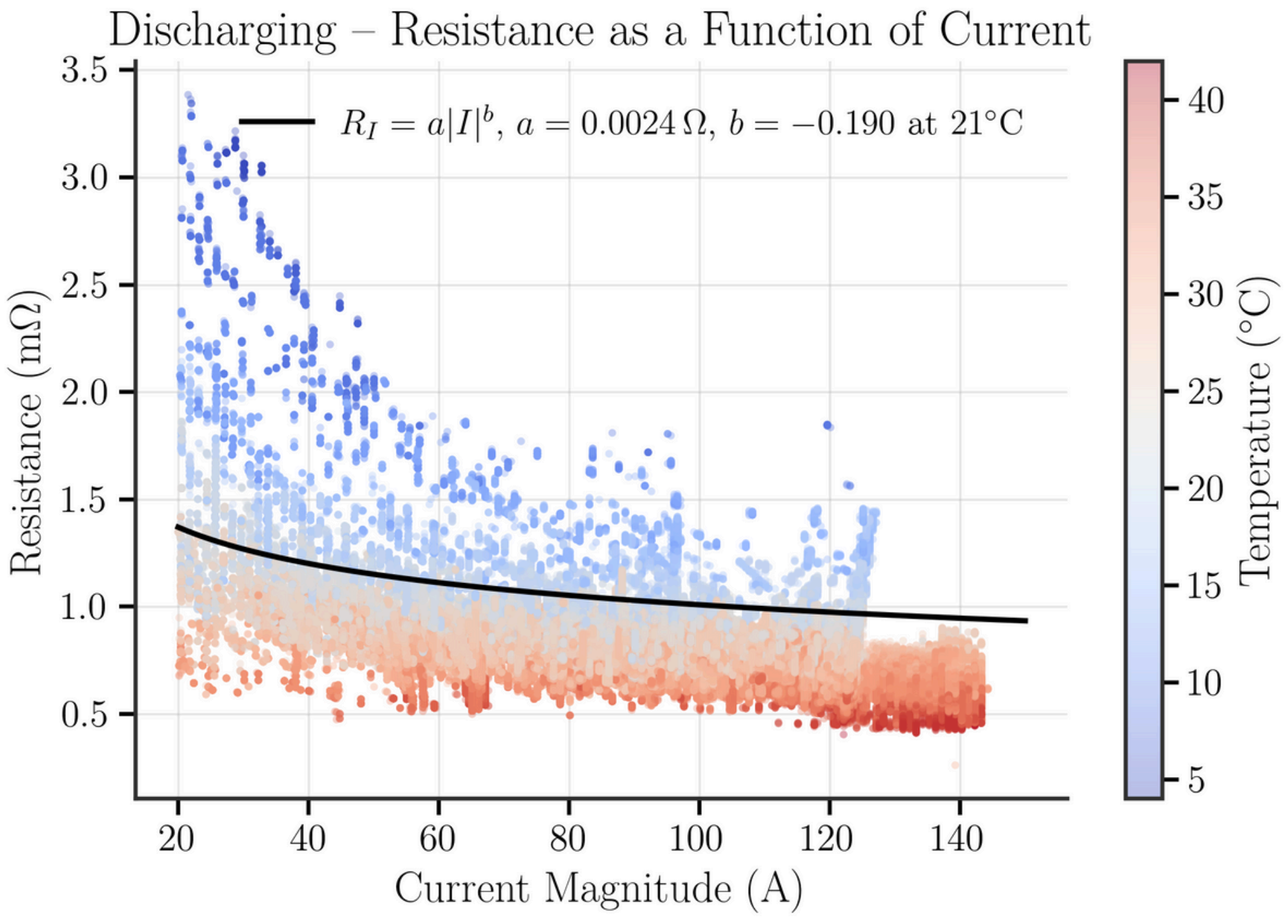


PRE-PROCESSING: TEST SECTIONS → **FIT 1**: TIME-DEPENDENT PARAMETERS

$$R_{\text{tot}}(T, SoC, I) = R_I(|I|, \sigma) \cdot f_{\text{SoC}}(SoC, \sigma) \cdot f_T(T, \sigma)$$

Resistance parameters

- Assume: separable dependencies
- Find: stable sections, all dynamics fully evolved
- Fit:
  - Current and temperature dependency parameters
  - SoC from residual (spline)

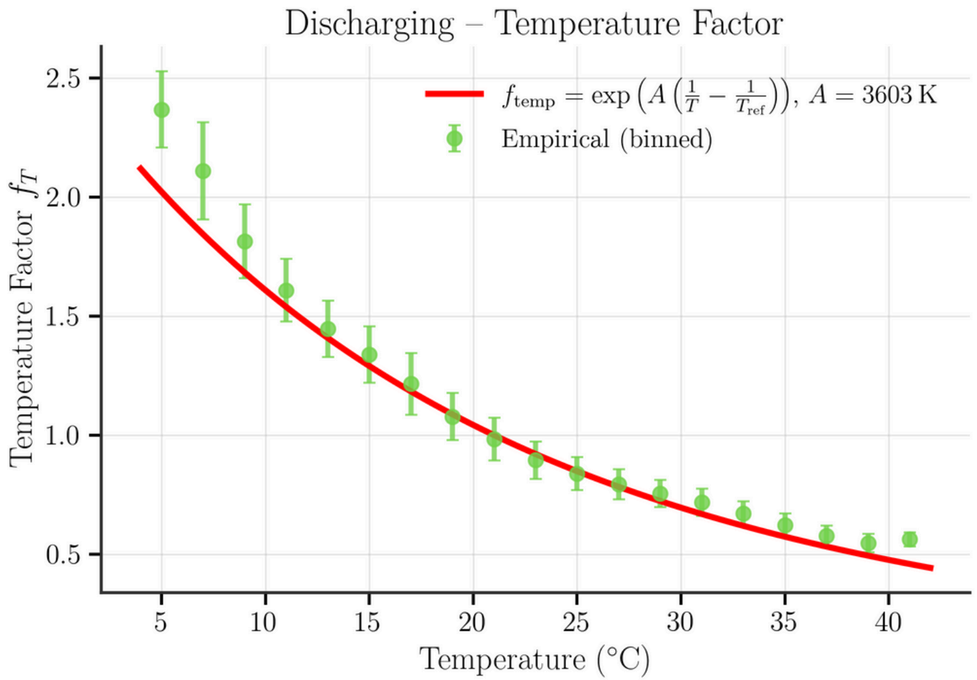


$$R_I(|I|, \sigma) = \begin{cases} a_\sigma |I|^{b_\sigma}, & |I| \geq I_0 \\ a_\sigma I_0^{b_\sigma} + a_\sigma b_\sigma I_0^{b_\sigma - 1}(|I| - I_0), & |I| < I_0 \end{cases}$$

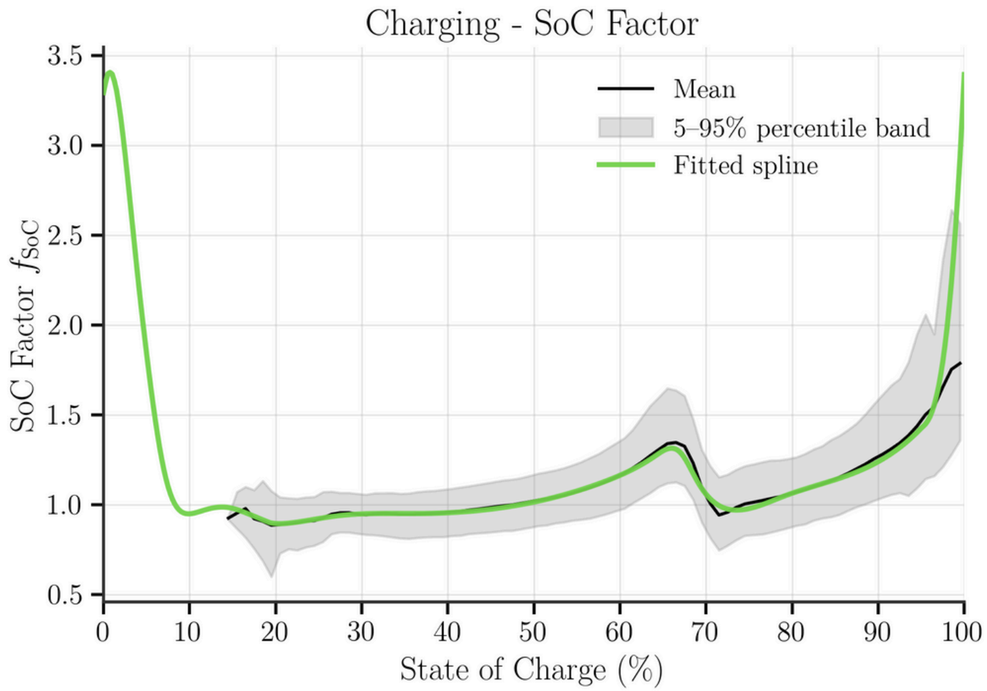
FIT 2: RESISTANCE TEMPERATURE AND CURRENT DEPENDENCIES → **FIT RESIDUAL**: SOC DEPENDENCY

$$R_{\text{tot}}(T, SoC, I) = R_I(|I|, \sigma) \cdot f_{\text{SoC}}(SoC, \sigma) \cdot f_T(T, \sigma)$$

Resistance parameters



$$f_T(T, \sigma) = \exp\left(A_\sigma\left(\frac{1}{T} - \frac{1}{T_{\text{ref}}}\right)\right)$$



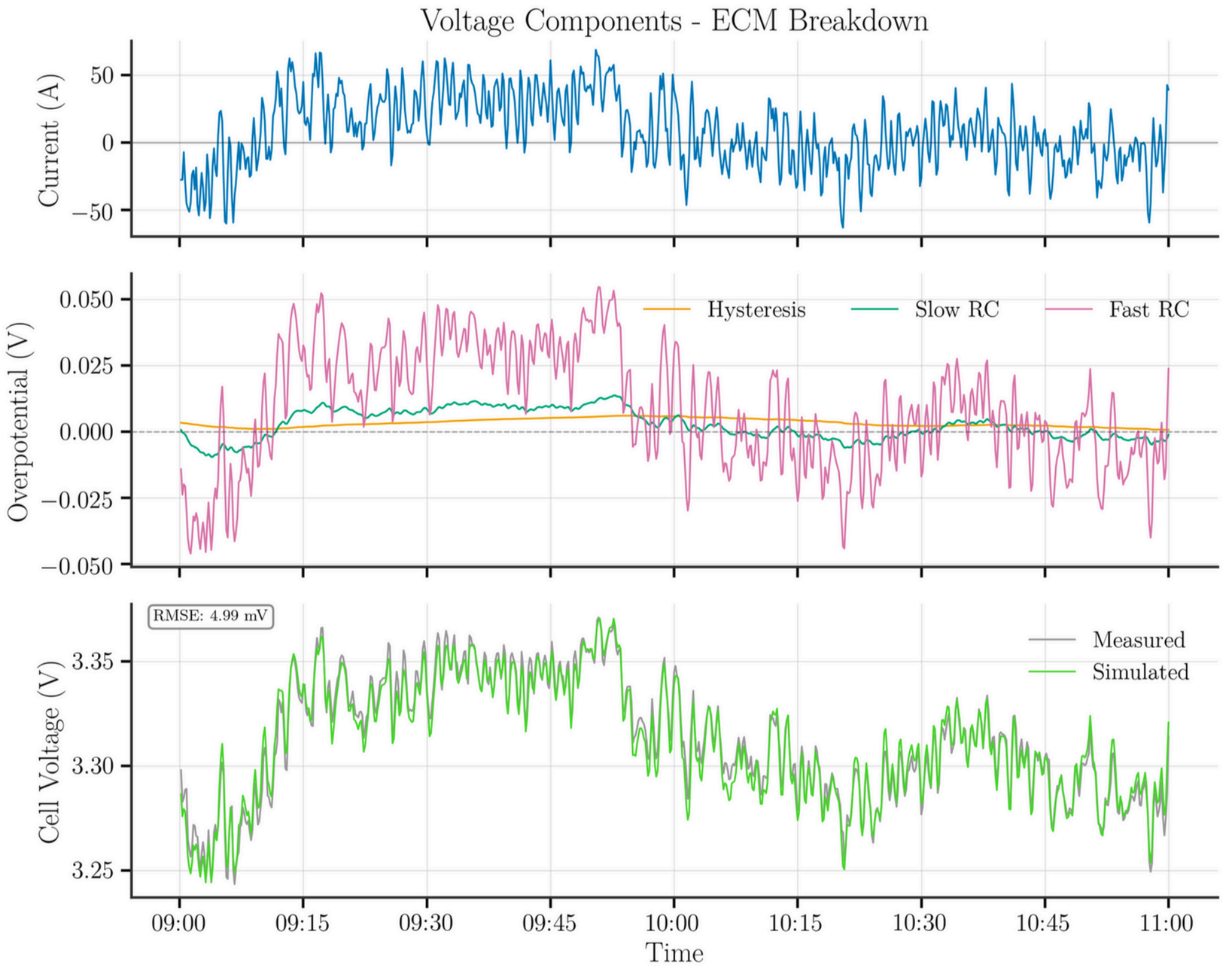
$$f_{\text{SoC}}(\text{SoC}, \sigma) = \frac{R_m}{R_{\text{pred}}(I, T, \sigma)}$$

FIT 2: RESISTANCE TEMPERATURE AND CURRENT DEPENDENCIES

FIT RESIDUAL: SOC DEPENDENCY

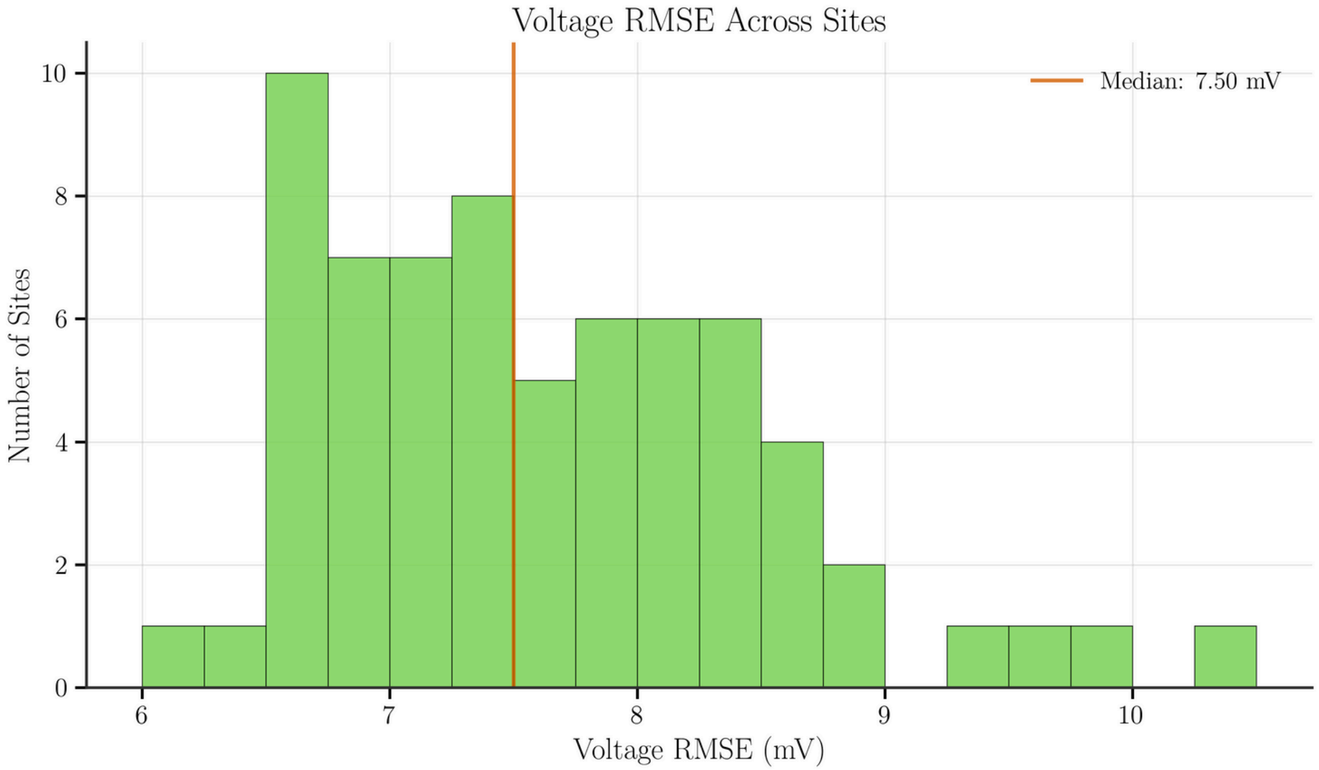
ECM: Does It Work?

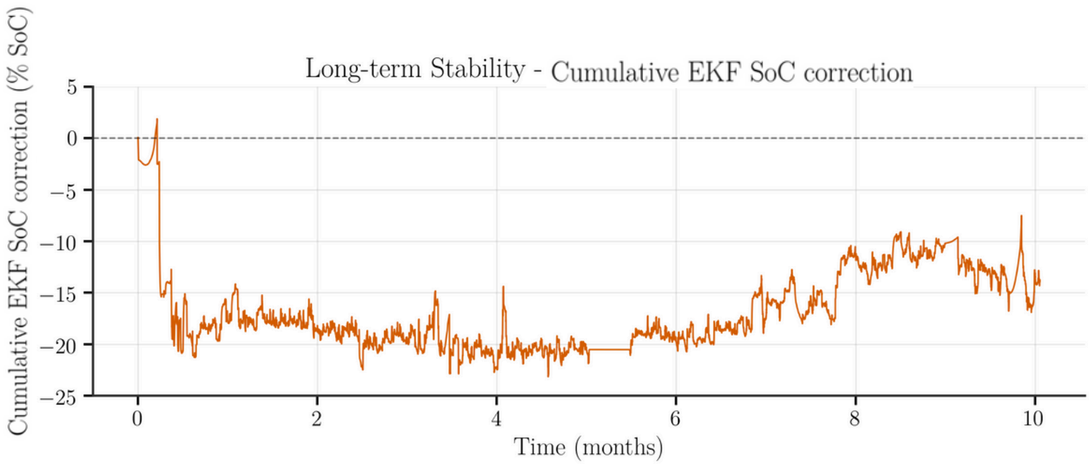
$$U_{\text{sim}} = U_0 + U_{\text{hyst}} + U_1 + U_2$$

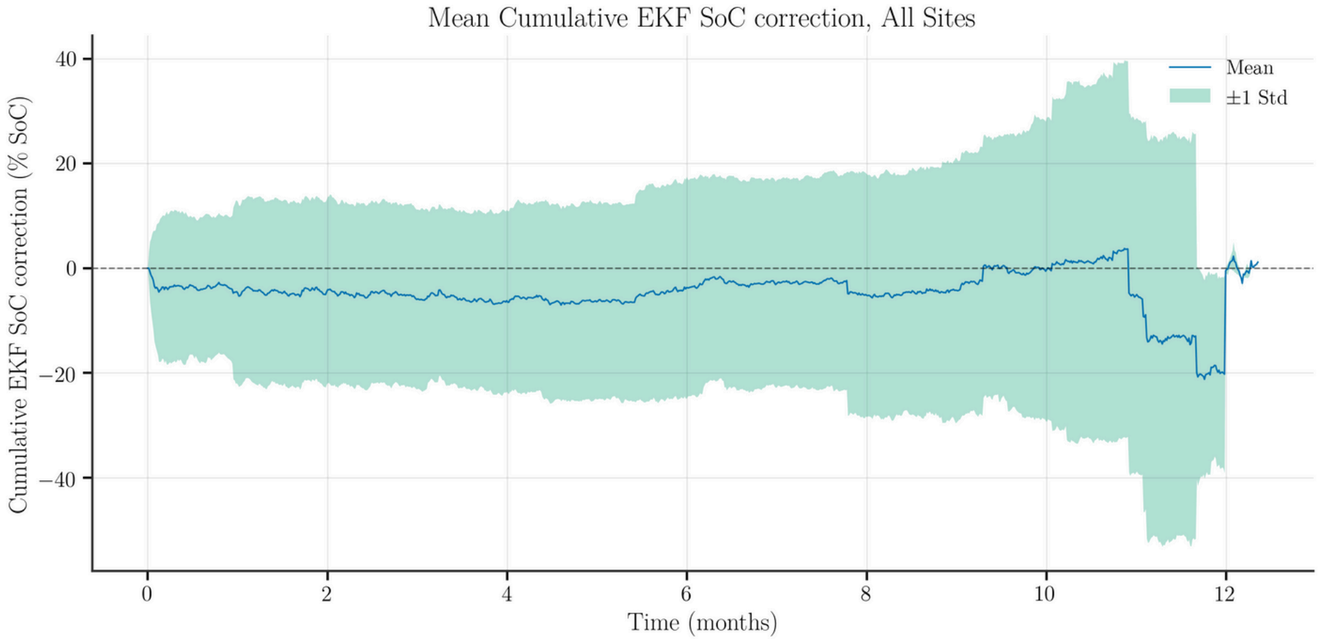


Results: SoC Model

- Across
  - 67 devices
  - 44 years of data
  - 56 GB
- Median voltage **RMSE 7.50 mV**
- Stable SoC corrections
- Literature often reports >10 mV RMSE

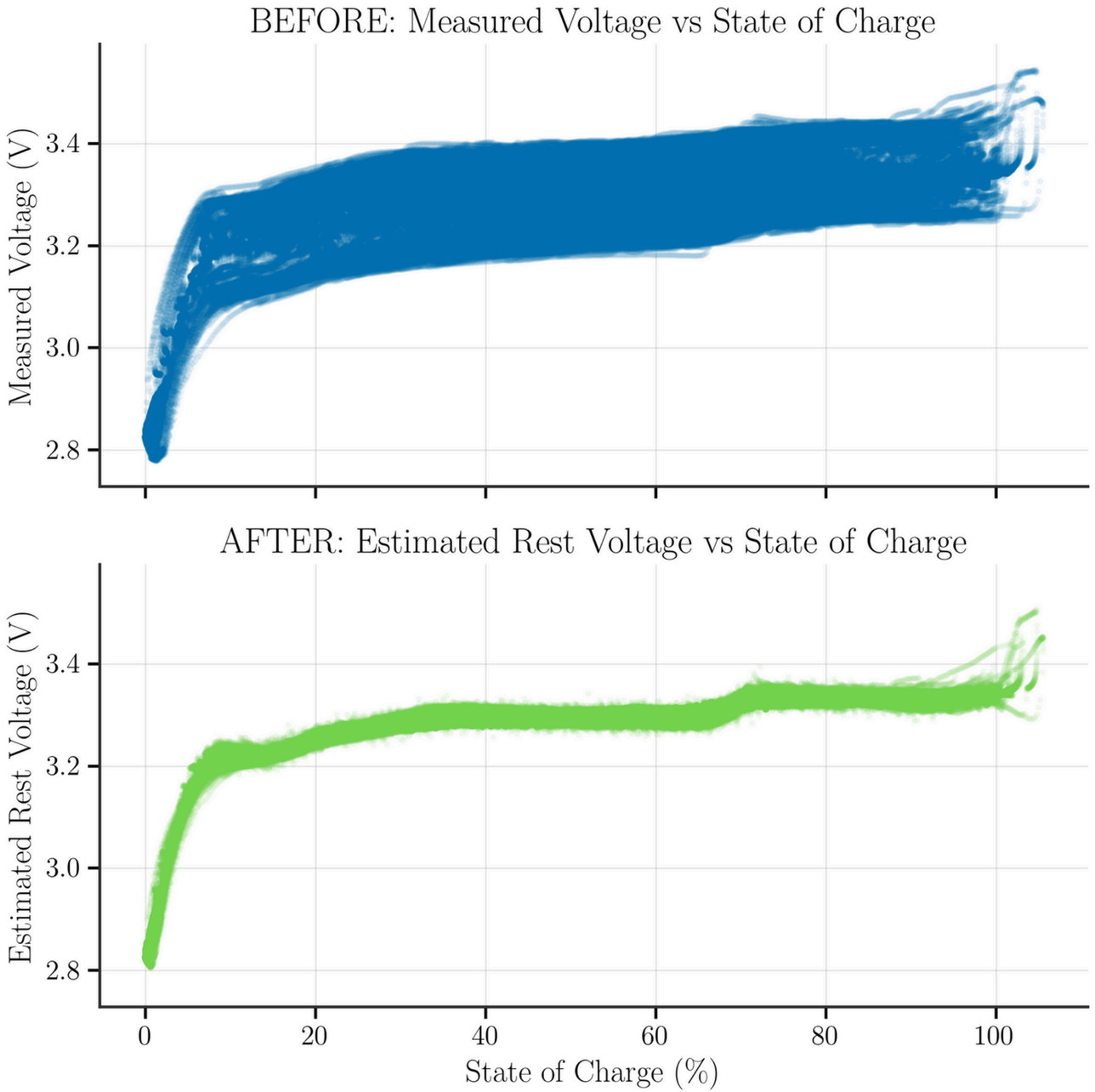






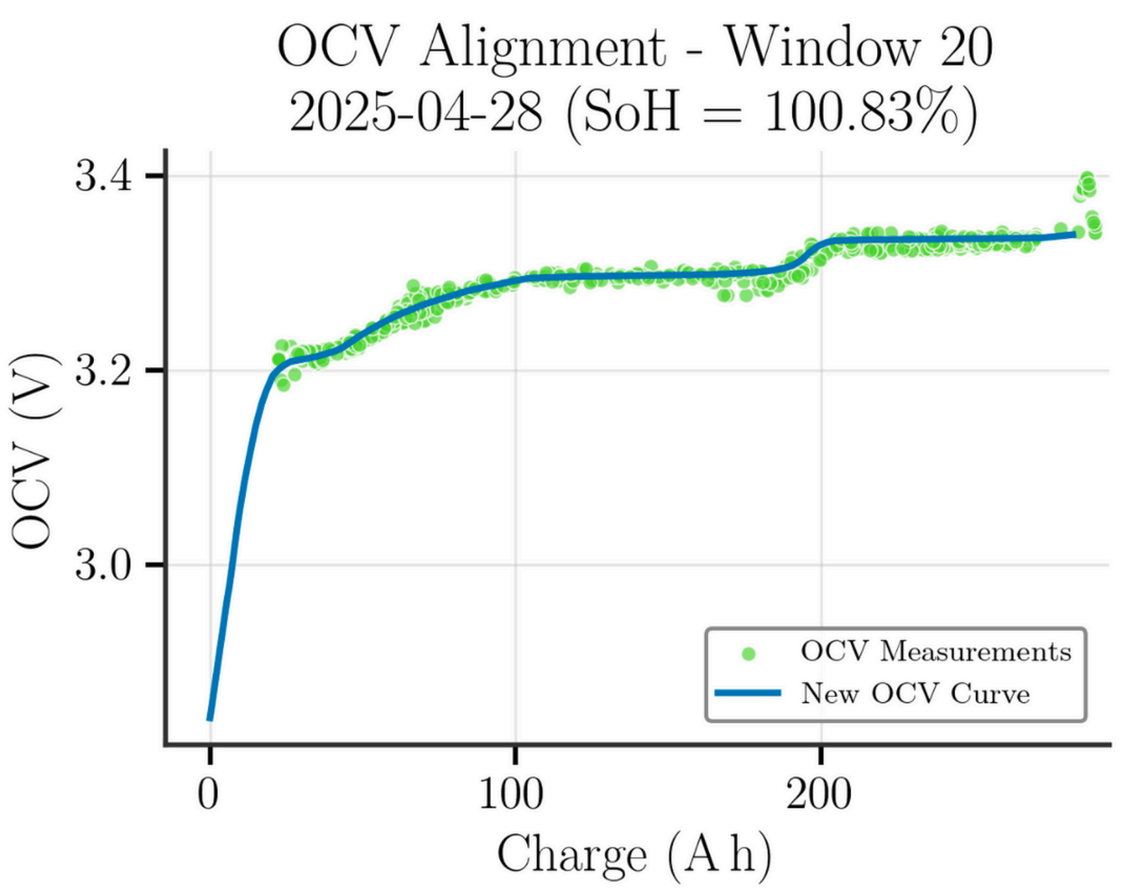
Why Does This Matter?

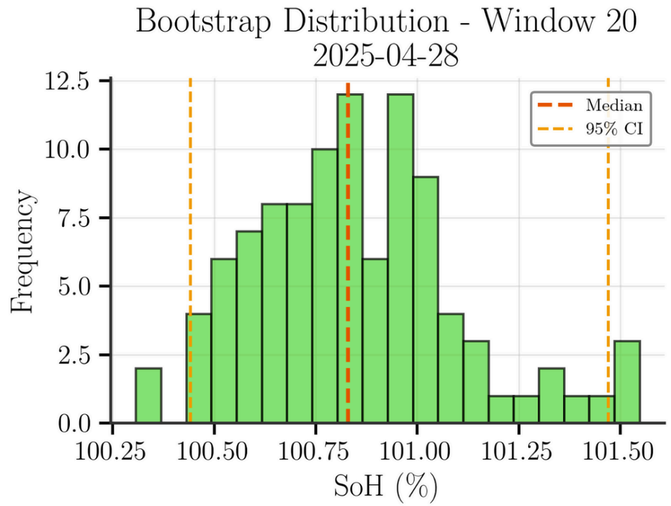
- Fit for our cells in real use
- No lab testing needed
- Adapts to real conditions
- Possibility for improvement through:
  - More data
  - More advanced models



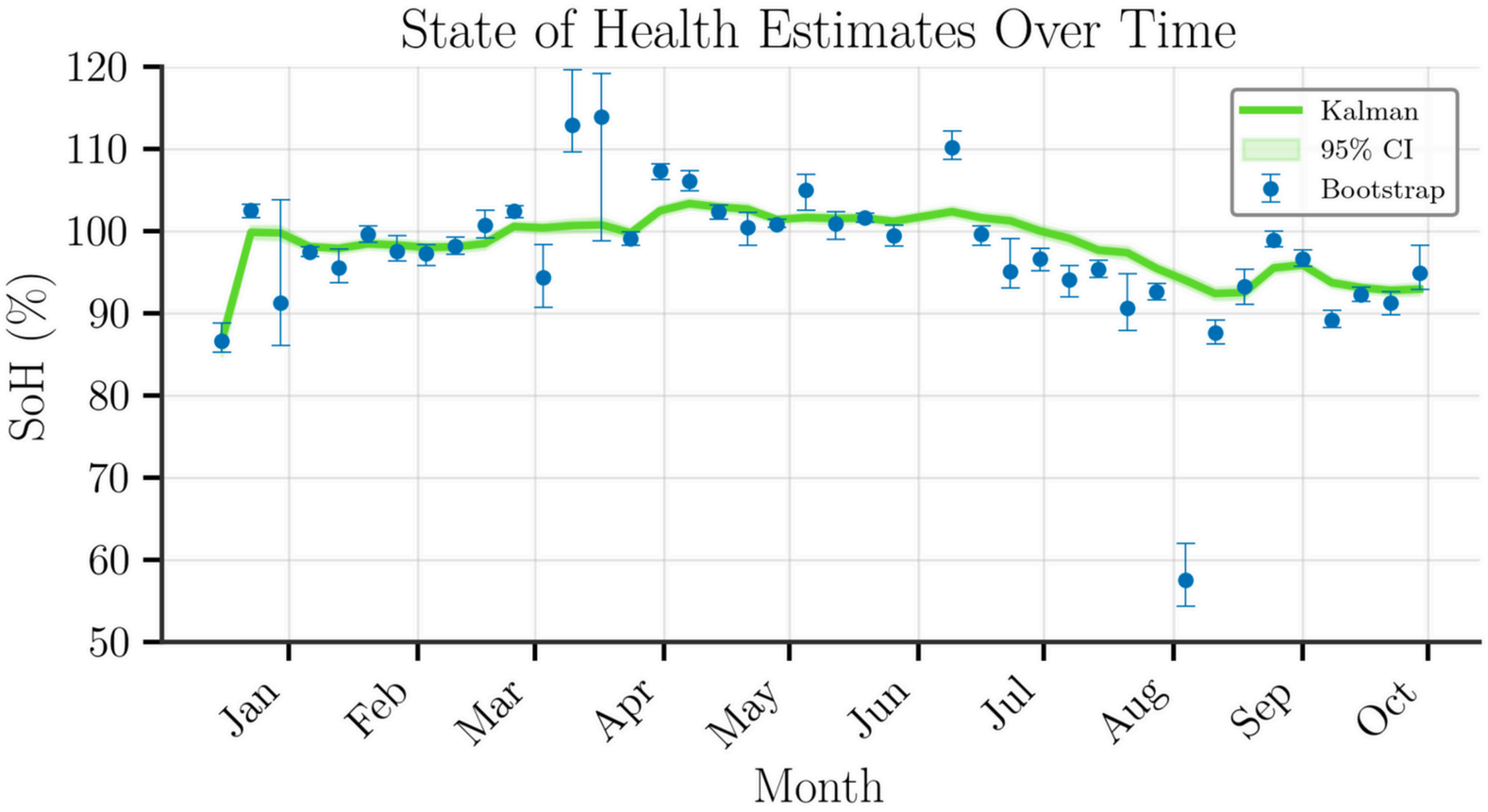
State of Health (SoH) Model

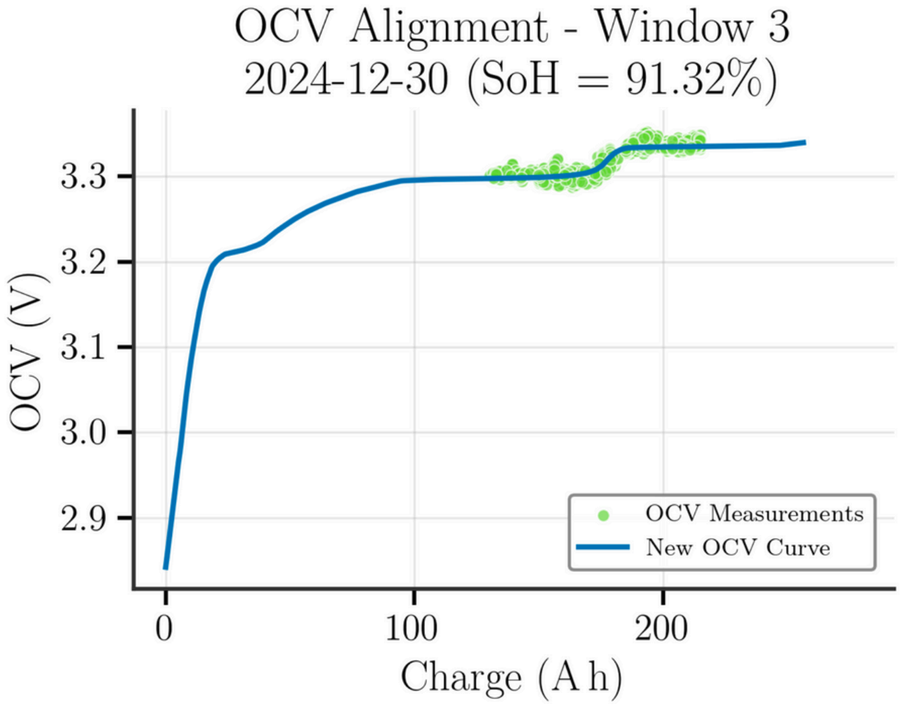
- Built **on top of the SoC model**
- Accurate SoC estimation
  - → basis for SoH estimation
- Method: OCV curve scaling
- Poor data quality (SoC > 20%)
- Two layers of uncertainty filtering
  - Bootstrapping
  - Kalman filter

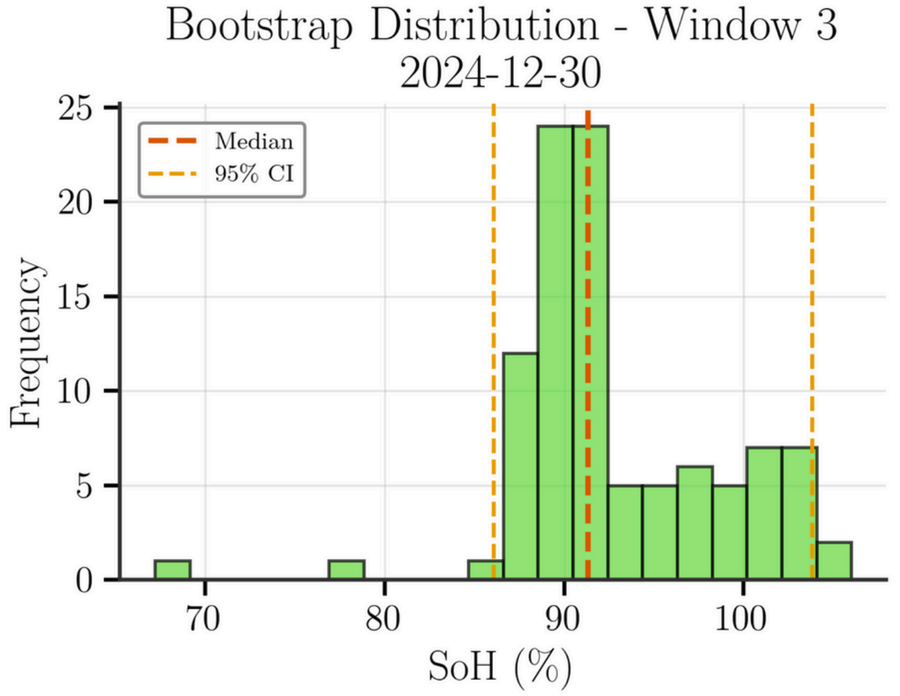




Results: SoH Model

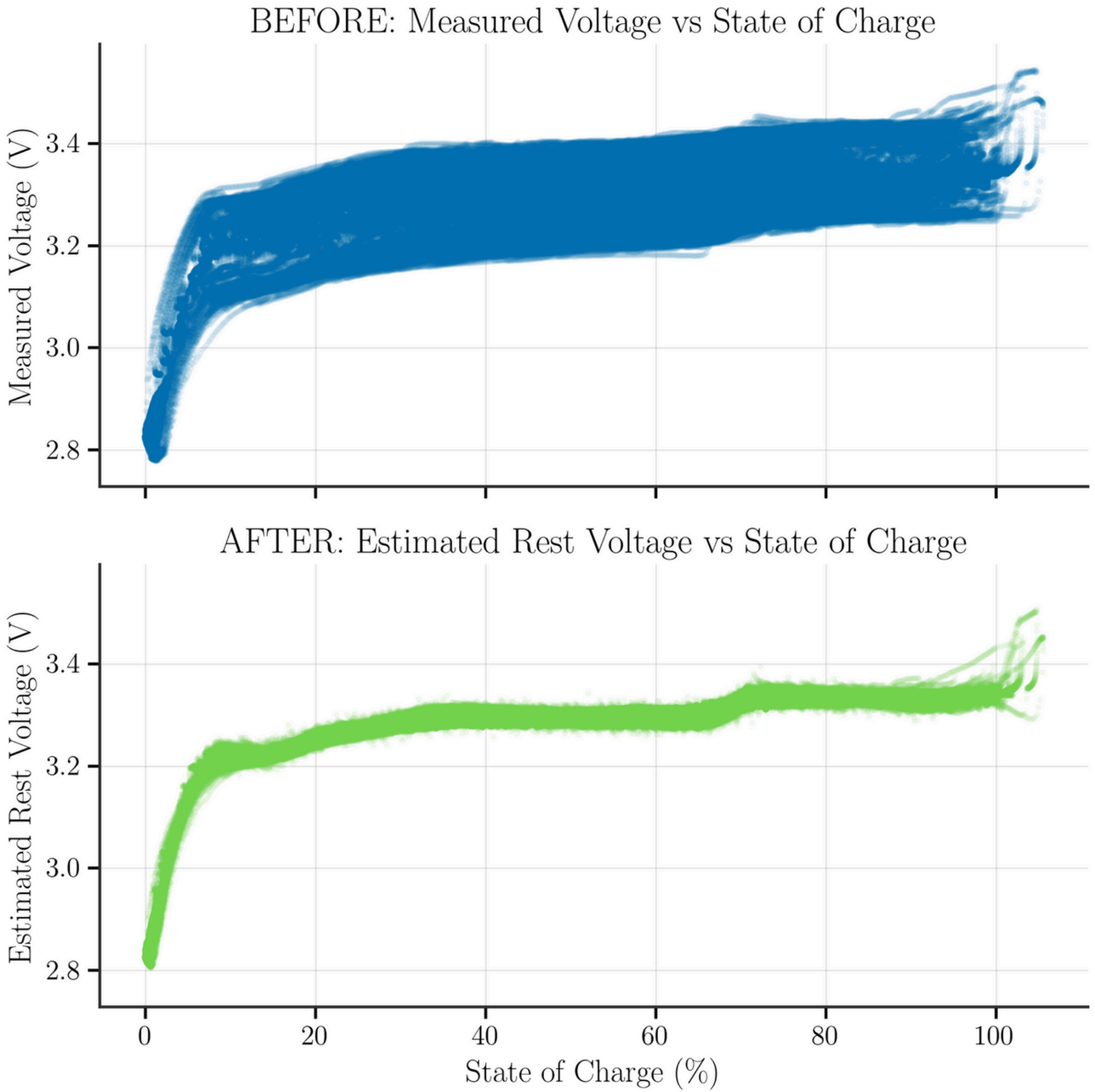






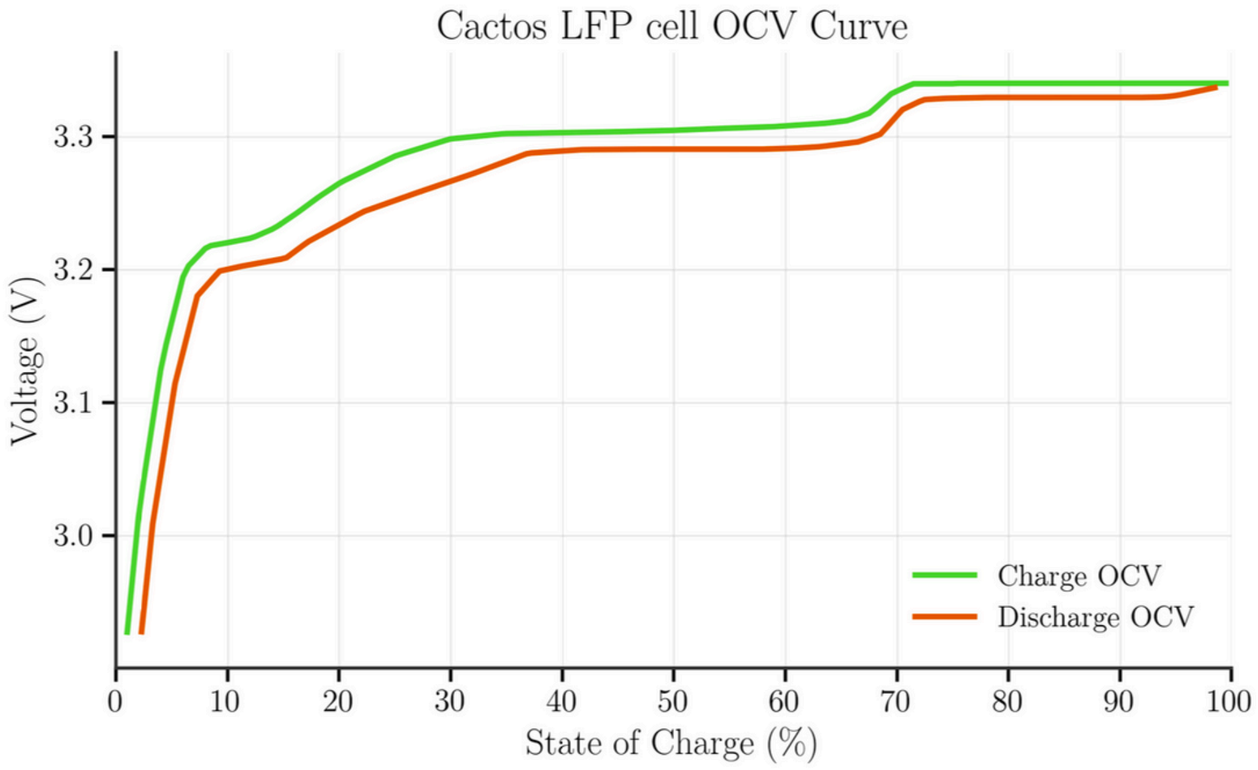
Key Takeaways & Future Research

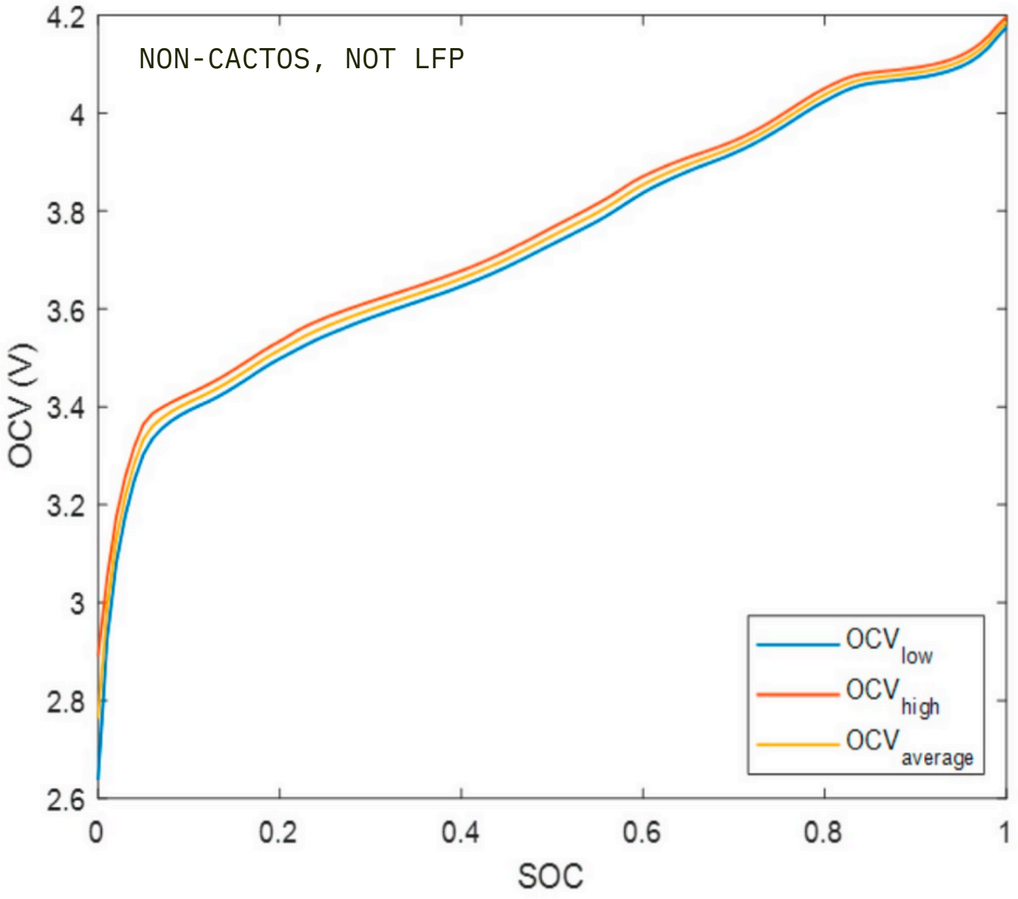
- Operational data brings challenges
- Data science, domain knowledge, and novel thinking
  - → Accurate SoC model
  - → Insights into the actual dependencies of the cells
  - → Production readiness
- Future research
  - Further SoH model validation
  - Aging dependencies
  - SoH estimates into the SoC model



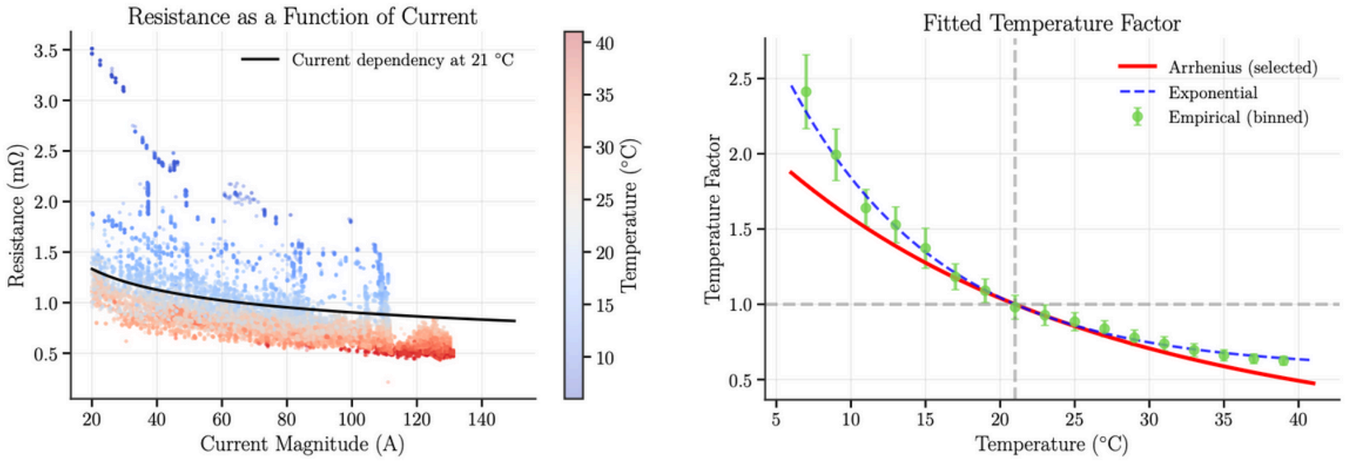
Additional Operational Data Issues

- Measurements not aligned in time
- Voltage calibration bias
- Measurement inaccuracies
  - Cell temperature
- OCV-SoC curve inaccuracy


Cactos LFP cell OCV Curve
Voltage (V)
3.3
3.2
3.1
3.0
0
10
20
30
40
50
60
70
80
90
100
State of Charge (%)
Charge OCV
Discharge OCV

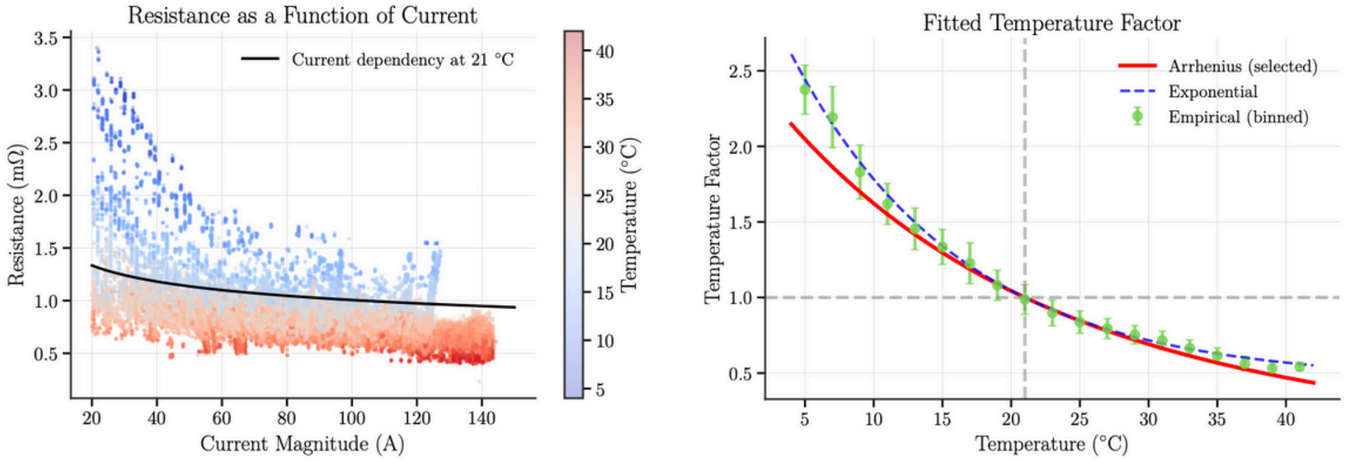

NON-CACTOS, NOT LFP
OCV (V)
4.2
4
3.8
3.6
3.4
3.2
3
2.8
2.6
0
0.2
0.4
0.6
0.8
1
SOC
OCV low
OCV high
OCV average

(b) NMC

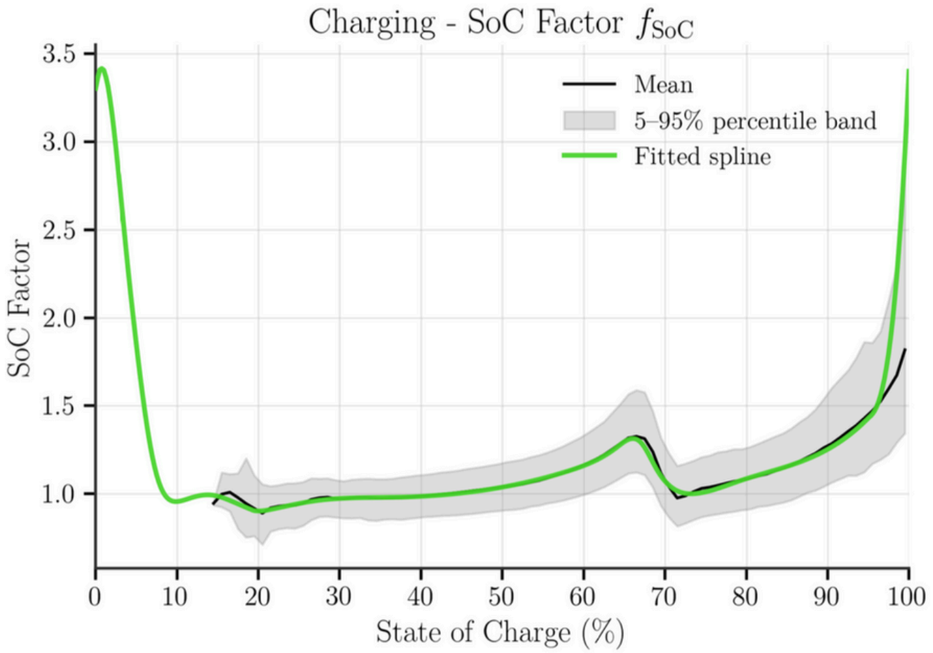


(a) Current dependency of resistance. (b) Temperature dependency of resistance.

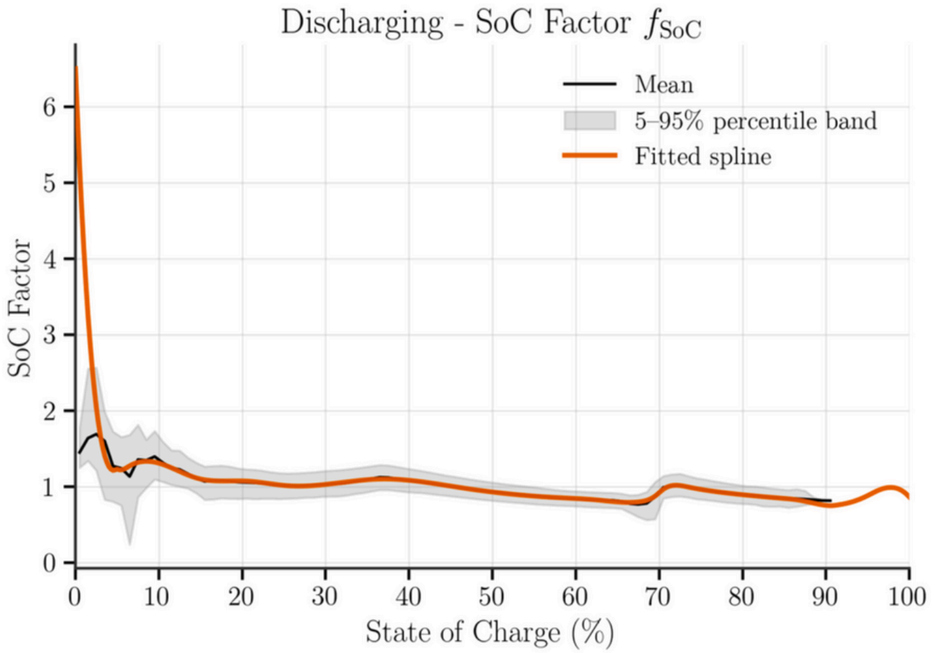
Figure 17: Fitted temperature and current dependencies of resistance for charging.



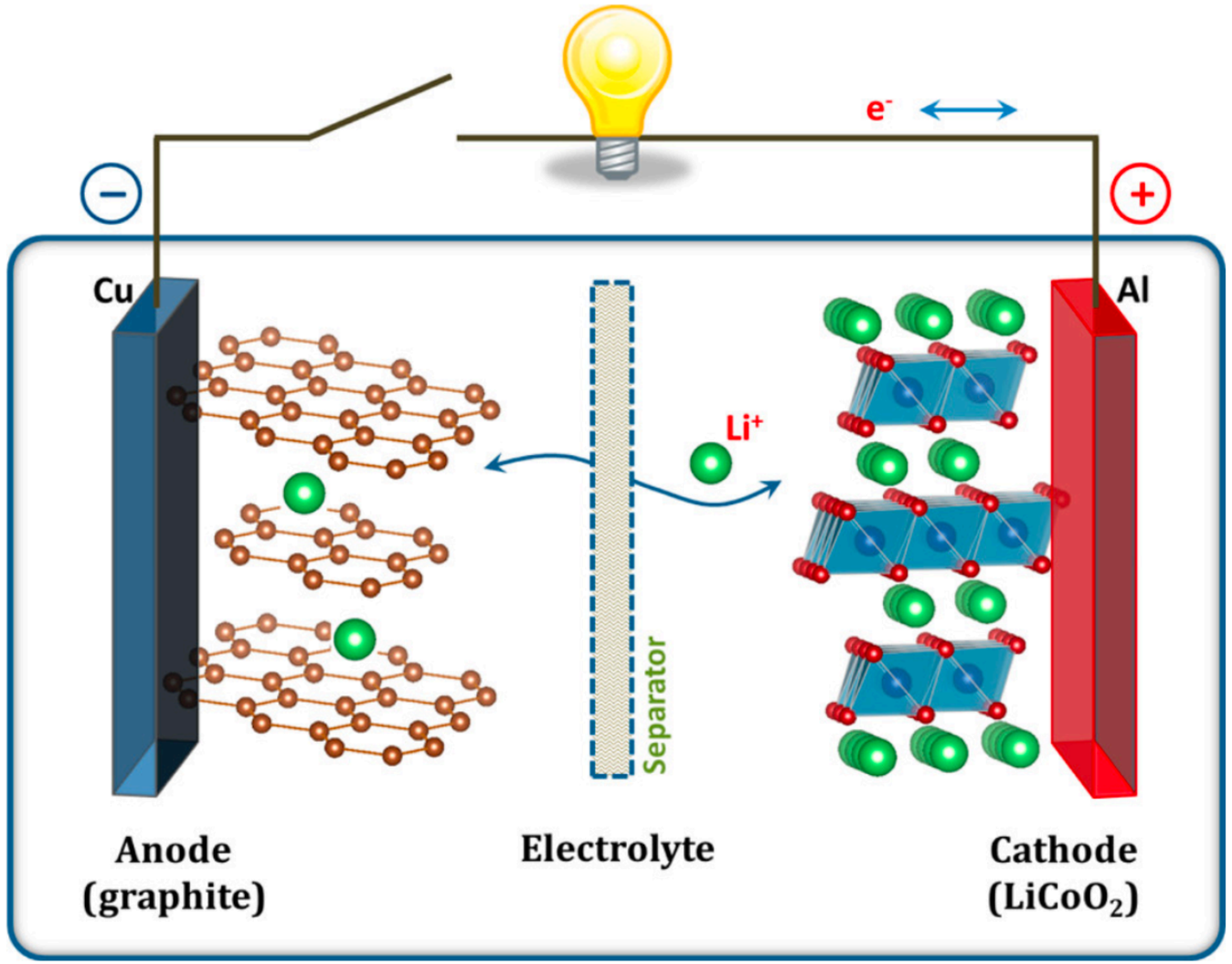
(a) Current dependency of resistance. (b) Temperature dependency of resistance.

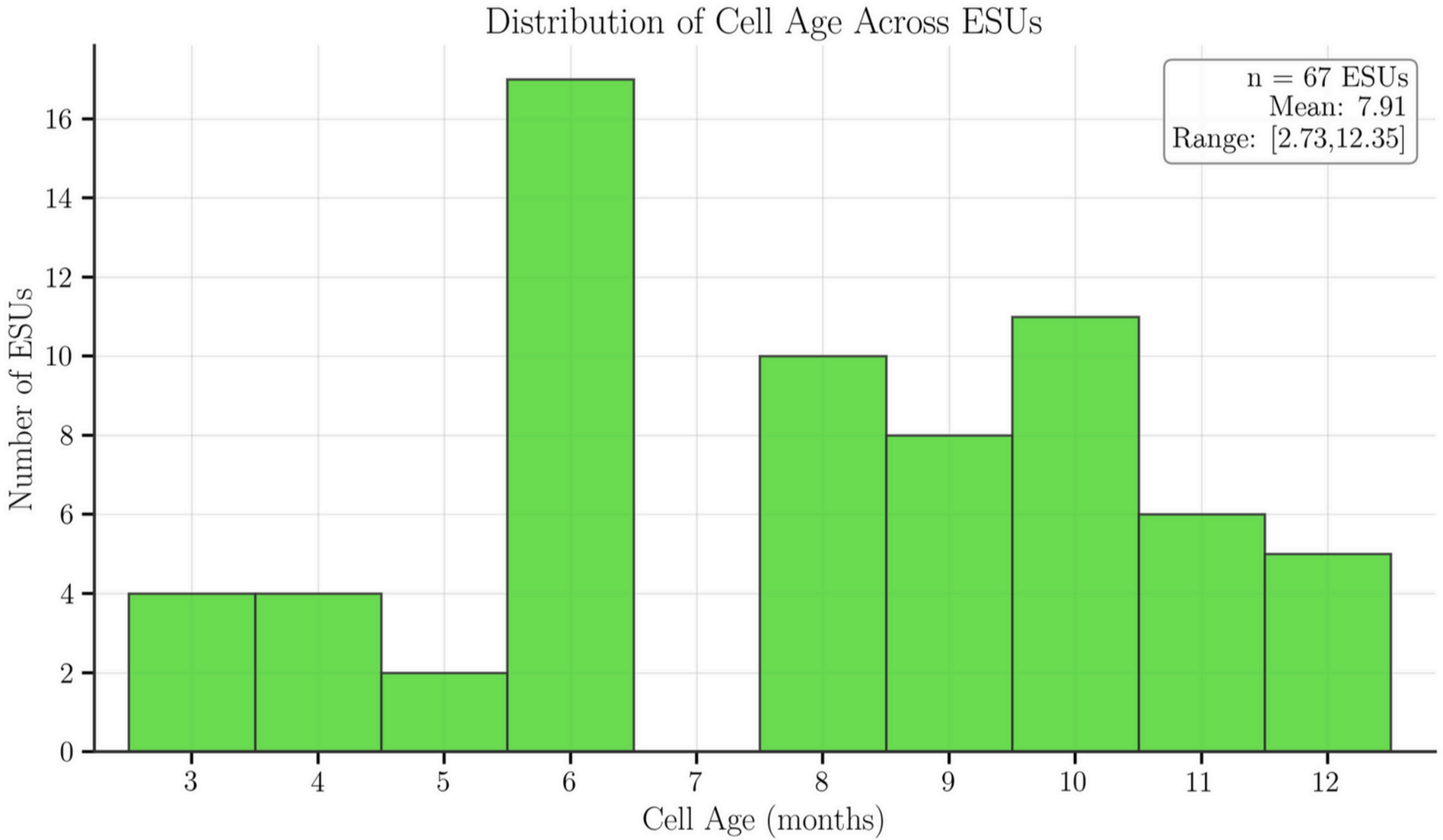


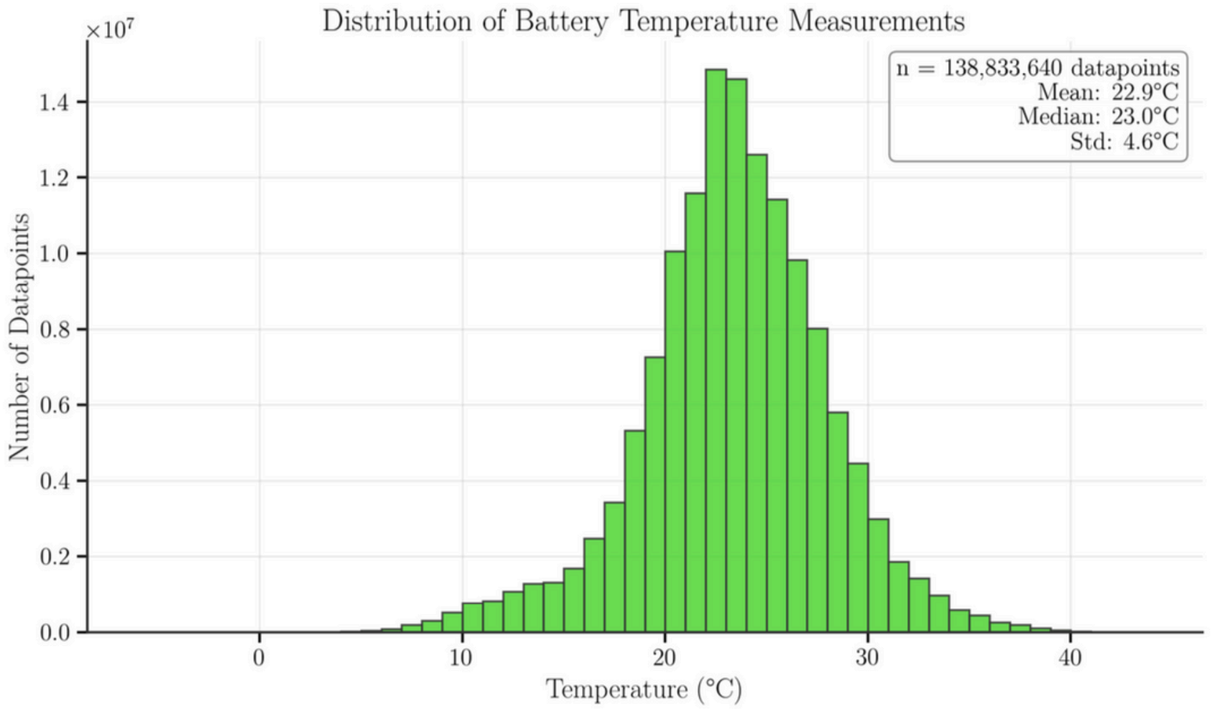
(a) SoC dependency during charging.

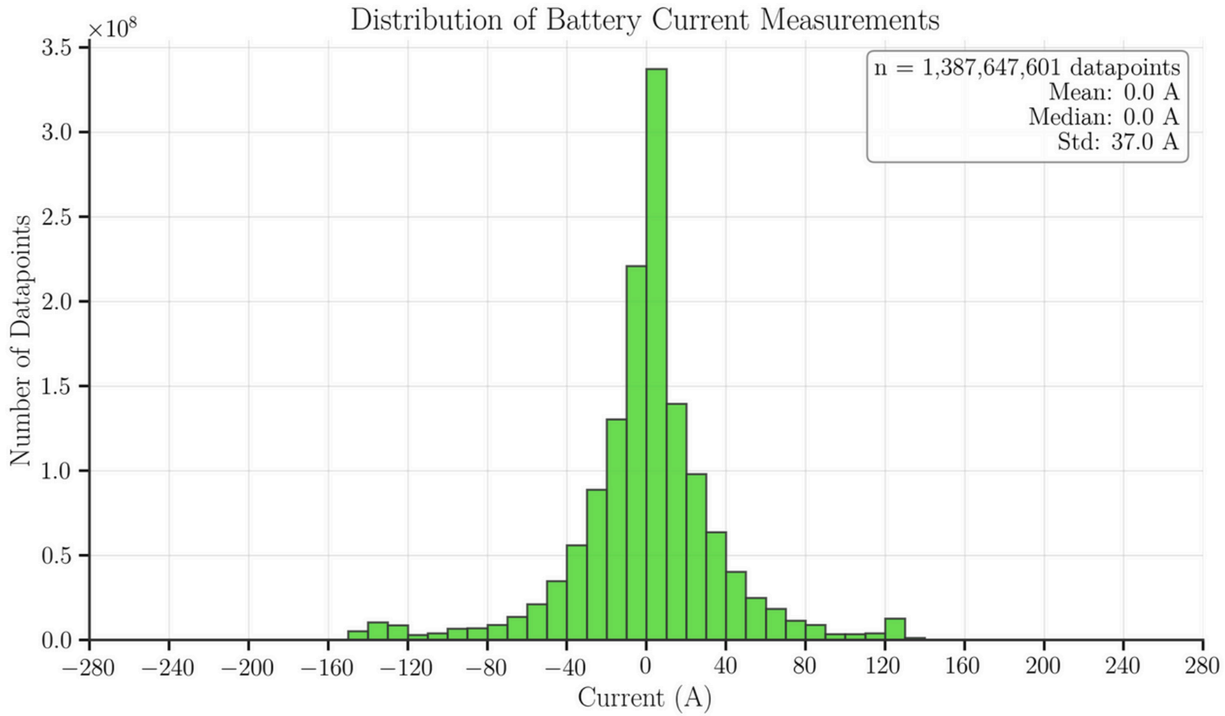


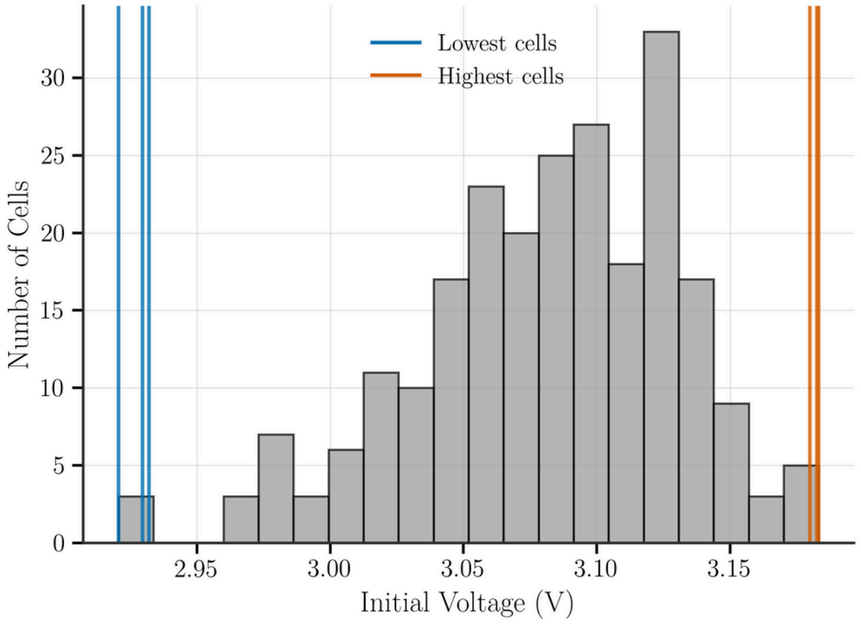
(b) SoC dependency during discharging.

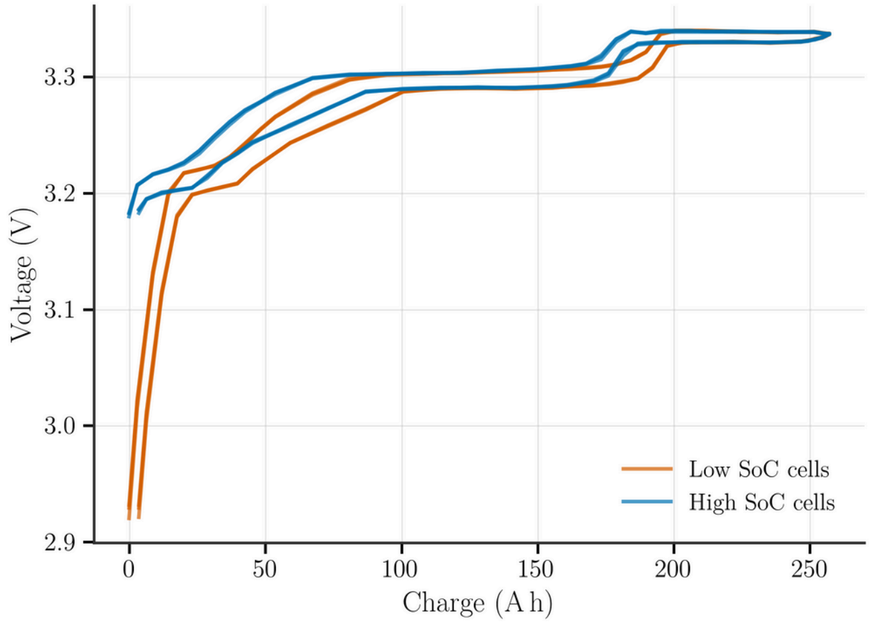

e⁻
Cu
Al
Li⁺
Separator
Anode
(graphite)
Electrolyte
Cathode
(LiCoO2)

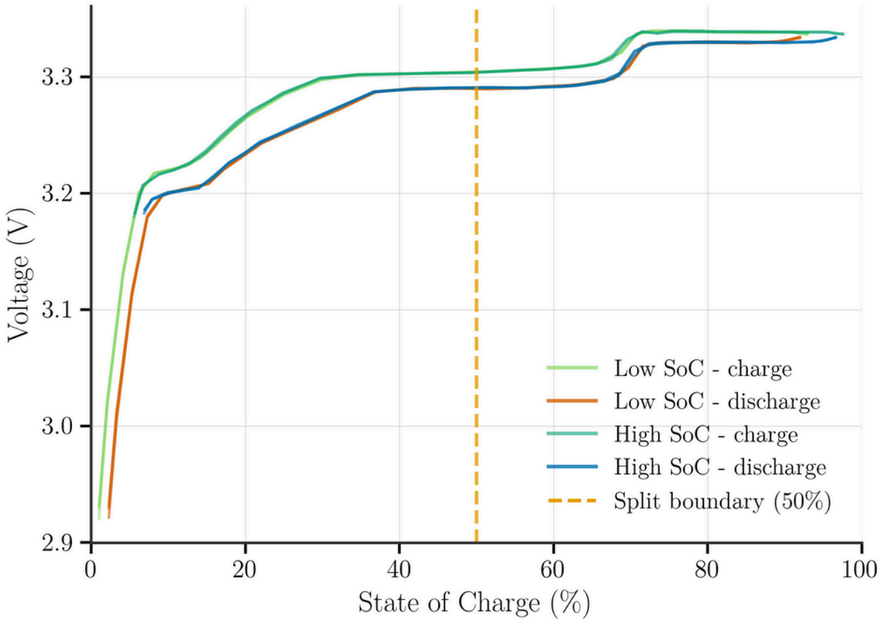

Distribution of Cell Age Across ESUs
n = 67 ESUs
Mean: 7.91
Range: [2.73,12.35]
Number of ESUs
0
2
4
6
8
10
12
14
16
3
4
5
6
7
8
9
10
11
12
Cell Age (months)

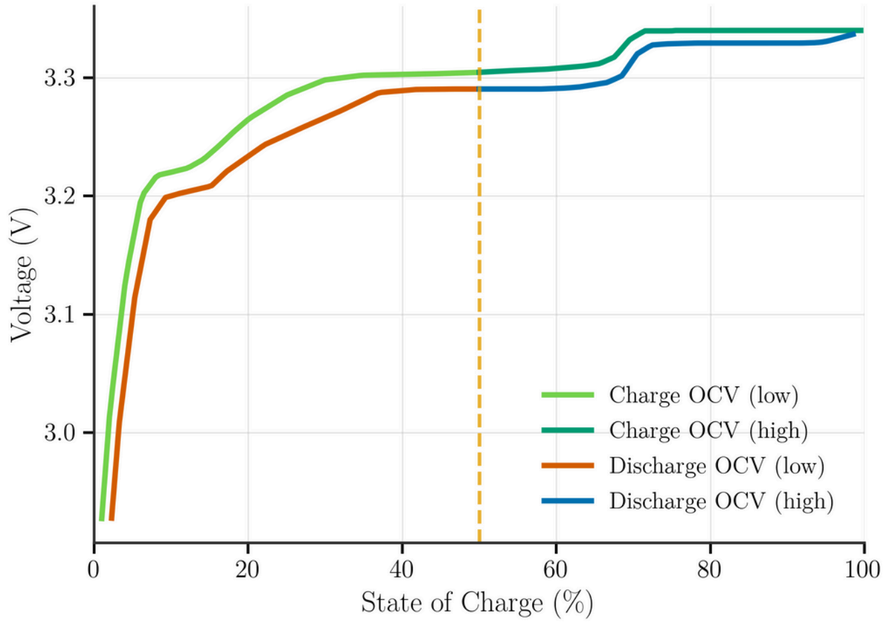

Distribution of Battery Temperature Measurements
n = 138,833,640 datapoints
Mean: 22.9°C
Median: 23.0°C
Std: 4.6°C
×10^7
Number of Datapoints
Temperature (°C)

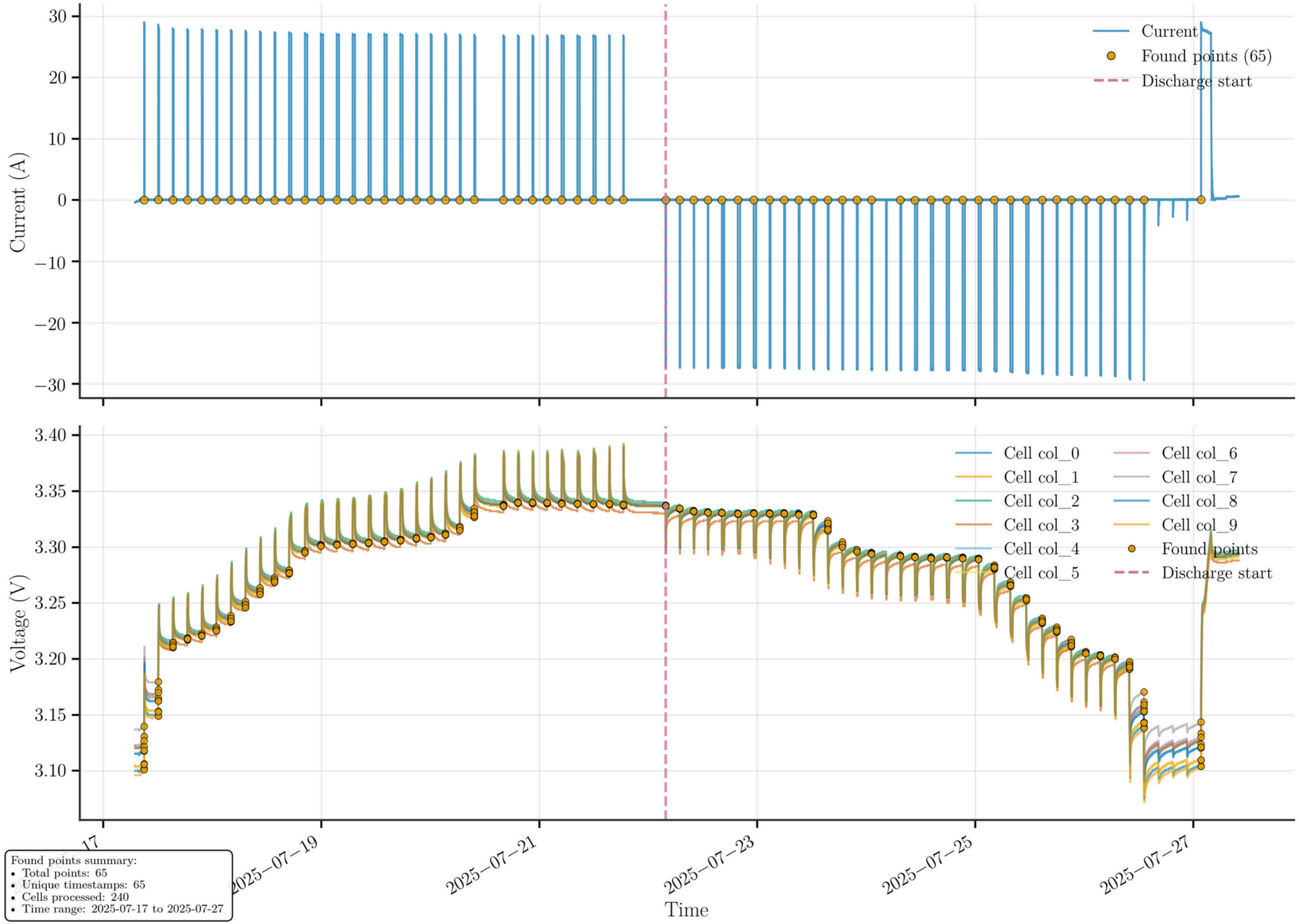

Distribution of Battery Current Measurements
n = 1,387,647,601 datapoints
Mean: 0.0 A
Median: 0.0 A
Std: 37.0 A
×10^8
Number of Datapoints
Current (A)


Lowest cells
Highest cells
Number of Cells
Initial Voltage (V)
0
5
10
15
20
25
30
2.95
3.00
3.05
3.10
3.15


Low SoC cells
High SoC cells
Voltage (V)
Charge (A h)
2.9
3.0
3.1
3.2
3.3
0
50
100
150
200
250


Low SoC - charge
Low SoC - discharge
High SoC - charge
High SoC - discharge
Split boundary (50%)
Voltage (V)
State of Charge (%)
2.9
3.0
3.1
3.2
3.3
0
20
40
60
80
100


Charge OCV (low)
Charge OCV (high)
Discharge OCV (low)
Discharge OCV (high)
Voltage (V)
State of Charge (%)
3.0
3.1
3.2
3.3
0
20
40
60
80
100


Current (A)
Current
Found points (65)
Discharge start
Voltage (V)
Cell col_0
Cell col_1
Cell col_2
Cell col_3
Cell col_4
Cell col_5
Cell col_6
Cell col_7
Cell col_8
Cell col_9
Found points
Discharge start
Time
2025-07-19
2025-07-21
2025-07-23
2025-07-25
2025-07-27
Found points summary:
• Total points: 65
• Unique timestamps: 65
• Cells processed: 240
• Time range: 2025-07-17 to 2025-07-27