Supplementary Information for: Battery Aging Assessment: From Critical Insights to Enhanced Diagnosis

Yunhong Che[1,2*], Joachim Schaeffer[1,3], Jinwook Rhyu[1], Liang Wu[1], Patrick A. Asinger[1], Minsu Kim[1], Jacob Sass[1], Rolf Findeisen[3], Martin Z. Bazant[1,4], William C. Chueh[5,6*], Richard D. Braatz[1*]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, 02139, MA, USA.
[2]Department of Energy, Aalborg University, Aalborg, 9220, Denmark.
[3]Control and Cyber-Physical Systems Laboratory, Technical University of Darmstadt, Darmstadt, 64289, Germany.
[4]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, 02139, MA, USA.
[5]Department of Energy Science and Engineering, Stanford University, Stanford, 94305, CA, USA.
[6]Applied Energy Division, SLAC National Accelerator Laboratory, Menlo Park, 94025, CA, USA.

*Corresponding author(s). E-mail(s): yunhche@mit.edu; wchueh@stanford.edu; braatz@mit.edu;

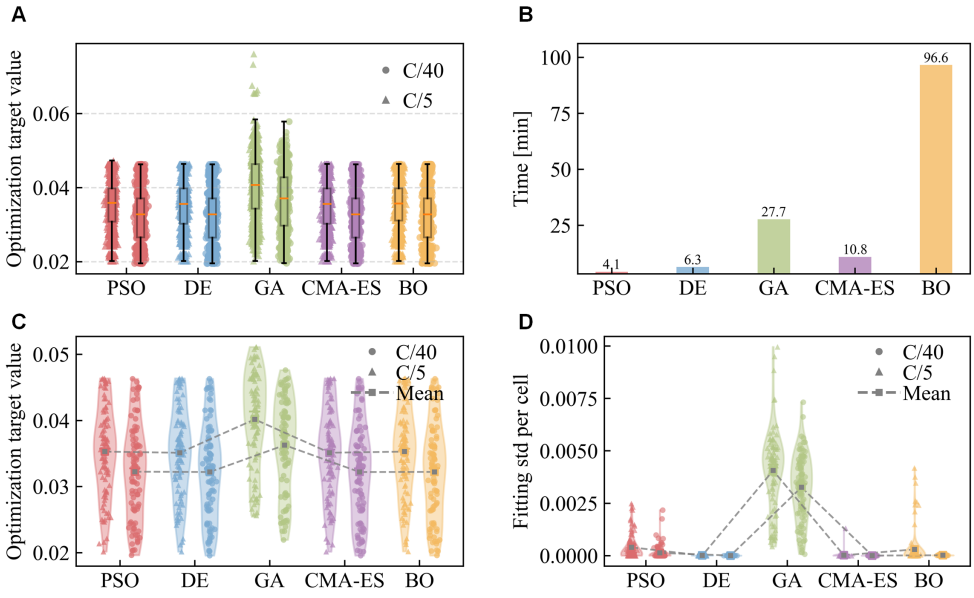

Figure S1 Fitting results with different algorithms based on 4 DOF for Dataset 1 A Fitting results for all the cells in 10 rounds. **B** Time cost for the whole optimization. **C** Best results in 10 rounds of fitting for all the cells in Dataset 1. **D** Standard deviations of the 10 rounds of fitting for all the cells in Dataset 1.

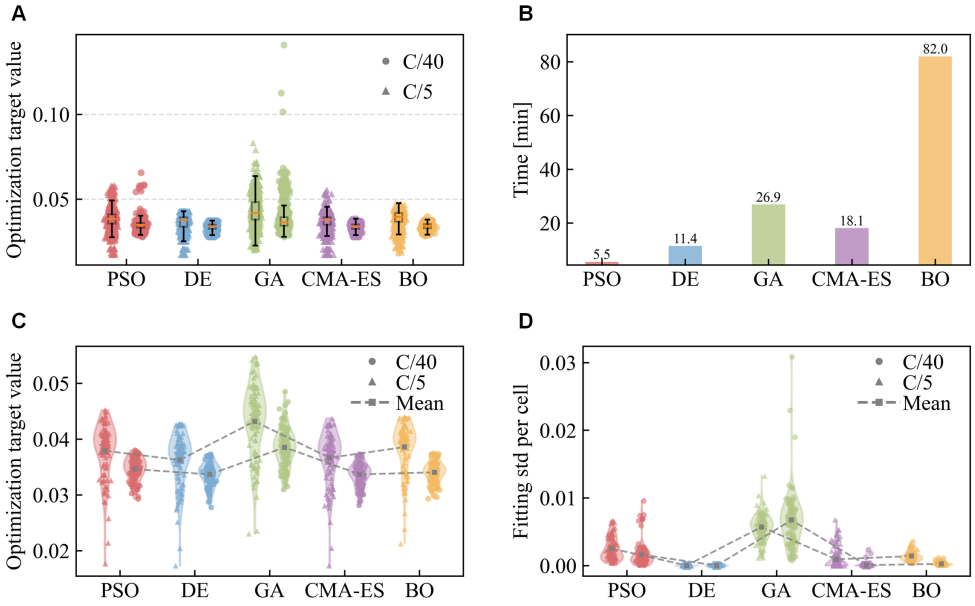

Figure S2 Fitting results with different algorithms based on 3 DOF for Dataset 1 A Fitting results for all the cells in 10 rounds. **B** Time cost for the whole optimization. **C** Best results in 10 rounds of fitting for all the cells in Dataset 1. **D** Standard deviations of the 10 rounds of fitting for all the cells in Dataset 1.

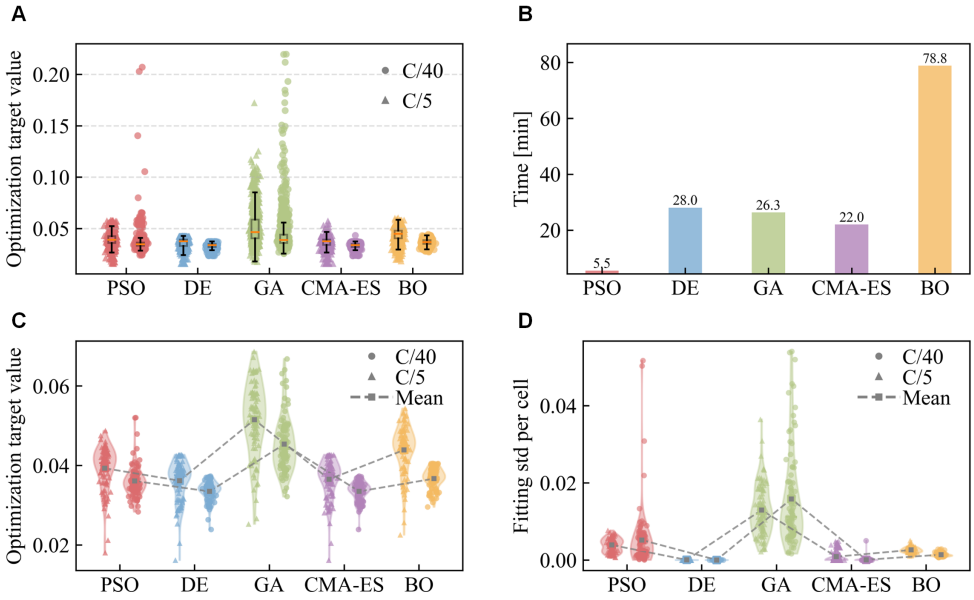

Figure S3 Fitting results with different algorithms based on 2 DOF for Dataset 1 A Fitting results for all the cells in 10 rounds. **B** Time cost for the whole optimization. **C** Best results in 10 rounds of fitting for all the cells in Dataset 1. **D** Standard deviations of the 10 rounds of fitting for all the cells in Dataset 1.

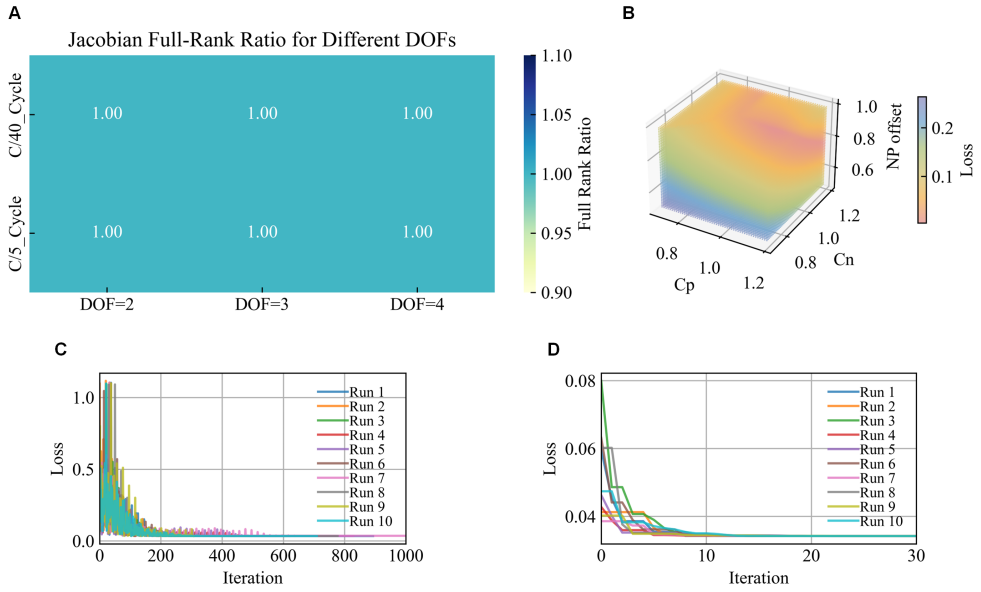

Figure S4 Identifiability analysis A Jacobian analysis. **B** Fitting errors with respect to the variations of the three parameters. **C** Loss variation with respect to iterations using the PSO algorithm in 10 rounds. **D** Loss variation with respect to iterations using the DE algorithm in 10 rounds.

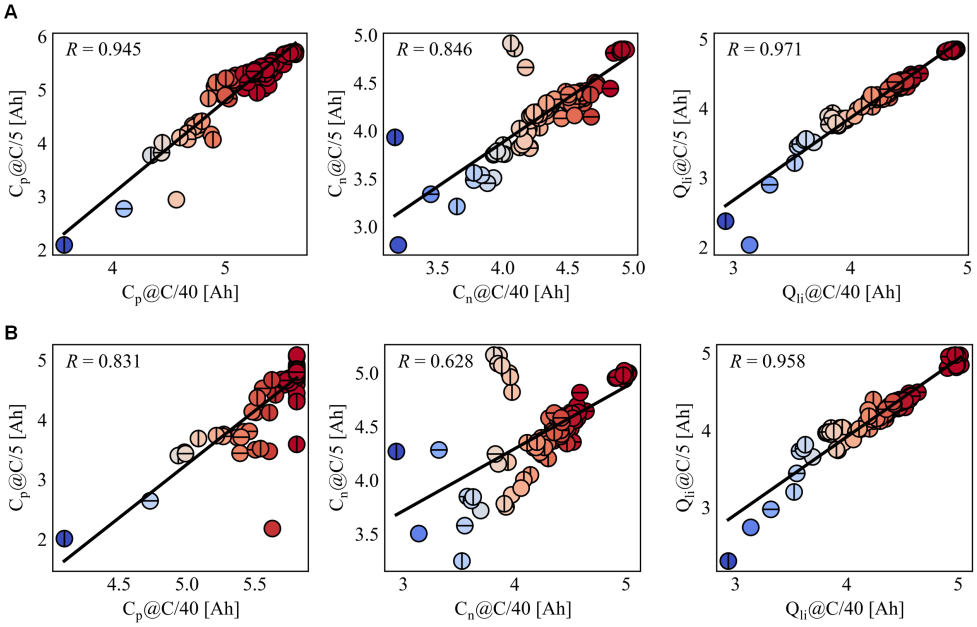

Figure S5 Relationships of the electrode health states between C/5 and C/40 A Results with the matched C-rates between the half cell and full cell OCVs. **B** Results with the mismatched C-rates between the half cell and full cell OCVs.

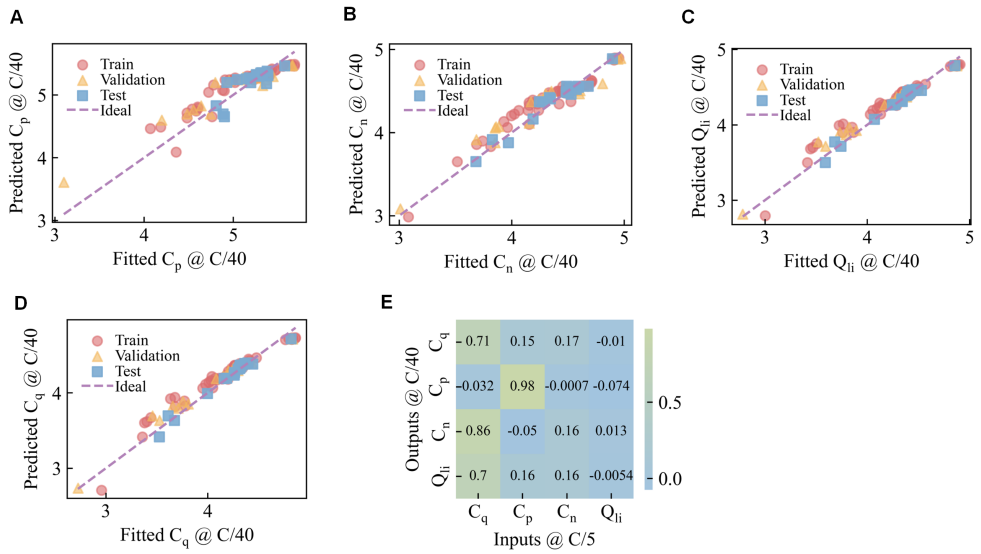

Figure S6 Results for the electrode health prediction with the best performance set A Predictions for the C_p. **B** Predictions for the C_n. **C** Predictions for the Q_{li}. **D** Predictions for the C_q. **E** Model interpretations of the feature contributions.

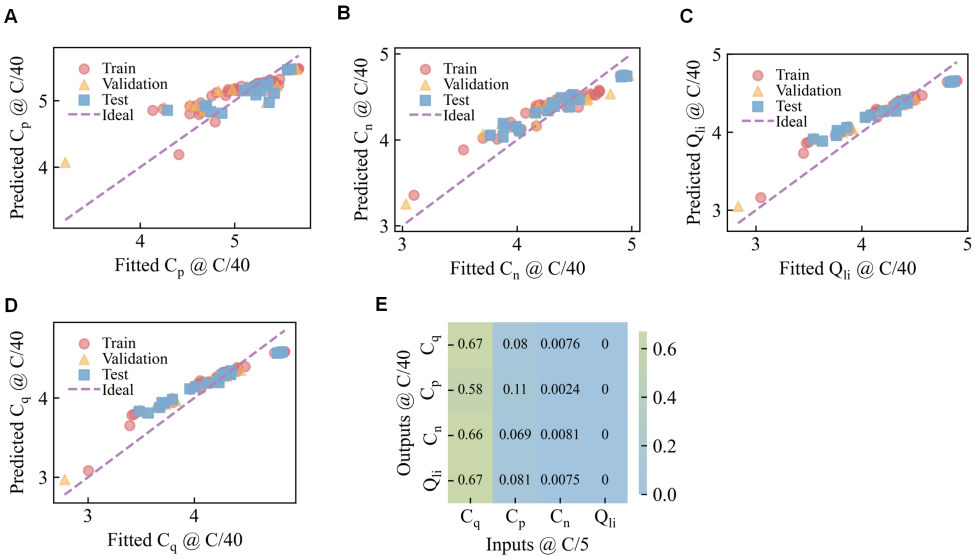

Figure S7 Results for the electrode health prediction with the worst performance set **A** Predictions for the C_p. **B** Predictions for the C_n. **C** Predictions for the Q_{li}. **D** Predictions for the C_q. **E** Model interpretations of the feature contributions.

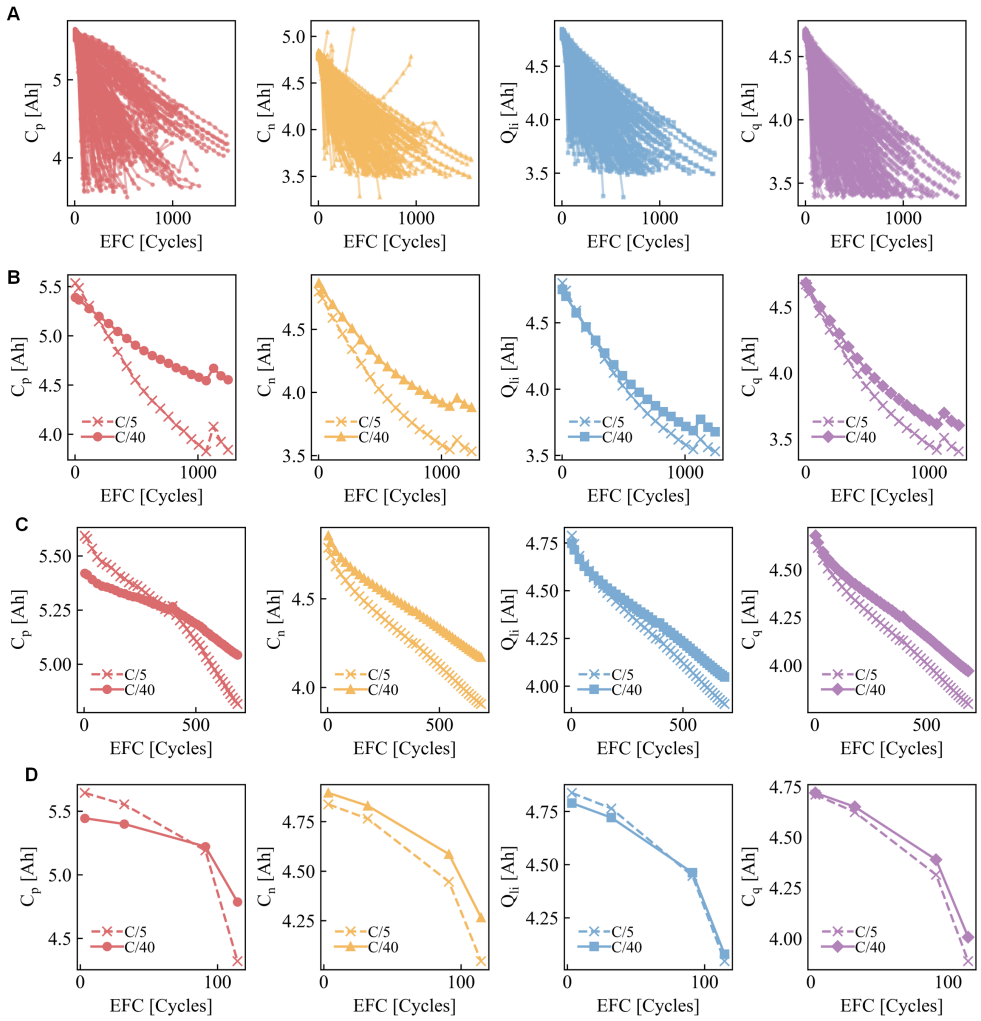

Figure S8 Results for the aging reconstruction with different trajectory patterns A Fitting results for all the cells under C/5. **B** Reconstruction of the aging patterns under C/40 using measurements from C/5 for one cell that has a sub-linear pattern and a lone lifetime with capacity regeneration. **C** Reconstruction of the aging patterns under C/40 using measurements from C/5 for one cell that has a sub-linear followed by super-linear pattern and a medium lifetime. **D** Reconstruction of the aging patterns under C/40 using measurements from C/5 for one cell that has a super-linear pattern and a short lifetime.

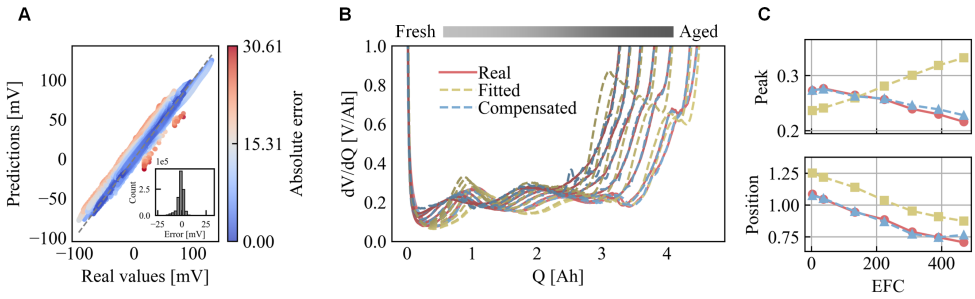

Figure S9 Residual compensations using machine learning from Dataset 1 to Dataset 2 **A** The compensation results for the testing cells in Dataset 2 and the error distributions. **B** Differential voltage curves of one representative cell during aging and the demonstration of the inefficient fittings and the enhanced fittings by machine learning. **C** Aging reconstruction capability of the original model and the enhanced model.

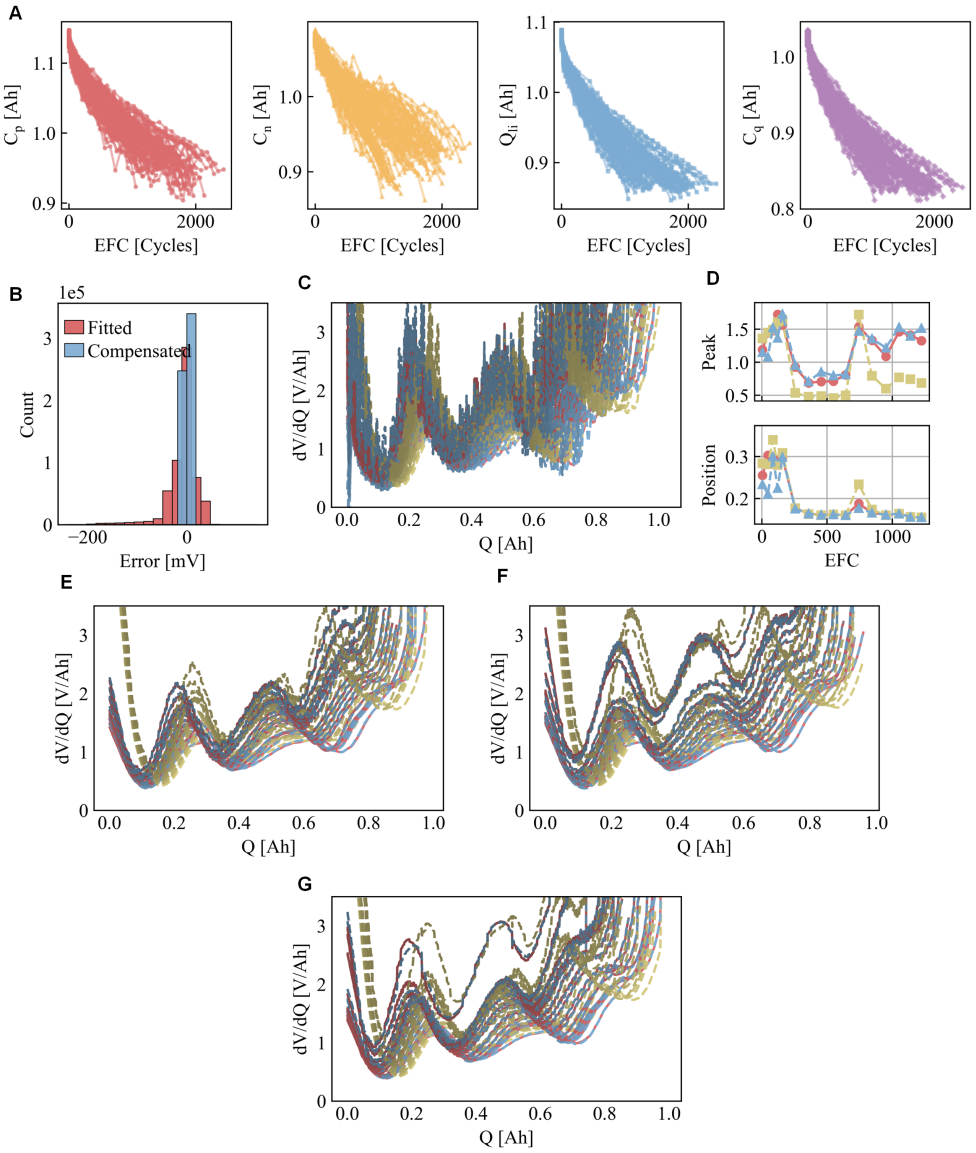

Figure S10 Residual compensations using machine learning from Dataset 3 **A** The DVF results for all the cells in Dataset 3. **B** The compensation results for the testing cells in Dataset 3 and the error distributions without curve smoothing. **C** Differential voltage curves of one representative cell during aging and the demonstration of the inefficient fittings and the enhanced fittings by machine learning without smoothing. **D** Aging reconstruction capability of the original model and the enhanced model for the in C. **E–G** Differential voltage curves reconstruction capability of the original model and the enhanced model for three representative cells aged with highway, urban, and real city loadings, respectively.

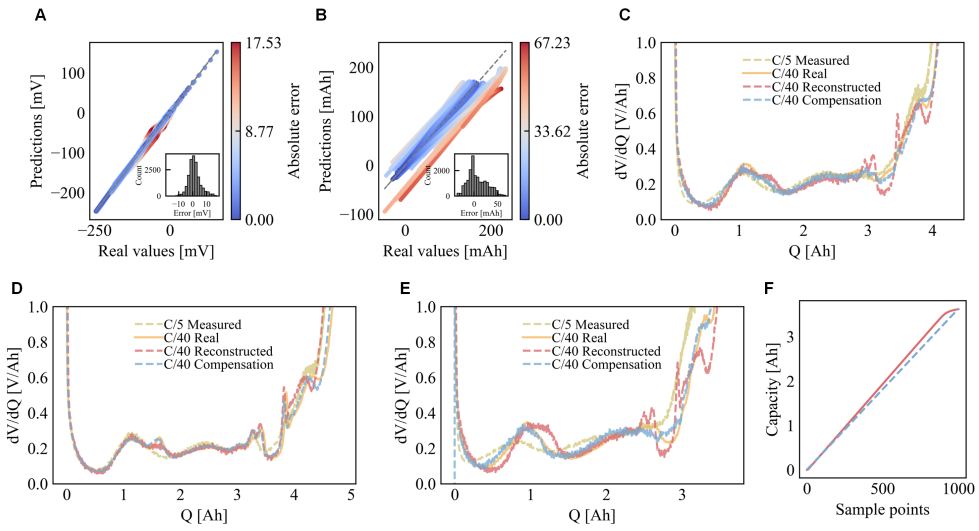

Figure S11 Residual compensations from C/5 to C/40 A The compensation results for the voltage residuals. **B** The compensation results for the capacity residuals. **C**–**E** Differential voltage reconstruction from C/5 to C/40 for three different representative cells. **F** Demonstration of the capacity residuals.

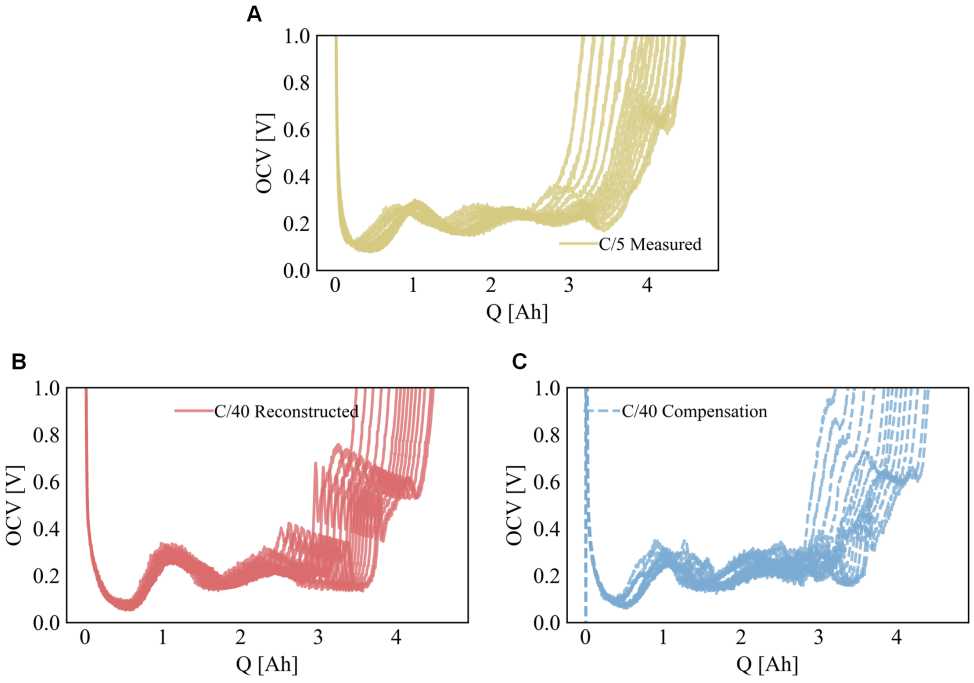

Figure S12 Reconstructions of differential voltage curves from C/5 to C/40 A Measured curves under C/5. **B** Reconstructed curves based on the C-rate adaptive predictions. **C** Enhanced reconstruction based on the residual compensations.

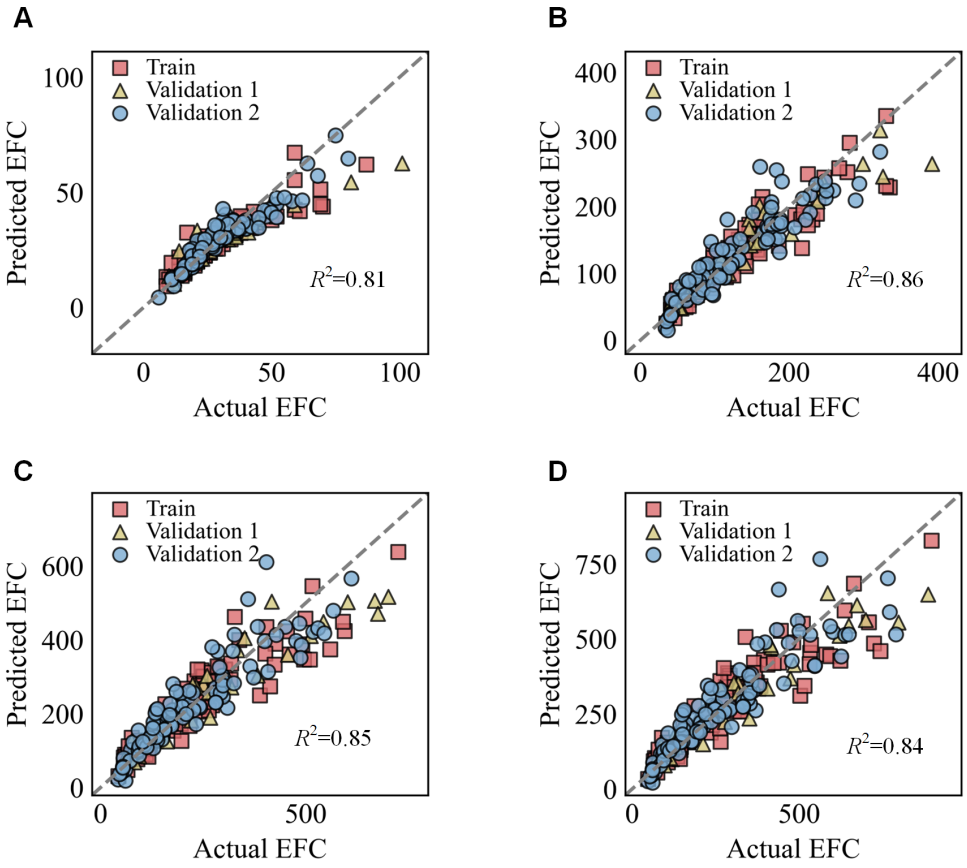

Figure S13 Cycle life prediction until different SOH thresholds. A EFC prediction until 95% SOH threshold. **B** EFC prediction until 90% SOH threshold. **C** EFC prediction until 85% SOH threshold. **D** EFC prediction until 82.5% SOH threshold.

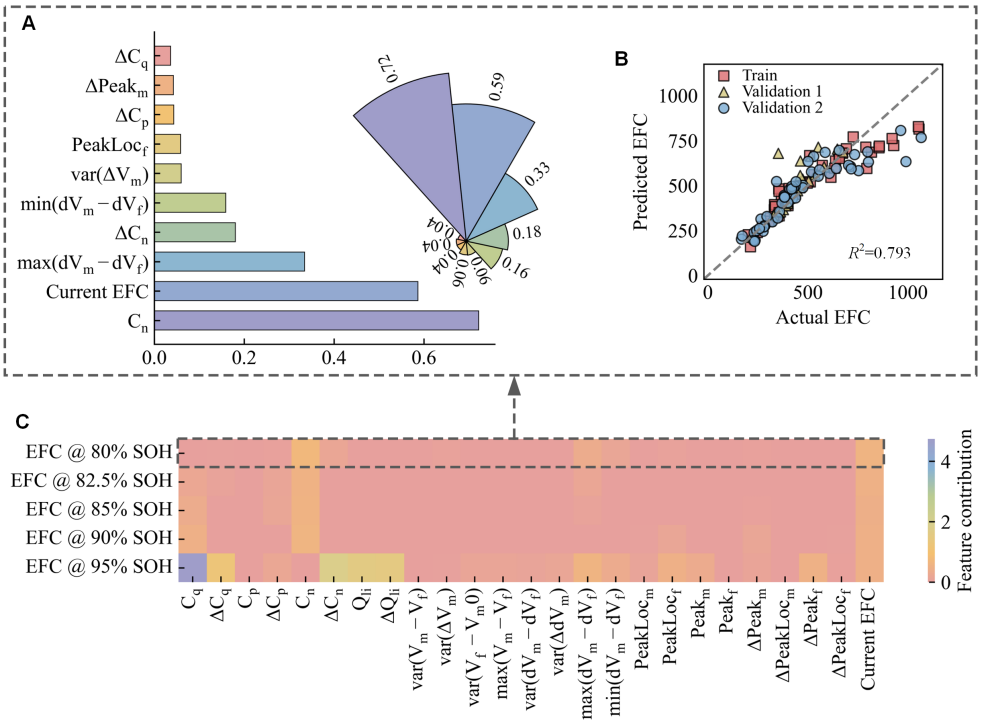

Figure S14 Cycle life prediction from SOH around 85% A Feature importance interpretation for the machine learning model in early cycle life until 80% SOH prediction. **B** Prediction results for the training, validation 1, and validation 2 in early cycle life until 80% SOH. **C** Feature importance interpretation for cycle life predictions until different SOHs.

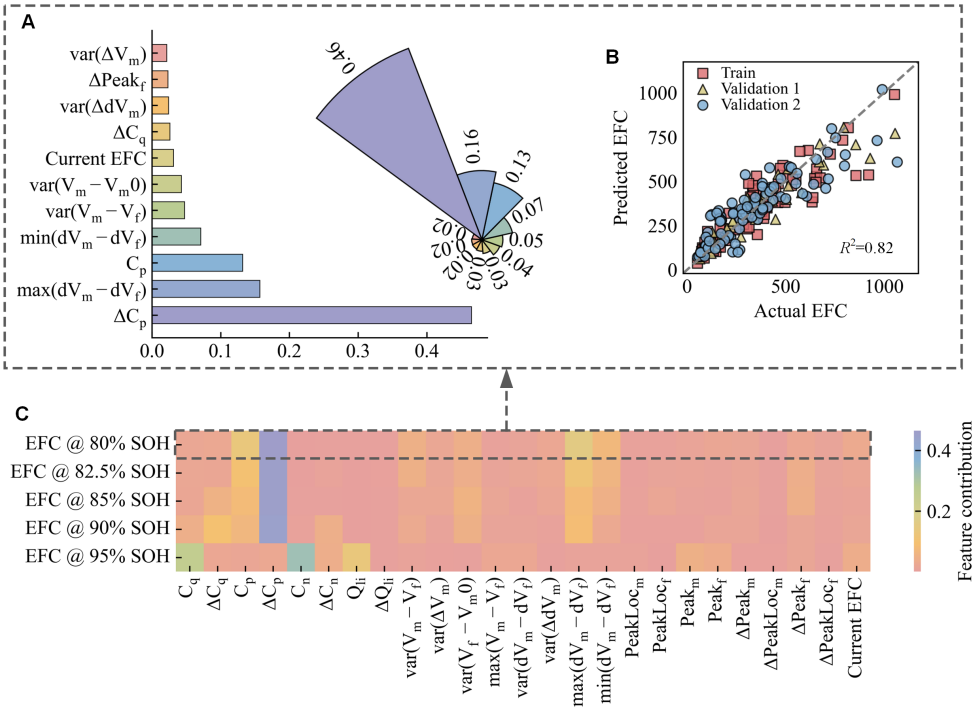

Figure S15 Cycle life prediction using XGboost and SHAP A Feature importance interpretation for the machine learning model in early cycle life until 80% SOH prediction from 30th EFC. **B** Prediction results for the training, validation 1, and validation 2 in early cycle life until 80% SOH from 30th EFC. **C** Feature importance interpretation for cycle life predictions until different SOHs.

Table 1 Summary of the C-rate adaptive prediction of electrode health.

Results	Parameters	RMSE(%)	MAE(%)
Best	C_q	0.96	0.70
	C_p	2.59	1.90
	C_n	1.38	1.15
	Q_{li}	0.86	0.60
Median	C_q	1.88	1.35
	C_p	2.84	2.51
	C_n	1.79	1.40
	Q_{li}	1.62	1.13
Worst	C_q	3.63	2.88
	C_p	4.23	3.18
	C_n	3.33	2.85
	Q_{li}	3.58	2.85

Table 2 Summary of the cycle life prediction using random cell split with 50% for testing.

SOH threshold	Trianing/Validation	RMSE(EFC)	MAE(EFC)	R^2
95% SOH	Training	5.24	3.36	0.84
	Validation 1	7.59	4.78	0.86
	Validation 2	4.46	3.00	0.91
90% SOH	Training	28.18	19.30	0.82
	Validation 1	26.30	19.81	0.91
	Validation 2	35.00	23.55	0.77
85% SOH	Training	57.97	41.02	0.83
	Validation 1	47.55	35.99	0.93
	Validation 2	74.49	53.00	0.75
82.5% SOH	Training	72.26	51.03	0.84
	Validation 1	66.29	53.24	0.90
	Validation 2	93.53	68.89	0.75
80% SOH	Training	93.07	64.96	0.83
	Validation 1	85.35	68.28	0.88
	Validation 2	110.00	81.80	0.76

Note S1 Brief descriptions of the optimization algorithms. In order to estimate the parameters with 2, 3, or 4 degrees of freedom, optimization algorithms are needed. In this work, we employ and compare the performances of 5 different algorithms, including particle swarm optimization (PSO), differential evolution (DE), Bayesian optimization (BO), genetic algorithms (GA), and covariance matrix adaptation evolution strategy (CMA-ES). A brief description of each method is given below.

PSO is a population-based stochastic optimization technique inspired by the social behavior of birds flocking or fish schooling. Each particle represents a potential solution and updates its position by combining its own best-known position and the global best-known position. The velocity and position update rules are [1–3]:

$$\mathbf{v}_i^{(t+1)} = \omega \mathbf{v}_i^{(t)} + c_1 r_1 (\mathbf{p}_i - \mathbf{x}_i^{(t)}) + c_2 r_2 (\mathbf{g} - \mathbf{x}_i^{(t)}) \tag{1}$$

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \mathbf{v}_i^{(t+1)} \tag{2}$$

where $\mathbf{v}_i^{(t)}$ and $\mathbf{x}_i^{(t)}$ denote the velocity and position of particle i at iteration t, $\mathbf{p}_i$ is the best position of particle i, $\mathbf{g}$ is the global best position found so far, ω is the inertia weight, c_1 and c_2 are acceleration coefficients, and $r_1, r_2 \sim \mathcal{U}(0,1)$ are random numbers.

DE is a simple yet powerful evolutionary algorithm that perturbs existing solutions based on the scaled difference between randomly selected individuals. The basic idea is to generate a trial vector by adding the weighted difference of two individuals to a third individual, followed by a crossover and selection step, [4–6]

$$\mathbf{v}_i = \mathbf{x}_{r1} + F \cdot (\mathbf{x}_{r2} - \mathbf{x}_{r3}) \tag{3}$$

$$\mathbf{u}_{i,j} = \begin{cases} \mathbf{v}_{i,j} & \text{if } \text{rand}_j < CR \text{ or } j = j_{\text{rand}} \\ \mathbf{x}_{i,j} & \text{otherwise} \end{cases} \tag{4}$$

Here, $\mathbf{x}_{r1}, \mathbf{x}_{r2}, \mathbf{x}_{r3}$ are randomly chosen individuals, F is a scaling factor, CR is the crossover rate, and $\mathbf{u}_i$ is the offspring. Selection is done by comparing fitness values between $\mathbf{u}_i$ and $\mathbf{x}_i$.

BO is a model-based global optimization method that builds a surrogate probabilistic model, such as a Gaussian Process (GP), of the objective function and uses an acquisition function to decide where to sample next [7–9]. The next point $\mathbf{x}_{\text{next}}$ is chosen by maximizing the acquisition function $a(\mathbf{x})$, such as Expected Improvement (EI):

$$\mathbf{x}_{\text{next}} = \arg\max_{\mathbf{x} \in \mathcal{X}} a(\mathbf{x}) \tag{5}$$

$$\text{EI}(\mathbf{x}) = \mathbb{E}[\max(f(\mathbf{x}) - f_{\text{best}}, 0)] \tag{6}$$

where f_{best} is the current best observation. The GP model provides both a mean and an uncertainty estimate, which guide the trade-off between exploration and exploitation.

GA is a bio-inspired heuristic based on the principles of natural selection and genetics. Individuals in a population undergo selection, crossover, and mutation to

evolve over generations [10, 11]. Given a fitness function $f(\mathbf{x})$, the process iteratively generates new populations,

$$\text{Selection: } \mathbf{x}_{\text{parent}} \sim P(f(\mathbf{x})) \tag{7}$$

$$\text{Crossover: } \mathbf{x}_{\text{child}} = \alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2 \tag{8}$$

$$\text{Mutation: } \mathbf{x}_{\text{mut}} = \mathbf{x}_{\text{child}} + \sigma \cdot \epsilon \tag{9}$$

where $\alpha \in [0, 1]$ is a mixing coefficient, $\epsilon \sim \mathcal{N}(0, 1)$, and σ is a mutation strength. GA is effective for exploring rugged, high-dimensional landscapes.

CMA-ES is an advanced evolutionary strategy that adapts the sampling distribution based on the past successful samples [12–14]. It uses a multivariate normal distribution to sample candidates and updates the covariance matrix to guide the search,

$$\mathbf{x}_k^{(t)} \sim \mathcal{N}(\mathbf{m}^{(t)}, {\sigma^{(t)}}^2 \mathbf{C}^{(t)}) \tag{10}$$

$$\mathbf{m}^{(t+1)} = \sum_{k=1}^{\mu} w_k \mathbf{x}_k^{(t)} \tag{11}$$

where $\mathbf{m}^{(t)}$ is the mean, $\sigma^{(t)}$ is the global step-size, $\mathbf{C}^{(t)}$ is the covariance matrix, and w_k are the weights assigned to the best μ solutions in the population. The covariance matrix captures the shape of the objective landscape and is continuously adapted to improve search efficiency.

Note S2 Feature descriptions. Three models are built in this work for the purpose of a C-rate-based adaptive prediction model, a residual compensation model, and a cycle life prediction model. The main features are from the mechanistic parameters obtained from the differential voltage fitting (DVF) model, with additional features of measurement.

C-rate-based adaptive prediction model takes the measured capacity at high C-rate, i.e., C_q, and the fitted electrode health parameters, i.e., C_p, C_n, Q_{li}, as the input features for the predictions of these four parameters at low C-rate.

Residual compensation model. In the scenario of compensation for the fitted curves to the real curve under the same C-rate, i.e., only the voltage residuals need to be compensated. In order to avoid data leakage (the residuals is the difference between the measured and fitted values) and make it suitable for following different C-rates compensations, only the fitted voltage from the DVF model and the three fitted electrode parameters are used as the features. In the application of the residual compensations from the high C-rate to the low C-rate, the predicted reconstructed voltage and the predicted three electrode states at the low C-rate are used as the features. In addition to the residuals from the voltage curve, the capacity from different C-rates also has discrepancies. From the C-rate adaptive prediction model, the capacity under the low C-rate can be predicted. Therefore, a predicted capacity curve of the Q-V curve can be generated by interpolation from 0 to the predicted capacity value. However, this curve is evenly spaced and still has residuals from the actual capacity curve as demonstrated in Figure S11F. Similarly, the predicted capacity and the predicted electrode states are used as input features for the capacity curve compensation.

Cycle life prediction model also takes the features from the fitted electrode parameters as the input features. The full cell capacity and electrode capacities from the second reference performance test (RPT) cycle are used, i.e., C_q, C_p, C_n, and Q_{li}. In order to reflect the degradation and variation in early stages, the difference of these four parameters between the second RPT and the initial RPT are included, i.e., ΔC_q, ΔC_p, ΔC_n, and ΔQ_{li}. In addition to the electrode features, as discussed in the main text, the key features from the differential voltage curve also reflect the degradation mechanisms. Therefore, the first peak value and the corresponding peak location (capacity value) of the fitted and measured curve are included in the feature matrix, i.e., $Peak_m$, $Peak_f$, $PeakLoc_m$, $PeakLoc_f$, as well as the difference between the curve from the second RPT and the initial one, i.e., $\Delta Peak_m$, $\Delta Peak_f$, $\Delta PeakLoc_m$, $\Delta PeakLoc_f$. Several other statistic values including the variance, minimum and maximum value of the voltage difference between the fitted curve and measured curve as well as their derivatives are extracted, which includes the statistic values between the fitted curve and measured curve, var(V_m-V_f), var(dV_m-dV_f), max(V_m-V_f), max(dV_m-dV_f), and min(dV_m-dV_f). The variations of the curves between the second and first RPT curve are also included, which are variations between the measured voltage curve of the two RPT cycles (var(ΔV_m)), the differential voltage curve of the two RPT cycles ((var(ΔdV_m))), and the variations between the fitted voltage to the initial measured voltage curves ((var(V_f-V_{m0}))).

References

[1] Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of ICNN'95-international Conference on Neural Networks, vol. 4, pp. 1942–1948 (1995). ieee

[2] Wang, D., Tan, D., Liu, L.: Particle swarm optimization algorithm: an overview. Soft computing **22**(2), 387–408 (2018)

[3] Eberhart, R., Kennedy, J.: Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp. 1942–1948 (1995). Perth, Australia

[4] Das, S., Suganthan, P.N.: Differential evolution: A survey of the state-of-the-art. IEEE transactions on evolutionary computation **15**(1), 4–31 (2010)

[5] Pant, M., Zaheer, H., Garcia-Hernandez, L., Abraham, A., *et al.*: Differential evolution: A review of more than two decades of research. Engineering Applications of Artificial Intelligence **90**, 103479 (2020)

[6] Das, S., Mullick, S.S., Suganthan, P.N.: Recent advances in differential evolution–an updated survey. Swarm and evolutionary computation **27**, 1–30 (2016)

[7] Pelikan, M.: Bayesian Optimization Algorithm, pp. 31–48. Springer, ??? (2005)

[8] Frazier, P.I.: A tutorial on bayesian optimization. arXiv preprint arXiv:1807.02811 (2018)

[9] Wang, X., Jin, Y., Schmitt, S., Olhofer, M.: Recent advances in bayesian optimization. ACM Computing Surveys **55**(13s), 1–36 (2023)

[10] Holland, J.H.: Genetic algorithms. Scientific american **267**(1), 66–73 (1992)

[11] Mitchell, M.: An Introduction to Genetic Algorithms. MIT press, ??? (1998)

[12] Iruthayarajan, M.W., Baskar, S.: Covariance matrix adaptation evolution strategy based design of centralized pid controller. Expert systems with Applications **37**(8), 5775–5781 (2010)

[13] Beyer, H.-G., Sendhoff, B.: Simplify your covariance matrix adaptation evolution strategy. IEEE Transactions on Evolutionary Computation **21**(5), 746–759 (2017)

[14] Hansen, N., Müller, S.D., Koumoutsakos, P.: Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). Evolutionary computation **11**(1), 1–18 (2003)