SI for Enhanced SOC Estimation for LFP Batteries: A Synergistic Approach Using Coulomb Counting Reset, Machine Learning, and Relaxation

Yunhong Che,+,‡,§, I Le Xu,+, § Remus Teodorescu,‡ Xiaosong Hu,*, II and Simona Onori*,+

+ Department of Energy Science and Engineering, Stanford University, 367 Panama Street, Stanford, CA 94305, USA.

‡ Department of Energy, Aalborg University, Fredrik Bajers Vej 7K, Aalborg 9220, Denmark.

II College of Mechanical and Vehicle Engineering, Chongqing University, 174 Shazheng St., Chongqing 40044, China.

§ These authors contributed equally.

I Visiting Scholar at the Department of Energy Science and Engineering, Stanford University, 367 Panama Street, Stanford, CA 94305, USA.

* E-mail: xiaosonghu@ieee.org; sonori@stanford.edu

Supplementary Figure S1

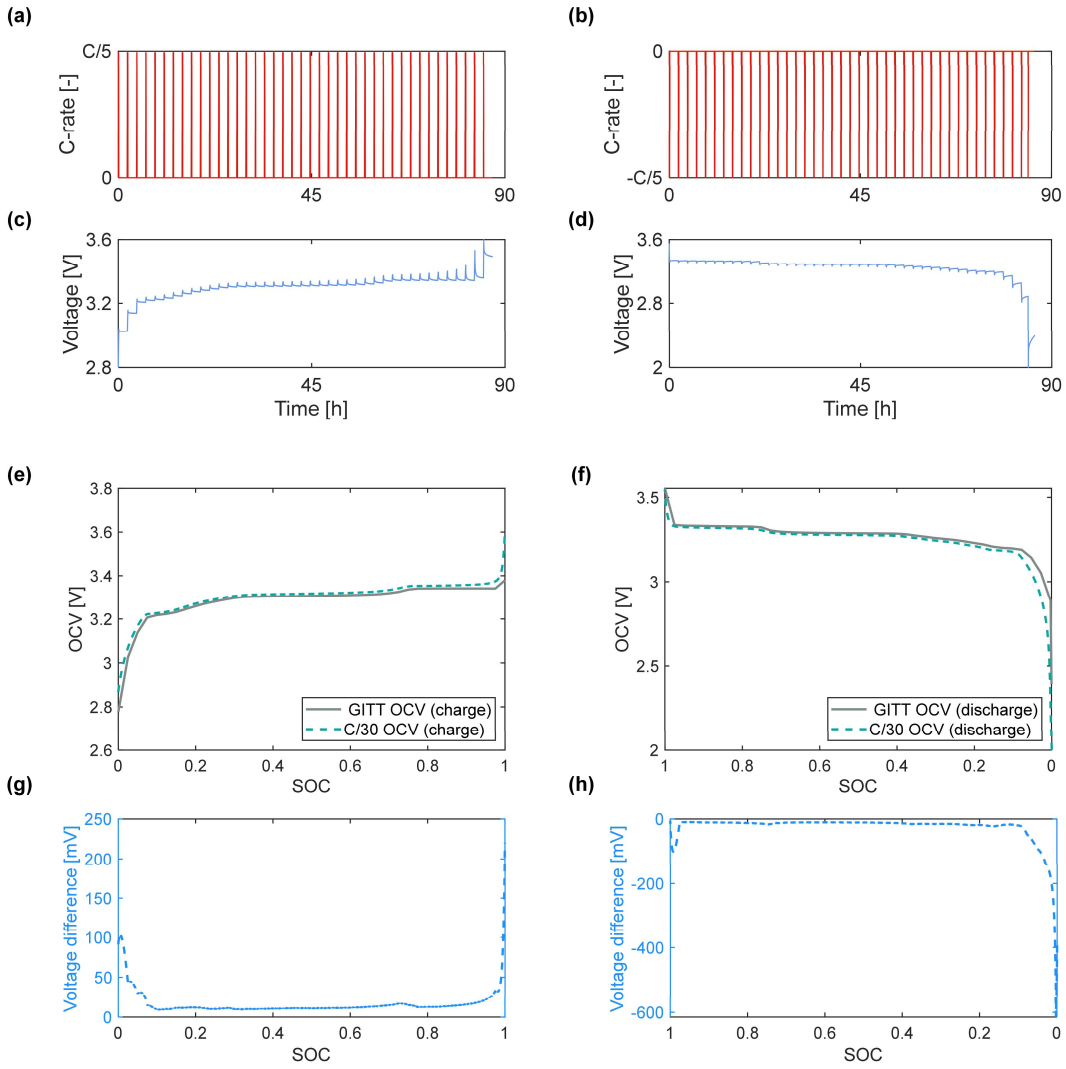

Figure S1. Comparison of pseudo OCV (from C/30) and OCV obtained from GITT test, OCV_{GITT}, at 25°C. (a) and (b) are the current profiles for the GITT test in charge and discharge, respectively. (c) and (d) are the measured voltages of the GITT test in charge and discharge, respectively. (e) and (f) show the pseudo OCV (from C/30) and the OCV_{GITT} test in charge and discharge, respectively. (g) and (h) demonstrate the voltage difference (in mV).

Supplementary Figure S2

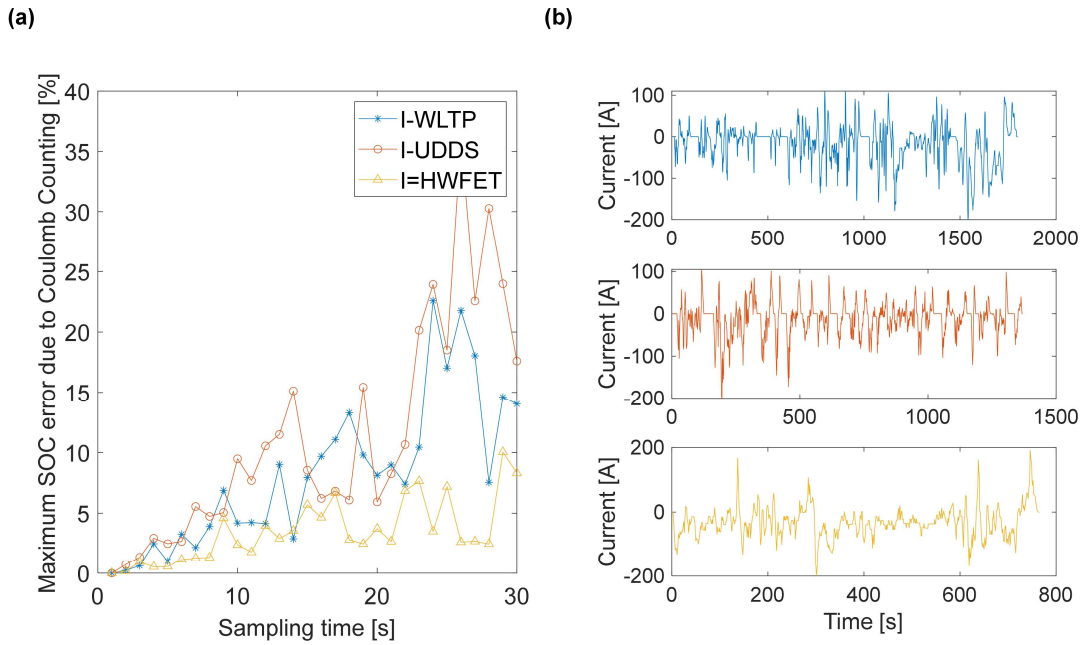

Figure S2 Maximum SOC errors caused by low sampling frequencies for different current profiles. The loading current with 1 Hz is chosen as the reference. Then the sampling time are increased from 1 second to 30 seconds to get the new current profiles. Next, the Coulomb counting method is used for the SOC calculation. Finally, the errors of the SOC under lower sampling frequencies are calculated by comparing with the reference SOC (using 1 Hz reference current profile for Coulomb counting). a) SOC calculation errors with different sampling frequencies for three different loading current profiles, b) reference current profiles with 1 Hz sampling frequency.

Supplementary Figure S3

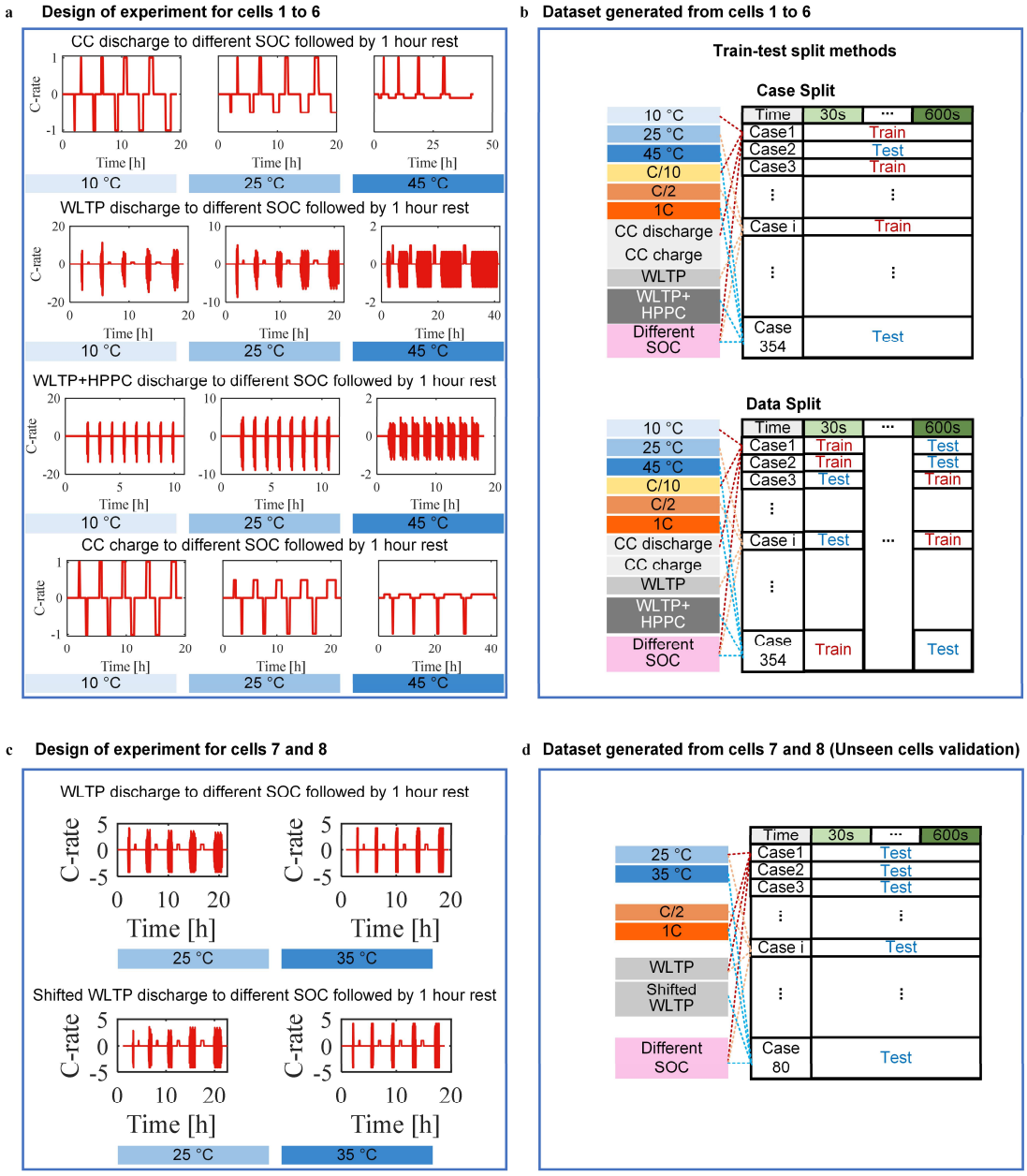

Figure S3: Design of experiment for cells 1 to 6. (a) design of experiment for cells 1 to 6; (b) dataset generated from cells 1 to 6 comprising of 354 cases under different loading profiles, average current rates, SOCs of resting, and temperatures. Two train-test splitting methods are employed: case-based splitting and data-based splitting. (c) design of experiment for cells 7 and 8; (d) dataset generated from cells 7 and 8, which has 80 cases. Cells 7 and 8 were designated as unseen cells, and all 80 generated cases were only used for testing.

Supplementary Figure S4

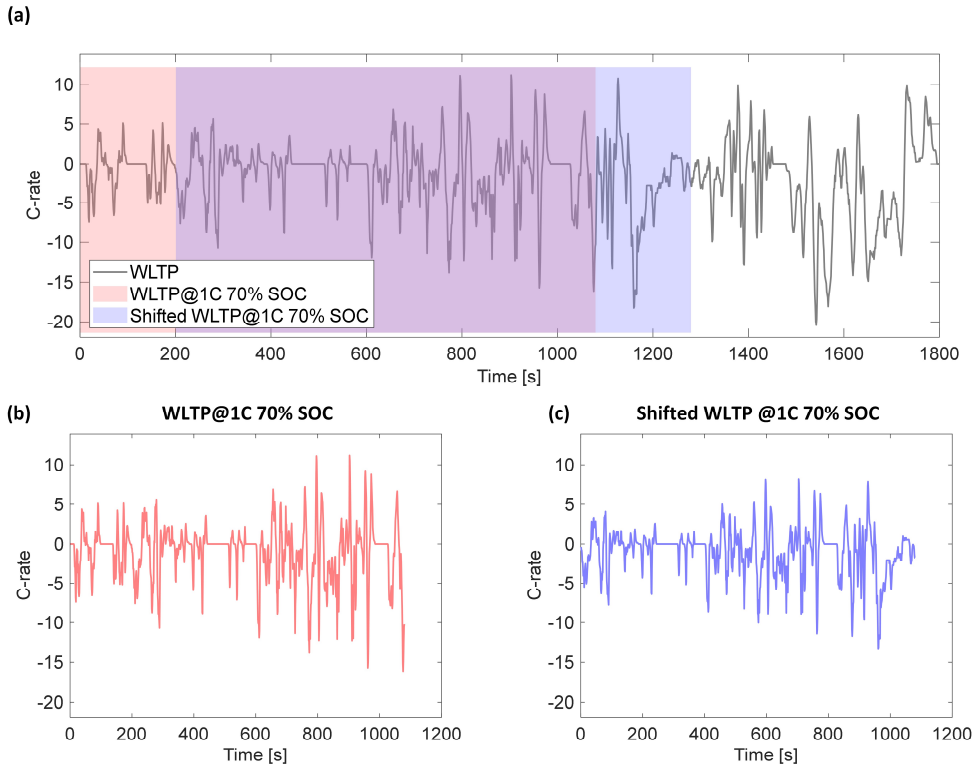

Figure S4. Illustration of WLTP and shifted WLTP profiles (used to discharge Cell 7 and 8) used in this study. (a) The original WLTP profile. (b) WLTP profile used to discharge the cell from 100% SOC to 70% SOC (with 1C average C-rate). The portion of the profile between 0 and 1080 seconds was used. (c) Shifted WLTP profile (with 1C average C-rate) used to discharge the cell 7 and cell 8 from 100% SOC to 70% SOC. The portion of the profile between 200 and 1280 seconds was used.

Supplementary Figure S5

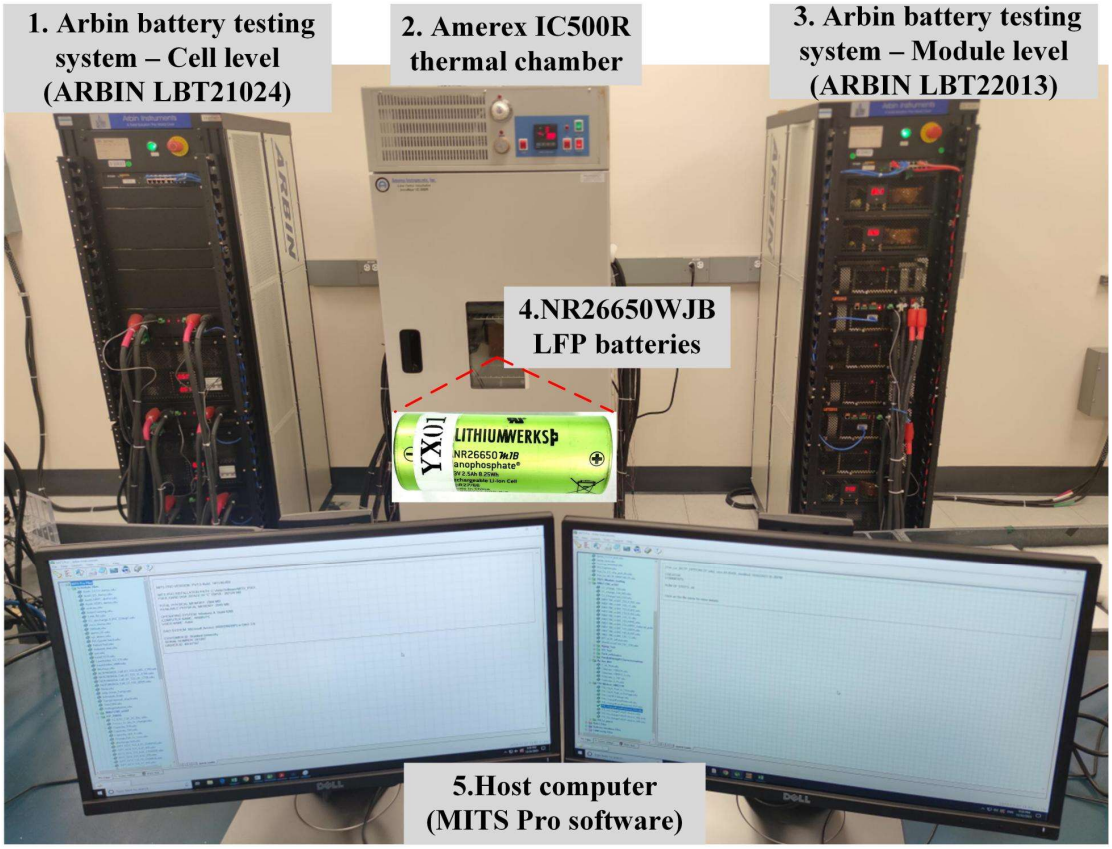

Figure S5: Equipment available at the Stanford Energy Control Lab [1]. 1 Arbin LBT21024, 2 Amerex IC500R thermal chamber, 3 Arbin LBT22013, 4 Host computer (MITS Pro software), 5 NR26650WJB LFP batteries. Lab location: 37.42676975682539, -122.17427385081344.

[1] Catenaro, E., Onori, S., Experimental data of lithium-ion batteries under galvanostatic discharge tests at different rates and temperatures of operation, Data in Brief, Volume 35, 106894, 2021.

Supplementary Figure S6

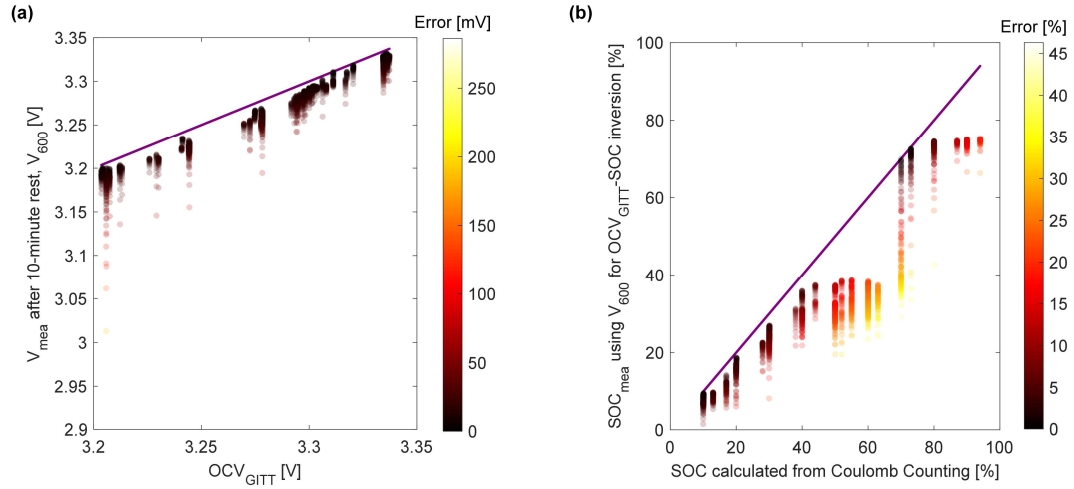

Figure S6: Measured voltage during rest and the observed SOC using the measured voltage for OCV-SOC look-up table checking. a) The measured voltage V_{mea} against OCV_{GITT}. The voltages are the measurements during 10 min rest, the color bars indicate the errors (i.e., OCV_{GITT} - V_{mea}). The OCV_{GITT} is obtained by using SOC calculated from Coulomb counting after discharge/charge the cell for OCV-SOC look-up table inversion. b) The SOC_{mea} against SOC from Coulomb counting, and the errors are calculated by SOC_{mea}- SOC. The SOC_{mea} is obtained by using V_{mea} for OCV-SOC loop-up table checking. The real SOC is obtained based on Colomb counting.

Supplementary Figure S7

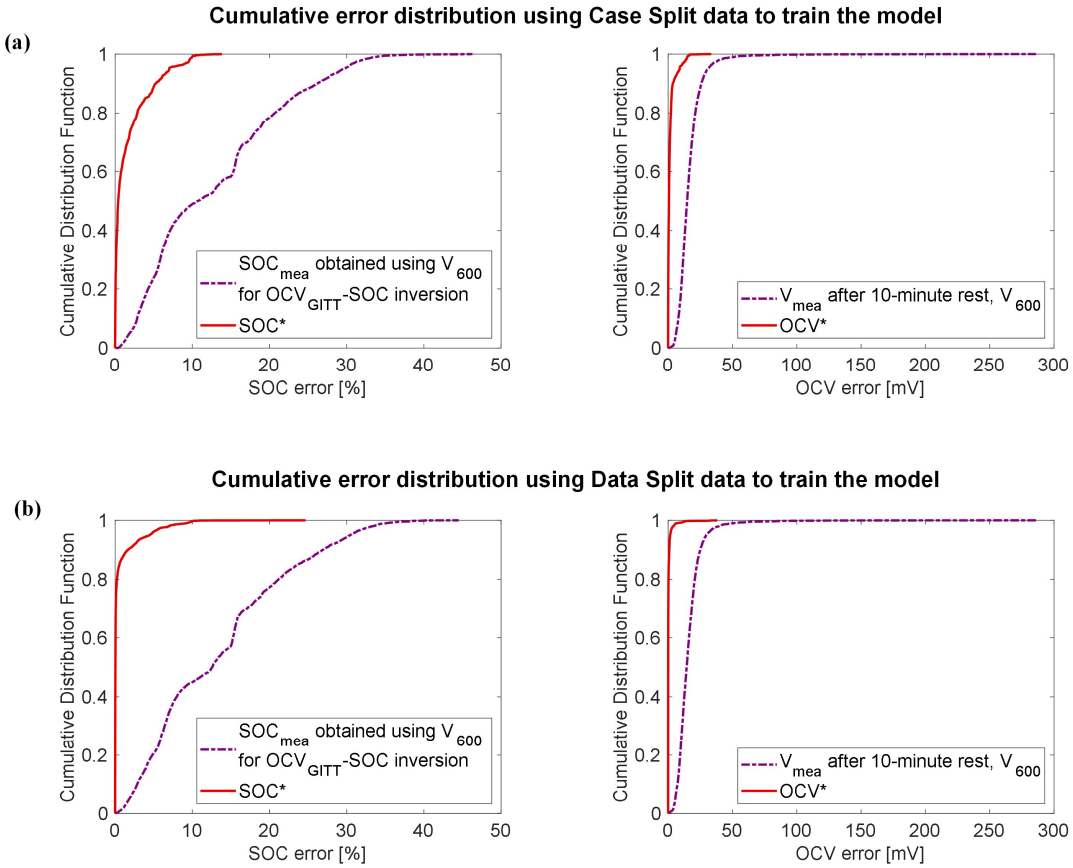

Figure S7. Cumulative distribution function for SOC and OCV errors. V_{mea} is the measured voltage after rest, and SOC_{mea} is the corresponding SOC determined from the OCV_{GITT}-SOC table using V_{mea}. OCV^* and SOC^* are the OCV and SOC predicted by the proposed machine learning method, respectively. (a) *Case Split*, (b) *Data Split*.

Supplementary Figure S8

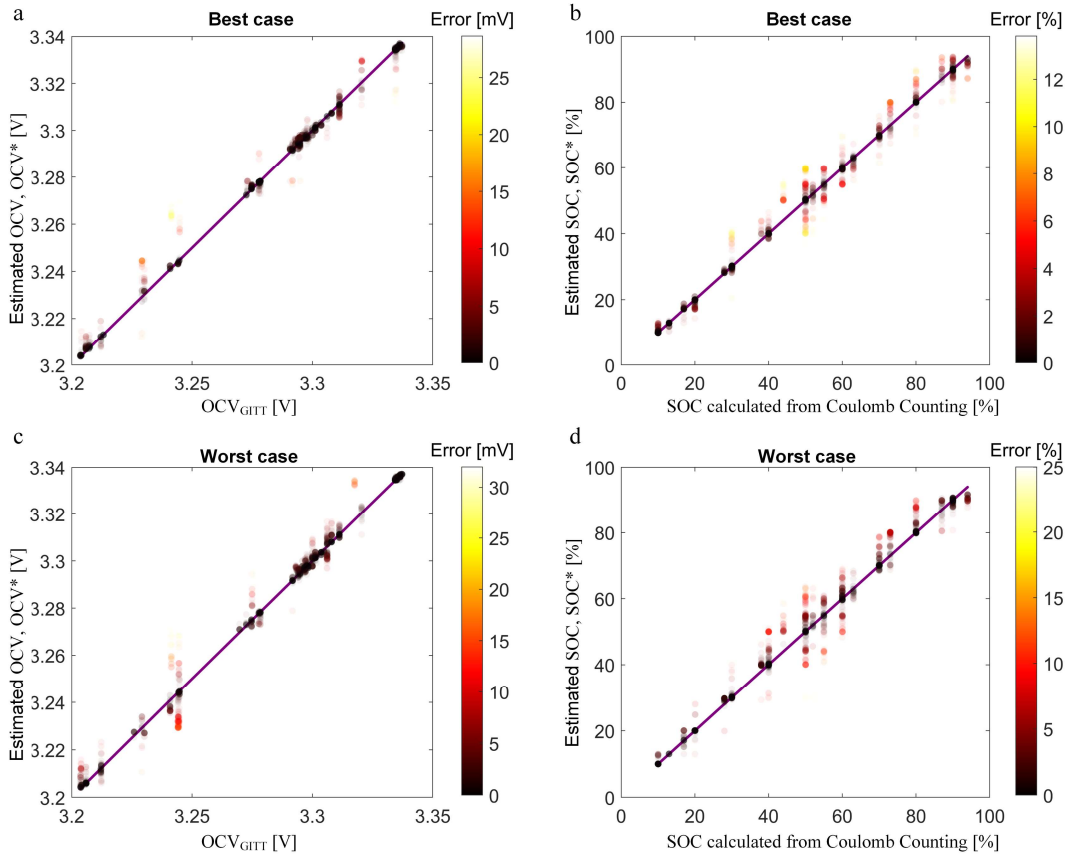

Figure S8: Illustrations of the results with the best performance and worst performance during the 10 times *Case Split* testing. The predicted OCV* and predicted SOC* are obtained based on the proposed machine learning model, and the color bars indicate the prediction errors. The real SOC is obtained based on Colomb Counting after charge/discharge the cell, and the OCV_{GITT} is obtained by using real SOC for OCV_{GITT}-SOC loop-up table inversion. The predicted a) OCV* and b) SOC* in the best case (i.e., smallest error). The predicted c) OCV* and d) SOC* in the worst case (i.e., biggest error).

Supplementary Figure S9

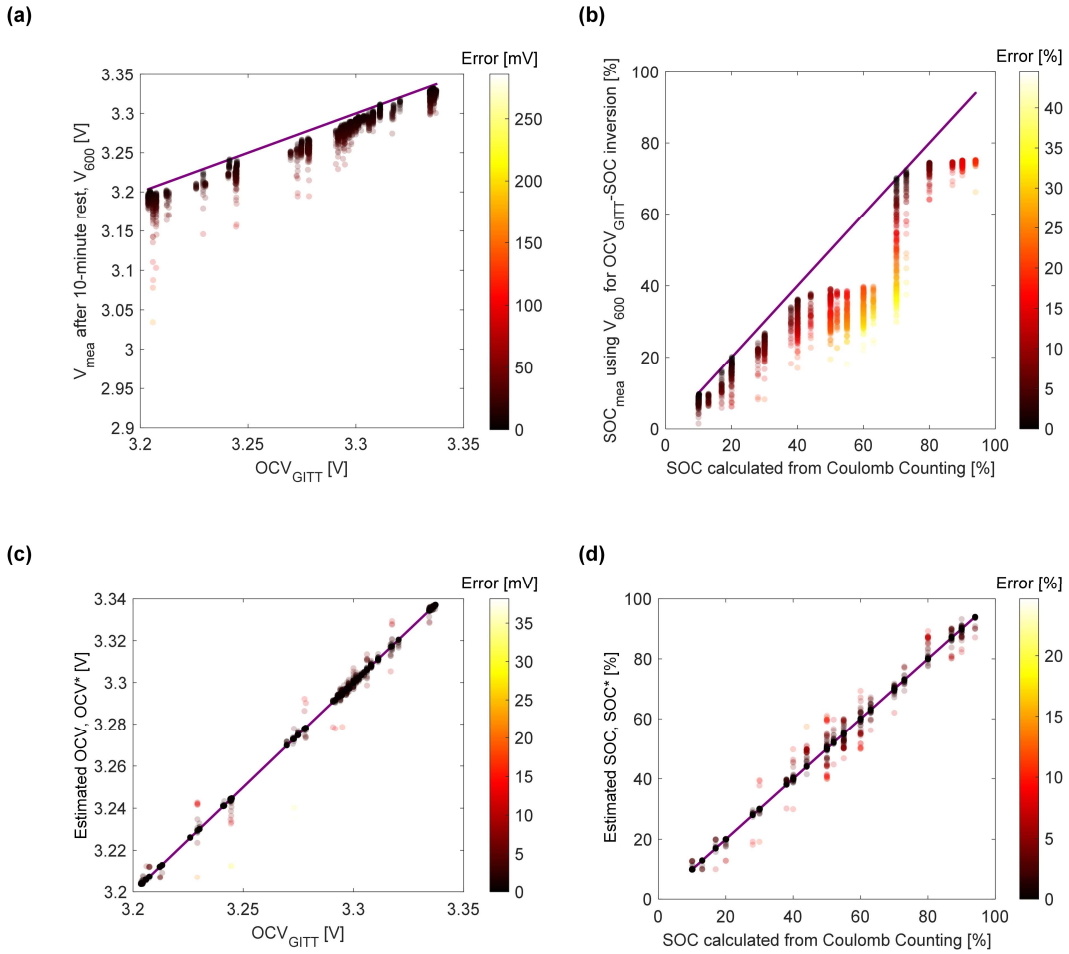

Figure S9: Prediction results with the *Data Split* testing after 10 min relaxation. a) the measured voltage against real OCV. b) observed SOC against real SOC. The results indicate the voltage errors can exceed 250 mV and the SOC_{mea} can have an error larger than 40%. c) the predicted OCV* based on the proposed machine learning model, which keeps most predictions having errors less than 20 mV. d) the predicted SOC* based on the proposed machine learning model, where most errors are less than 10%.

Supplementary Figure S10

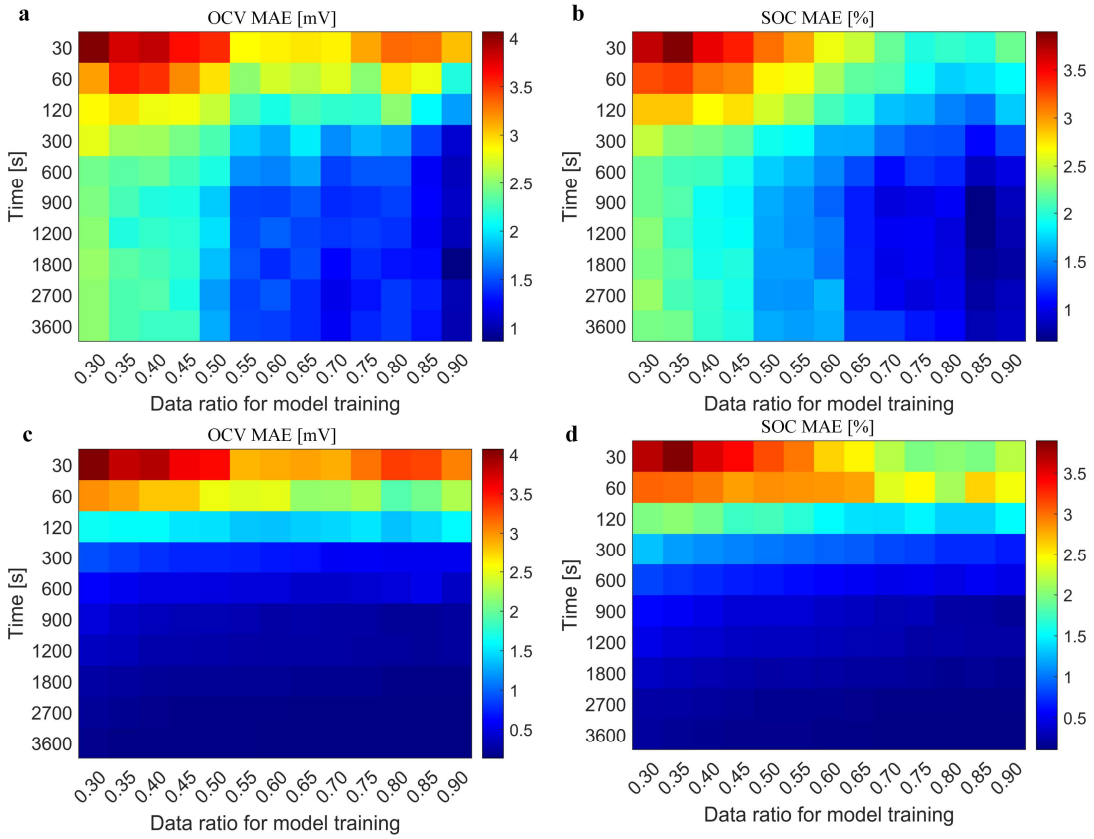

Figure S10. Robustness evaluations. a/b, MAE of calibrated voltage/SOC with different ratios of data for model training and different relaxation times in case random split validations; c/d, MAE of calibrated voltage/SOC with different ratios of data for model training and different relaxation times in data random split validations.

Supplementary Figure S11

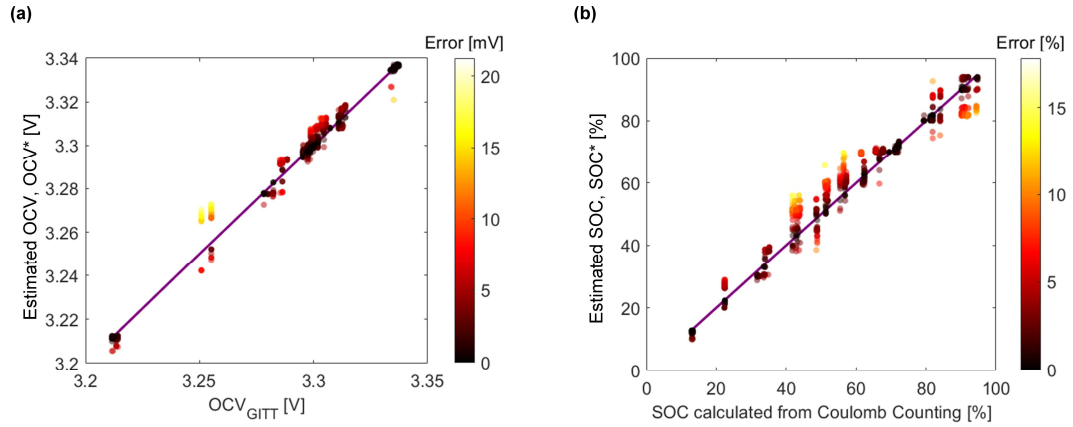

Figure S11 The predicted OCV and SOC for cell 7 and cell 8. The training data are from cells 1-6. (a) the prediction results of OCV*. (b) the prediction results of SOC*. Results indicate all the errors for OCV* are less than 22 mV and those of SOC* are less than 18%.

Supplementary Figure S12

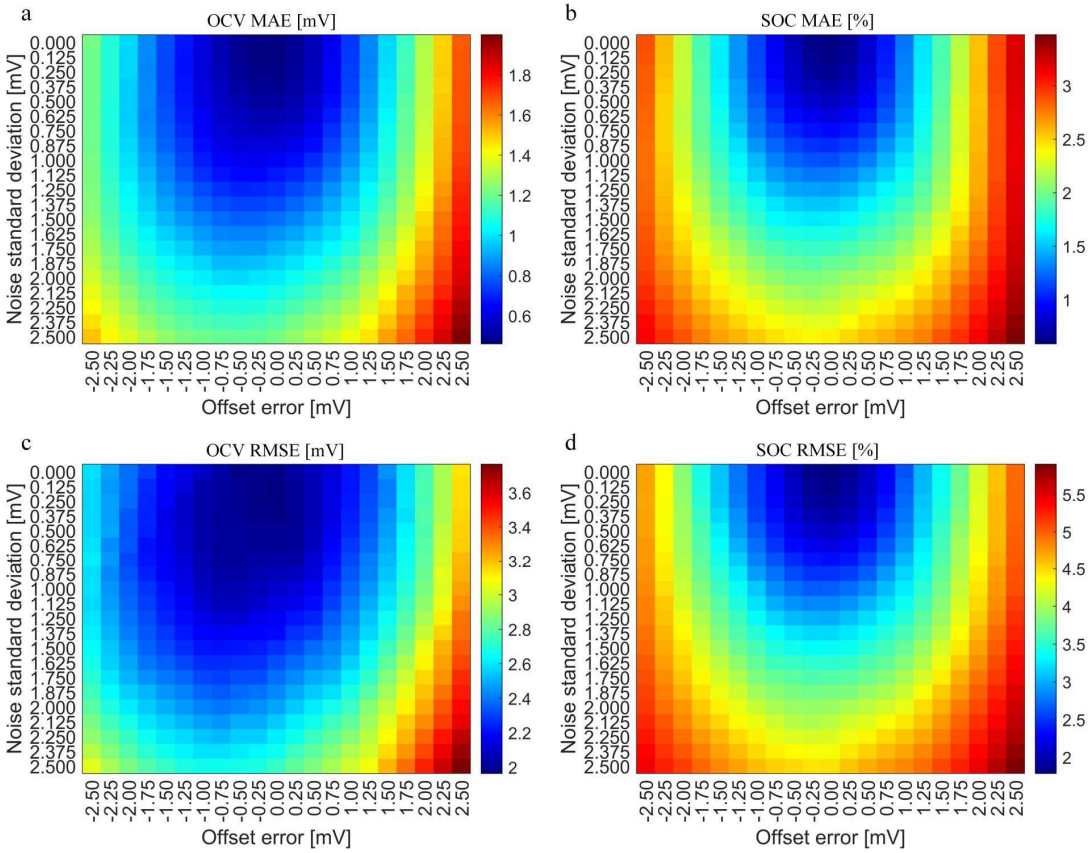

Figure S12. Robustness evaluations under effects of noises. a/b, MAE for the calibrated voltages and SOCs with different offset errors and measurement noises; c/d, RMSE for the calibrated voltages and SOCs with different offset errors and measurement noises.

Supplementary Table S1

Table S 1: Definition of extracted features

Feature set	Feature	Meaning
F_b	Rest time t	Lasting time during voltage relaxation
	Measured voltage V_{mea}	Measured voltage during voltage relaxation
	Estimated SOC_{ob}	SOC obtained by checking the OCV-SOC table using V_{mea}
F_i	Initial voltage V_{10}	Initial voltage during rest, measured after resting the cell for 10 seconds
	Mean current I_m	Mean current of the discharge or charge current
	Resistance R	Calculated as $(V_{30}-V_{10})/I_m$
	Indication of historical current condition I_{flag}	Indication of CC condition (0) or dynamic condition (1)
F_t	Environmental temperature T_{env}	Temperature chamber setting values
	Variation of the surface temperature dT	Temperature variation during voltage relaxation

Supplementary Note 1 Machine learning

This paper proposes a machine learning pipeline for battery SOC resetting during short-term voltage relaxation, the overall flowchart is demonstrated by Figure 3 and 4 in the main manuscript. Experiments are carefully designed and conducted to gather the data under different scenarios. The features are extracted from the measured data during the voltage relaxation to serve as the input of the machine learning pipeline.

Random forest is adopted as the machine learning tool to construct the proposed pipeline. The mapping relationships between input and output Y are learned by a group of trees, each prediction from the trees depicted as:

$$Y_{\mathrm{i}} = h\left(\mathbf{x};\theta_i\right),=1,\cdots M \tag{1}$$

where x represents the observed input (covariate) vector with associated random vector M and the θ_{i} are independent and identically distributed random vectors.

The final predictions of the random forest are obtained by averaging all the tree predictors as:

$$\overline{h}\left(x\right)=\left(\frac{1}{M}\right)\sum_{i=1}^{M}h\left(x;\theta_i\right) \tag{2}$$

For sub-model 1, the input contains the extracted features from the voltage relaxation. The corrections of sub-model 1 are also augmented to the input as features for sub-model 2 to construct an incremental model. Finally, the corrections from the two sub-models are augmented together for the final corrections of the fusion model. The number of trees is selected as 100 in this paper. The 10 folds validation for case random split condition is conducted with different numbers of trees used in the random forest. According to the accuracy on training and testing data and considering the computational burden, 100 is selected as the tree number in this paper.

In the evaluations of the proposed method, two kinds of random split strategies are adopted, which are named as data random split and case random split respectively. For each case, one battery is discharge or charge to different SOC with different current modes and average rates under different temperature. For example, the battery is discharged to 50% SOC using WTLP under 25 ℃ is recognized as one case. The same setting of current and temperature to another SOC or the same loading under another temperature are recognized another case. Therefore, the case random split randomly split all the cases into training and testing datasets. In data random split condition, the data of all the cases are put together firstly. Then, the whole data is randomly split into training and testing datasets according to the preset ratio of data for training and testing. Note that, the data in the testing datasets are never appeared in the training datasets for both of the two split strategies.

Supplementary Note 2 Different testing cases

In this section, we provide detailed information on how different testing cases were configured.

Cell 1-6

We used four loading profiles, three temperature conditions (10 ℃, 25℃, and 45℃), and three average current rates (1C, 0.5C, and 0.1 C) to generate 354 cases using Cell 1-6. The general calculation function of the cases for each loading profile can be described as follows:

$$N = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$

where n_C, n_T, and n_{SOC} denotes the number of average current rates, number of temperature conditions, and number of SOC points, n_{miss} means the missing cases and n_{add} is the additional added cases. The case generation for each loading profile is detailed below.

WLTP: We tested WLTP for three average current rates ($n_C = 3$) and three temperatures ($n_T = 3$) with two groups of SOC points ($n_{SOC} = 10$), i.e., [90%, 70%, 50%, 30%, 10%] and [80%, 60%, 55%, 40%, 20%]. Since we didn't use data from experiments to discharge the cell to 50% SOC with average current rates of 1C and 0.5C under 25℃ and 45 ℃ ($n_{miss} = 2 * 2 = 4$). We repeated the experiment for the SOC [90%, 70%, 50%, 30%, 10%] under 25℃ and 45 ℃ using average C rates of 1C and 0.5 C ($n_{add} = 2 * 2 * 5 = 20$). Therefore, the original WLTP cases contain,

$$N_{\mathrm{WLTP1}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 3 * 3 * 10 - 4 + 20 = 106$$

To verify the effectiveness of our model at other SOC points, we added more experiments using average 1 C and 0.5 C ($n_C = 2$) to discharge the battery until 10 random SOC points at [13%, 94%, 17%, 87%, 28%, 73%, 38%, 63%, 44%, 52%] ($n_{SOC} = 10$) under three temperatures ($n_T = 3$), which include 60 cases,

$$N_{\mathrm{WLTP2}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 2 * 3 * 10 - 0 + 0 = 60$$

Thus, the total number of cases for WLTP is $N_{\mathrm{WLTP}} = N_{\mathrm{WLTP1}} + N_{\mathrm{WLTP2}} = 106 + 60 = 166$

WLTPHPPC: The WLTPHPPC loading profiles with three average current rates ($n_C = 3$) are used to discharge the cell intermittently from 90% SOC to 10% SOC with a 10% SOC interval, i.e., 9 SOC points (n_{SOC}=9), under three temperatures ($n_T = 3$). Also, we didn't conduct experiment to discharge the cell to 10% SOC points with the average 1C current rate under three temperatures ($n_{miss} = 3$). Therefore, the cases are,

$$N_{\mathrm{HPPCWLTP}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 3 * 3 * 9 - 3 = 78$$

CC discharge: As the same as WLTP, i.e., three current rates ($n_C = 3$), three temperature conditions ($n_T = 3$), and ten SOC points ($n_{SOC} = 10$) are used for the case calculation. While some tests are not included. The SOC points at [80%, 60%, 55%, 40%, 20%] using 0.1 C under 10 ℃ and 25℃ are not conducted and the SOC points at [80%, 60%, 55%, 40%, 20%] for all three current rates under 45 ℃ are not conducted. Therefore, the cases with CC discharge are,

$$N_{\mathrm{CC\_discharge}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 3 * 3 * 10 - (1 * 2 * 5 + 1 * 3 * 5) + 0 = 65$$

CC charge: For CC charge test, the 5 SOC points ($n_{SOC} = 5$) at [80%, 60%, 50%, 40%, 20%] using three current rates ($n_C = 3$) under three temperatures ($n_T = 3$) are conducted,

$$N_{\mathrm{CC\_charge}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$

$$= 3 * 3 * 5 - 0 + 0 = 45$$

Therefore, the total cases for Cell 1-6 are 354, which is the sum of all the cases above,

$$N_{\mathrm{C1-C6}} = N_{\mathrm{WLTP}} + N_{\mathrm{HPPCWLTP}} + N_{\mathrm{CC_{discharge}}} + N_{\mathrm{CC_{charge}}}$$
$$= 166 + 78 + 65 + 45 = 354$$

Cell 7-8

To validate the proposed machine learning pipeline for additional unseen conditions, we added more experiments using two new cells, i.e., Cell 7-8. We use WLTP and shifted WLTP with two average current rates (1C and 0.5 C), i.e., $n_C = 2$, under two temperatures (25 ℃ and 35℃), i.e., $n_T = 2$, to discharge the cell until ten different SOC points at [90%, 70%, 50%, 30%, 10%] and [80%, 60%, 55%, 40%, 20%], i.e., $n_{SOC} = 10$. Therefore, 80 cases are generated

$$N_{\mathrm{WLTP}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 2 * 2 * 10 - 0 + 0 = 40$$
$$N_{\mathrm{ShiftWLTP}} = n_C * n_T * n_{SOC} - n_{miss} + n_{add}$$
$$= 2 * 2 * 10 - 0 + 0 = 40$$
$$N_{\mathrm{C7-C8}} = N_{\mathrm{WLTP}} + N_{\mathrm{ShiftWLTP}}$$
$$= 40 + 40 = 80$$

Supplementary Note 3 Experimental procedures

The experiments conducted in this work include three main parts, capacity test, OCV tests, and training/validation tests. Experiments are conducted at the Stanford Energy Control Lab (Figure S5), and the testing temperatures were set to 10℃, 25℃, 35℃ and 45℃.

In this study, battery capacity is determined by fully charging the cell from 0% SOC to 100% SOC using C/5 current (i.e., Q_{ch}) and fully discharging it from 100% SOC to 0% SOC using C/5 current (i.e., Q_{dis}). This test was performed at 10℃, 25℃ and 45℃ for cells 1 to 6, and was performed at 25 ℃ and 35 ℃ for cells 7 and 8. Then, battery capacity is defined as:

$$Q = \frac{Q_{ch} + Q_{dis}}{2} \tag{3}$$

The calculated capacity for all 8 cells at different temperatures is shown in Figure S13.

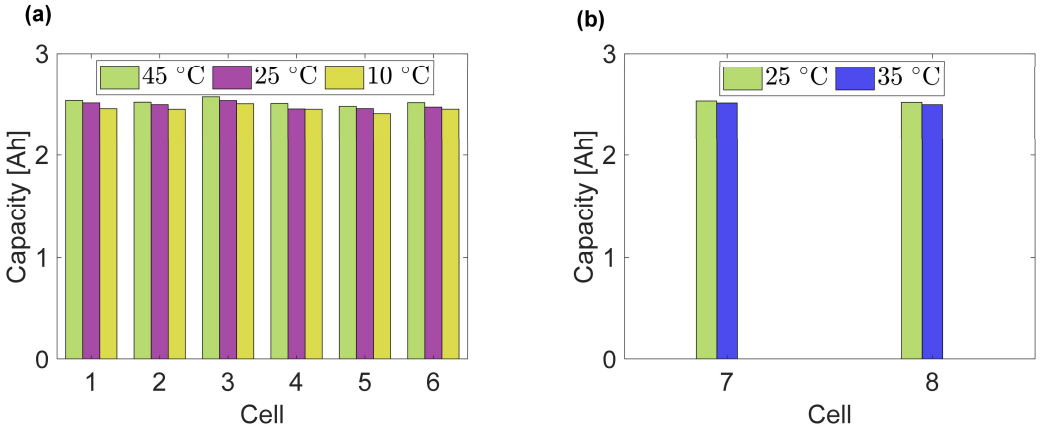

Figure S13. Battery capacity. (a) Cell capacity at 10 ℃, 25 ℃ and 45 ℃ for Cells 1 to 6. (b) Cell capacity at 25 ℃ and 35 ℃ for Cells 7 and 8.

For OCV tests, the main purpose is to obtain the reference OCV-SOC curve, which is used for the SOC estimation. The batteries were firstly fully charged or discharged, and then the C/5 discharge or charge pulses at 2.5% SOC followed by 2-hour rest were repeated to obtain the OCV_{GITT} curves. In addition, the pseudo OCV tests with C/30 was also conducted for comparisons (Figure S1).

In training/validation tests (for Cells 1 to 6), the 1C CC-CV charge protocol is used to fully charge the batteries and rest for 2 hours. The CC discharge and WLTP protocols are used to discharge the batteries down to different SOCs, and then rest for 1 hour to obtain the relaxation data. Different SOC points, three different average current rates (C/10, C/2, and 1C), and three different temperatures (10℃, 25℃ and 45℃) are used in the experiment. HPPC+WLTP test is also conducted in this study. In this test procedure, the battery was firstly fully charged. Then, WLTP protocol was used to discharge the battery to 90% SOC, and rest 1 hour. After that, HPPC test (1 charge and 1 discharge pulse) was conducted. The HPPC+WLTP test was repeated (SOC decreases 10% for each repeat) until SOC reaches 10%. Moreover, the CC charge test was also conducted in this study. The battery was firstly fully discharged with 1C CC to 2 V, and then the voltage was held till discharge current is lower than C/30. Then after 2 hours rest, the batteries were charged to different SOC points and rest for 1 hour. It should be noted that the current rates for WLTP and WLTP+HPPC

represent the average current values.

For unseen cell tests (for Cells 7 and 8), experiments were conducted at 25℃ and 35℃ with average current rates of C/2 and 1C. Both WLTP and shifted WLTP (see Supplementary Figure S4) were used to discharge the cells.

Table S2: Battery specifications

Battery manufactory	Lithiumwerks, China
Battery nominal capacity	2.5 Ah (0.5C rate)
Cathode	LFP
Anode	Graphite
Nominal voltage	3.3 V

Supplementary Note 4 Maximum prediction error:

In the proposed method, features extracted from data after a 30 to 600 second rest period are used as input. While this enhances the feasibility of the proposed method for real-world applications, it also significantly increases the dimensionality of the input features. Since we employed lightweight machine learning algorithms (i.e., random forest regressions), their capacity to handle high-dimensional features is inherently limited. As shown in Fig. 5a and 5b, there are a few outliers in the parity plot. To further enhance prediction accuracy, one possible approach is to extract features specifically from data collected during a fixed rest period (i.e., 30s). In Figure S14, the prediction results of the proposed machine learning method are presented using features extracted from rest times of 30, 60, 120, and 600 seconds, respectively. Compared to the results in Fig. 5b of the main manuscript, the maximum SOC prediction error is reduced to lower than 10%.

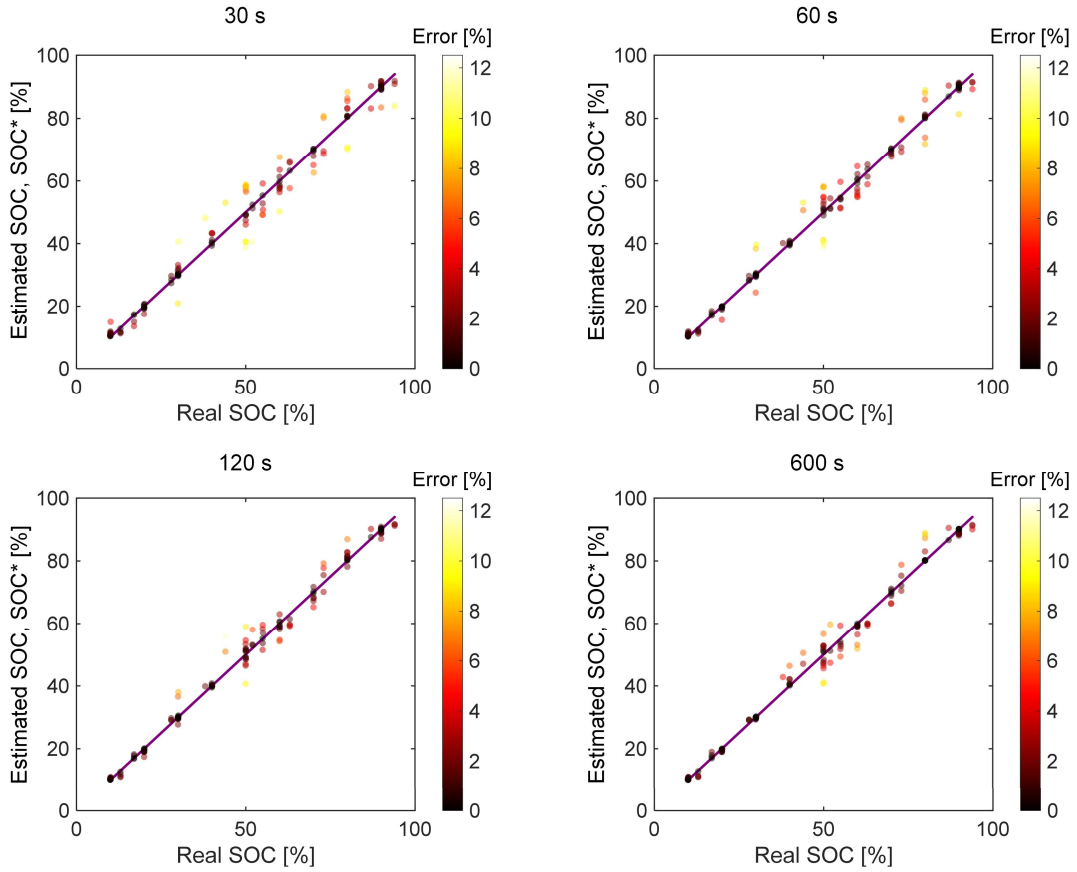

Figure S14. SOC prediction errors using fixed time for feature extraction from rest data: (a) Results using features extracted after 30s rest, (b) Results using features extracted after 60s rest, (c) Results using features extracted after 120s rest, and (d) Results using features extracted after 600s rest.

Supplementary Note 5

$$
\begin{aligned}
RMSE_{SOC} &= \sqrt{\frac{1}{M}\sum_{k=1}^{M}\left(SOC(k)-SOC^{*}(k)\right)^{2}} \\
RMSE_{OCV} &= \sqrt{\frac{1}{M}\sum_{k=1}^{M}\left(OCV(k)-OCV^{*}(k)\right)^{2}} \\
MAE_{SOC} &= \frac{1}{M}\sum_{k=1}^{M}\left(\left|SOC(k)-SOC^{*}(k)\right|\right) \\
MAE_{OCV} &= \frac{1}{M}\sum_{k=1}^{M}\left(\left|OCV(k)-OCV^{*}(k)\right|\right)
\end{aligned}
\tag{4}
$$

where k is the index for the k-th testing data, and M is the number of total testing data. SOC(k) and SOC^{*} (k) are the reference SOC and predicted SOC by the proposed ML model for the k-th testing data, respectively. OCV (k) and OCV* (k) are the reference OCV and predicted OCV for the k-th testing data, respectively.

Supplementary Note 6 Comparison with other machine learning methods

To evaluate the proposed machine learning pipeline with different machine learning algorithms, linear regression (LR), support vector machine (SVR), Gaussian process regression (GPR), LRBoost, and the proposed method are compared. In Figure S15, the left y-axis represents prediction errors and the right y-axis represents the computational burden. The results show ensemble learning based methods (LSBoost) and GPR have better accuracy compared to LR and SVR. Among these models, the proposed method shows the best accuracy and has a medium computational burden.

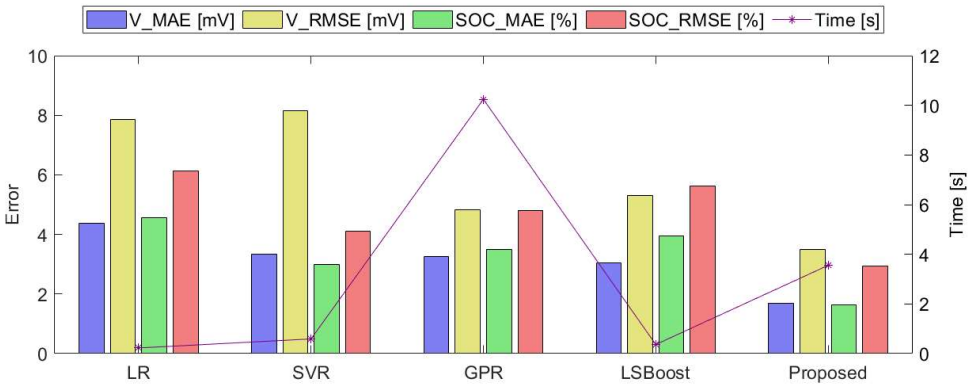

Figure S15. Comparative results of accuracy and computational burden for different ML models.