

## **Supplementary Information for**

### **Long-sequence voltage series forecasting for internal short circuit early detection of Lithium-ion batteries**

*Binghan Cui, Han Wang, Renlong Li, Lizhi Xiang, Jiannan Du, Huaian Zhao, Sai Li, Xinyue Zhao,*

*Geping Yin, Xinqun Cheng, Yulin Ma, Hua Huo, Pengjian Zuo, Guokang Han\*, Chunyu Du\**

MIIT Key Laboratory of Critical Materials Technology for New Energy Conversion and Storage,  
School of Chemistry and Chemical Engineering, Harbin Institute of Technology, Harbin 150001,  
China.

Lead contact

\*Correspondence: cydu@hit.edu.cn (C.D.),

gkhan@hit.edu.cn (G.H.)

## **Supplementary Note 1. Battery specifications and test protocols**

In the normal and ISC test, the batteries charged from 2.75 V in a uniform constant current-constant voltage (CC-CV) mode at different charging rates to 4.2 V, and a current cutoff of C/5 then discharged from 4.2 V at different CC rates, DST conditions, and random tests to 2.75 V. Here 1 C is the required current for fully charging or discharging the nominal capacity in 1 hour. The data of the ISC test when the equivalent ISC resistance of the No. 45 battery is  $30\Omega$  is lost. A LAND CT2001 is employed for battery cycling and voltage, and current collecting. The discharging voltage and charge capacity curves of batteries are shown in Figure S2-6.

## **Supplementary Note 2. Voltage and power sequences prediction, and ISC detection based on deep learning methods**

The performance of the proposed method is benchmarked against four widely used deep learning methods, including convolutional neural network (CNN), deep neural network (DNN), gate recurrent unit neural network (GRU), and long-short-term memory neural network (LSTM). Each method uses the same input data as our proposed method.

The above methods are also implemented in Python 3.9 with Pytorch 1.11. The computation is executed based on a RTX 3090 graphics processing unit. The key hyperparameters of each method are listed in Table S8. CNN and DNN have the best computational efficiency but the worst performance. The computational efficiency and performance of GRU are both bad. LSTM has the best performance in these baseline methods and the worst computational efficiency. Compared with our proposed methods, LSTM has better computational efficiency but worsen performance, which can lead to big ISC detection errors.

Table S8. The model architecture of the baseline methods.

Model architecture	Baseline methods			
	CNN	DNN	GRU	LSTM

<b>First layer for voltage and power prediction (neural number)</b>	1D convolutional layer (1536)	Dense layer (1536)	GRU layer (1536)	LSTM layer (1536)
<b>Second layer for voltage and power prediction (neural number)</b>	1D convolutional layer (262144)	Dense layer (262144)	Dense layer (262144)	Dense layer (262144)
<b>Final layer for voltage prediction (neural number)</b>	1D convolutional layer (512)	Dense layer (512)	Dense layer (512)	Dense layer (512)
<b>Final layer for power prediction (neural number)</b>	1D convolutional layer (1024)	Dense layer (1024)	Dense layer (1024)	Dense layer (1024)
<b>Activation function for voltage and power prediction</b>	Elu	Elu	Elu	Elu
<b>First layer for ISC detection (neural number)</b>	1D convolutional layer (1024)	Dense layer (1024)	GRU layer (1024)	LSTM layer (1024)
<b>Second layer for ISC detection (neural number)</b>	1D convolutional layer (262144)	Dense layer (262144)	Dense layer (262144)	Dense layer (262144)
<b>Final layer for ISC detection (neural number)</b>	1D convolutional layer (480)	Dense layer (480)	Dense layer (480)	Dense layer (480)
<b>Activation function for ISC detection</b>	Sigmoid	Sigmoid	Sigmoid	Sigmoid

### Supplementary Note 3. Introduction of proposed deep learning methods

#### Positional encoding layer

The positional encoding layer is proposed to add position information to the inputs. We first computed the  $\theta$  by:

$$\theta = \frac{\text{position}}{2000^{\text{all\_position}}} \quad (1)$$

Where  $\text{position}$  is the spatial location of the time series,  $\text{all\_position}$  is the length of the time series, and  $\theta$  is the angle that will be computed by the sine and cosine function of the  $\text{position}$ .

Then, compute the position information of encoder inputs and decoder inputs according to the spatial location of voltage, current, and time series, and odd and even positions are calculated using the sine and cosine functions, respectively. The position information is computed by:

$$\text{position}_{\text{enc}} = L_{\text{enc}} * \sin \theta \quad \text{if } \frac{\text{position}}{2} = 0 \quad (2)$$

$$\text{position}_{\text{dec}} = (\frac{L_{\text{enc}}}{2} + L_{\text{pred}}) * \sin \theta \quad \text{if } \frac{\text{position}}{2} = 0 \quad (3)$$

$$\text{position}_{\text{enc}} = L_{\text{enc}} * \cos \theta \quad \text{if } \frac{\text{position}}{2} \neq 0: \quad (4)$$

$$position_{dec} = (\frac{L_{enc}}{2} + L_{pred}) * \cos \theta \quad if \frac{position}{2} \neq 0: \quad (5)$$

Where  $position_{enc}$  and  $position_{dec}$  are the position information of encoder inputs and decoder inputs,  $L_{enc}$  and  $L_{pred}$  are the length of time series of the encoder and the length of prediction results, respectively.

## 1D convolutional layer

The 1D convolutional layer incorporates a few filters, each of which outputs its dot product between the encoder or decoder inputs in a given window and its weights. The window called kernel size moves along the input to condense information. Detailed information on the 1D convolutional layer used in the encoder part and decoder part is presented in Table S14.

## Multi-head attention layer

The multi-head attention layer is based on self-attention. The self-attention is defined based on the tuple inputs, such as query, key, and value, which perform the scaled dot-product as shown in Figure S21A. The scaled dot product is defined as

$$\mathcal{A}(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (6)$$

Where  $Q$ ,  $K$ , and  $V$  are the matrix in which the first dimension is the batch size, the second dimension is the length of the time series, the third dimension is the number of *heads*, and the fourth dimension is the result of dividing  $d$  by *heads*.  $K^T$  and  $d$  are the transpose of  $K$  and the input dimension. *Softmax* is the function that outputs the probability distribution of samples. The self-attention can capture the long-term interdependent features better in the time series compared with RNN, GRU, and LSTM, which can improve the prediction accuracy. In our study, the  $d$  and *heads* are 512 and 8, respectively.

## Normalization

To avoid gradient vanishing, and overfitting, and improve the training speed, we use batch normalization and layer normalization methods during the training process of the proposed method. The batch normalization normalizes the

samples of each batch size in each neuron, and the layer normalization normalizes all neurons of each sample in each layer. In both batch normalization and layer normalization, the mean and standard deviation of inputs are 0 and 1, respectively.

### **Residual connections**

Residual connections connect the outputs of multiple layers including the multi-head attention layer, 1D convolutional layer, and so on, and inputs (Figure S21D) to prevent the gradient vanishing, which can enhance the performance of the proposed methods. The residual connections are represented as “Add&Norm” in the flowchart of our proposed methods.

### **Encoder layer**

The components of the encoder layer are shown in Figure S21B. The detailed information on two 1D convolutional layers is presented in Table S14.

### **Conv layer**

The components of the conv layer are shown in Figure S21C. Detailed information on the 1D convolutional layer is presented in Table S14. The maximum pooling layer is incorporated following batch normalization. The maximum element is extracted from its input over a given pool size of 3. We use the maximum pooling layer to privilege the superior ones with dominating features and make a focused self-attention feature map in the next layer.

### **Early stopping strategy**

After calculating the training loss and the validation loss. The model parameters will continue to be updated when the training loss and the validation loss simultaneously decrease. But stop to be updated when the train loss decreases and the validation loss increases.

### **Elu activation**

Elu activation is used in our proposed methods to enhance the nonlinear computing capability. The elu

activation is defined as

$$Elu(x) = e^x - 1 \quad x < 0 \quad (7)$$

$$\{Elu(x) = x \quad x > 0 \quad (8)$$

Where  $x$  is the input. When  $x$  is less than 0, we use Eq. 7 for calculation, otherwise we use Eq. 8 for calculation.

### Sigmoid activation

Sigmoid activation is used in our proposed methods to enhance the nonlinear computing capability and output the probability of the ISC occurrence. The sigmoid activation is defined as

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

Where  $x$  is the input.

## Supplementary Figures

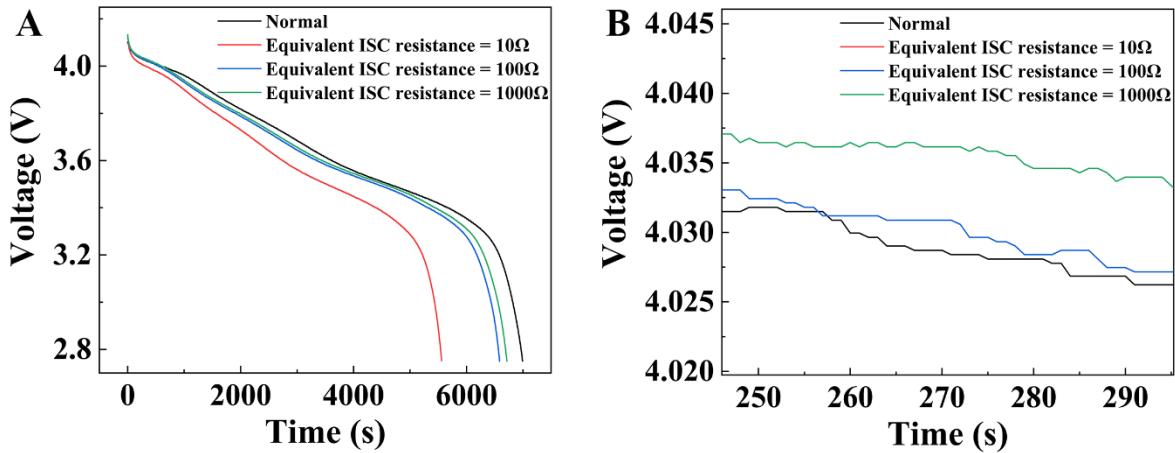


Figure S1. (A) Discharge voltage over the entire SOC range of NCM811 batteries. (B) Discharge voltage from 245 s to 295 s of NCM811 batteries.

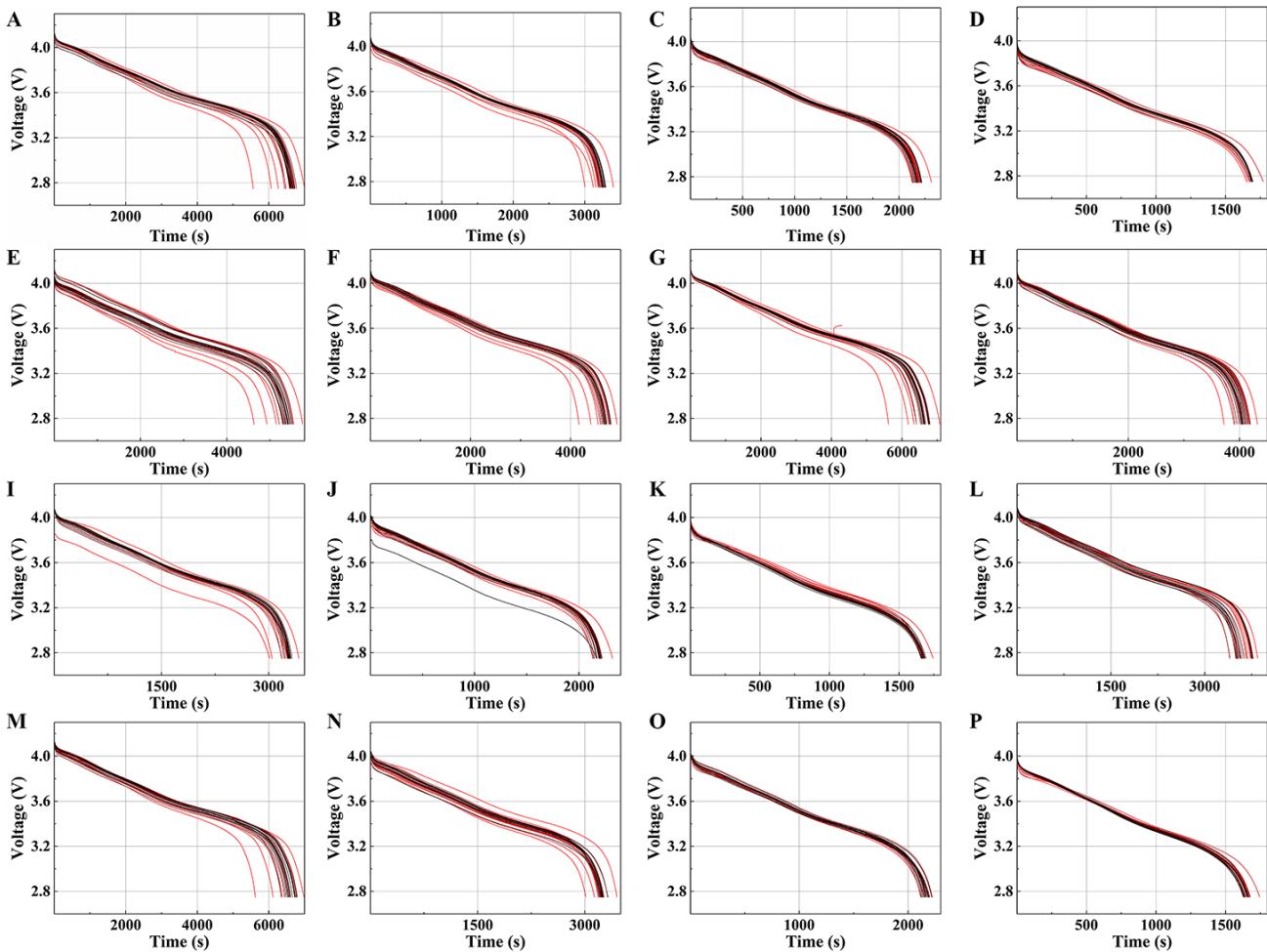


Figure S2. Battery discharge voltage curves versus cycle number of No. 1 battery (A), No. 2 battery (B), No. 3 battery (C), No. 4 battery (D), No. 5 battery (E), No. 6 battery (F), No. 7 battery (G), No. 8 battery (H), No. 9

battery (I), No. 10 battery (J), No. 11 battery (K), No. 12 battery (L), No. 13 battery (M), No. 14 battery (N), No. 15 battery (O), and No. 16 battery (P).

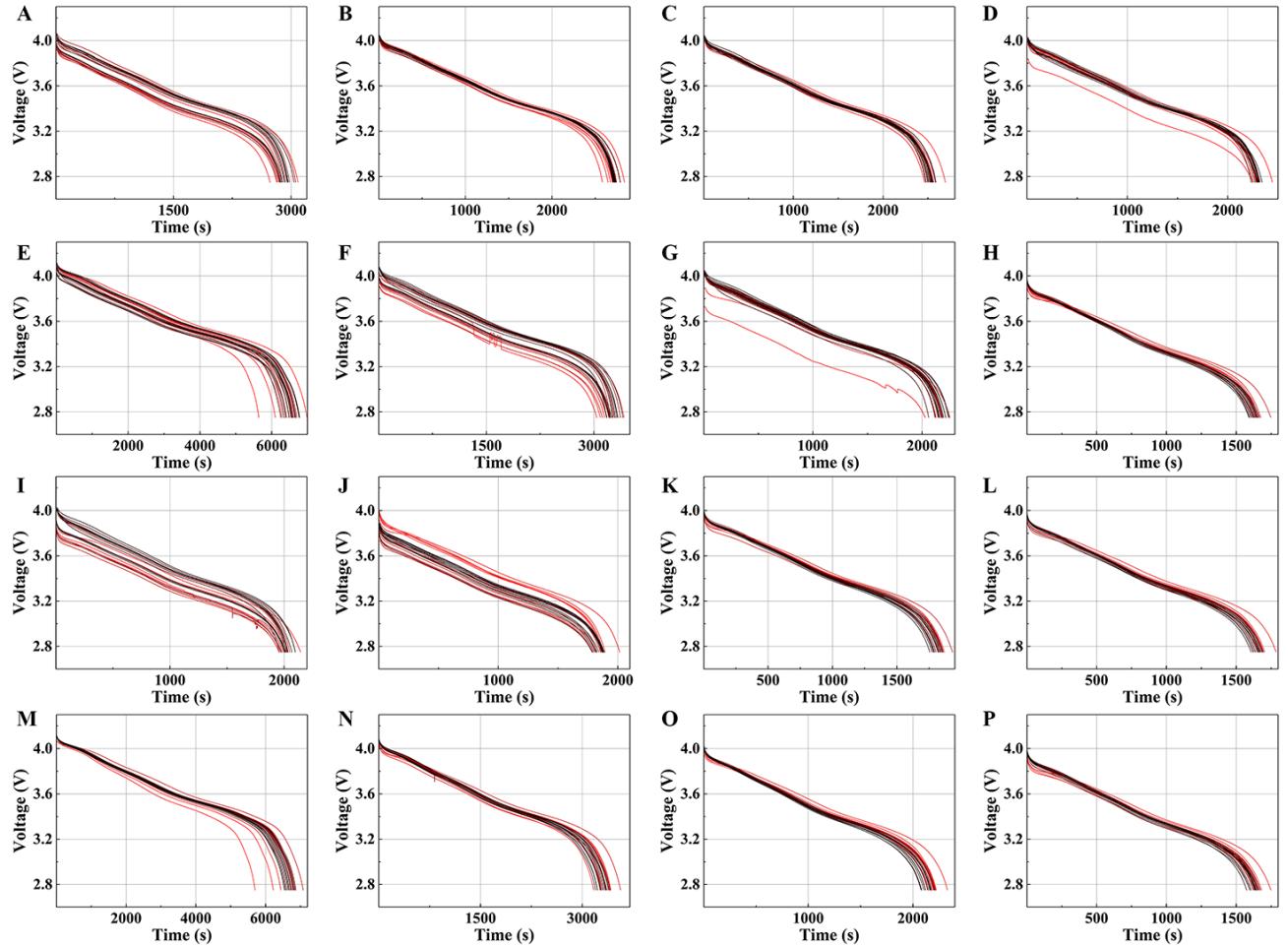


Figure S3. Battery discharge voltage curves versus cycle number of No. 17 battery (A), No. 18 battery (B), No. 19 battery (C), No. 20 battery (D), No. 21 battery (E), No. 22 battery (F), No. 23 battery (G), No. 24 battery (H), No. 25 battery (I), No. 26 battery (J), No. 27 battery (K), No. 28 battery (L), No. 29 battery (M), No. 30 battery (N), No. 31 battery (O), and No. 32 battery (P).

Figure S4. Battery discharge voltage curves versus cycle number of No. 33 battery (A), No. 34 battery (B), No. 35 battery (C), No. 36 battery (D), No. 37 battery (E), No. 38 battery (F), No. 39 battery (G), No. 40 battery (H), No. 41 battery (I), No. 42 battery (J), No. 43 battery (K), No. 44 battery (L), No. 45 battery (M), No. 46 battery (N), No.

47 battery (O), and No. 48 battery (P).

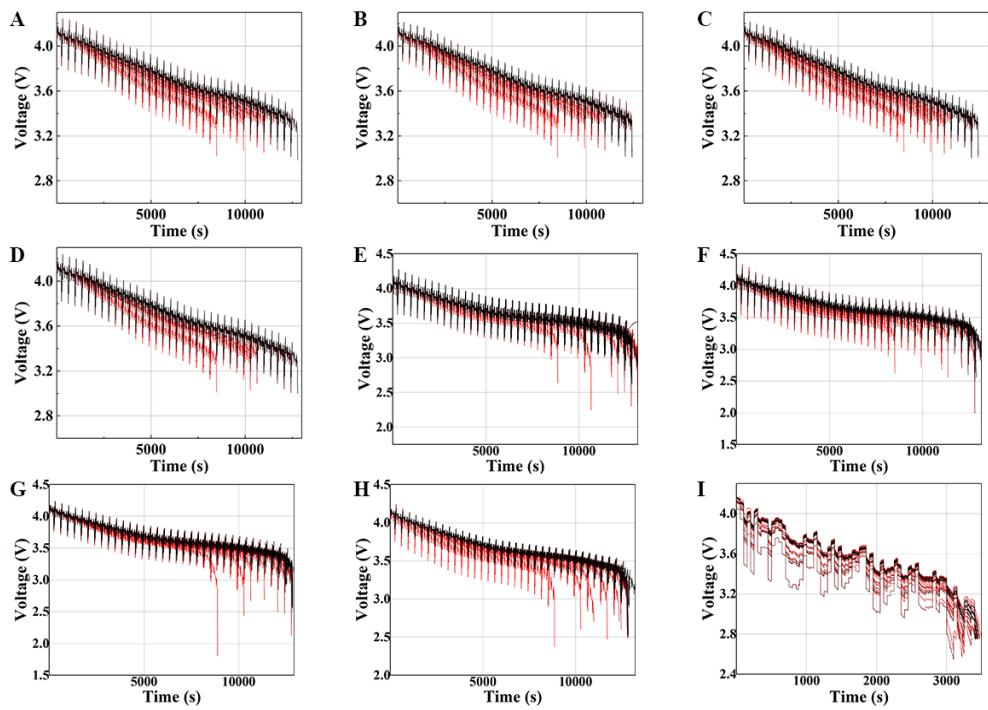


Figure S5. Battery discharge voltage curves versus cycle number of No. 49 battery (A), No. 50 battery (B), No. 51 battery (C), No. 52 battery (D), No. 53 battery (E), No. 54 battery (F), No. 55 battery (G), No. 56 battery (H), and No. 57 battery (I).

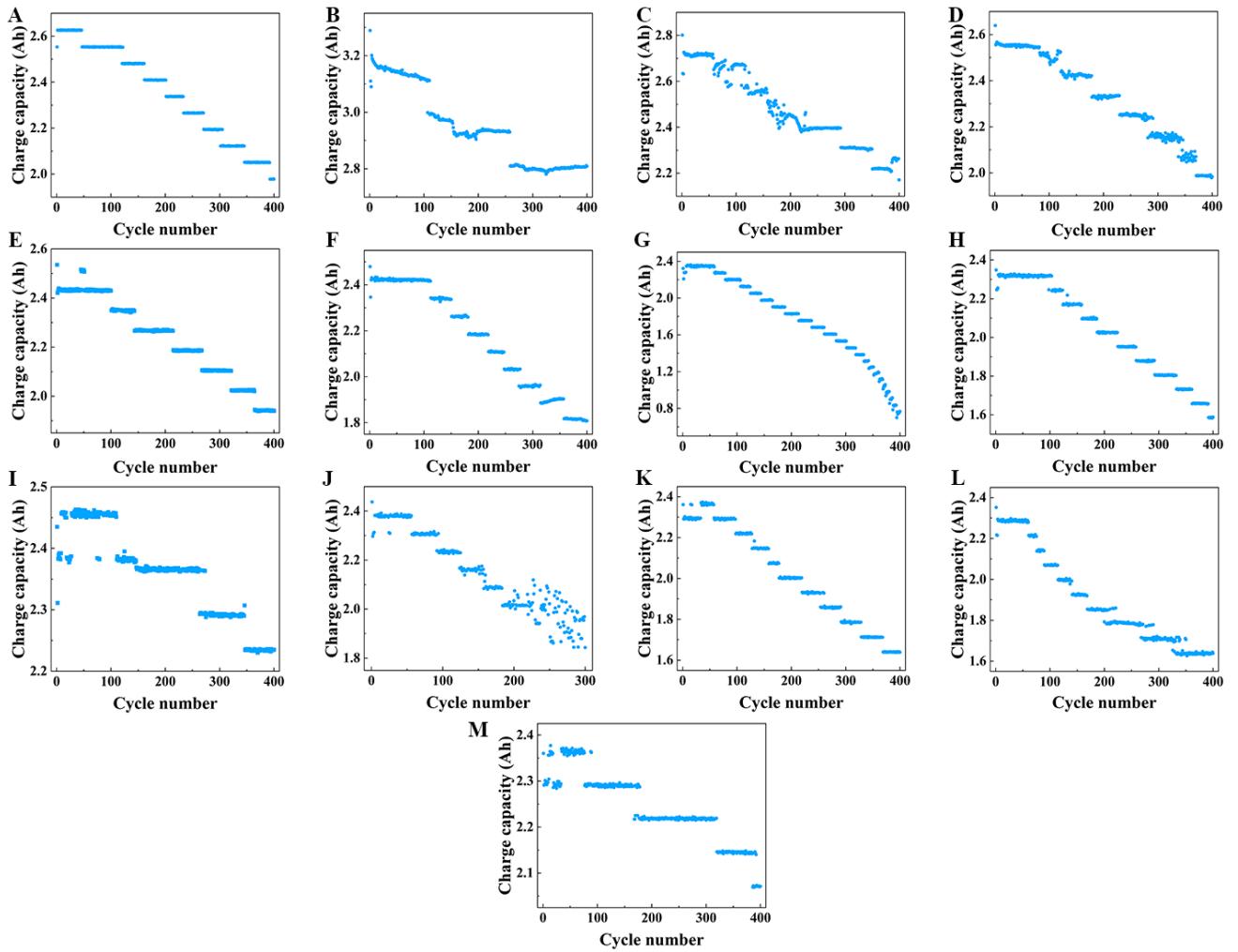


Figure S6. Battery charge capacity curves versus cycle number of No. 58 battery (A), No. 59 battery (B), No. 60 battery (C), No. 61 battery (D), No. 62 battery (E), No. 63 battery (F), No. 64 battery (G), No. 65 battery (H), No. 66 battery (I), No. 67 battery (J), No. 68 battery (K), No. 69 battery (L), and No. 70 battery (M).

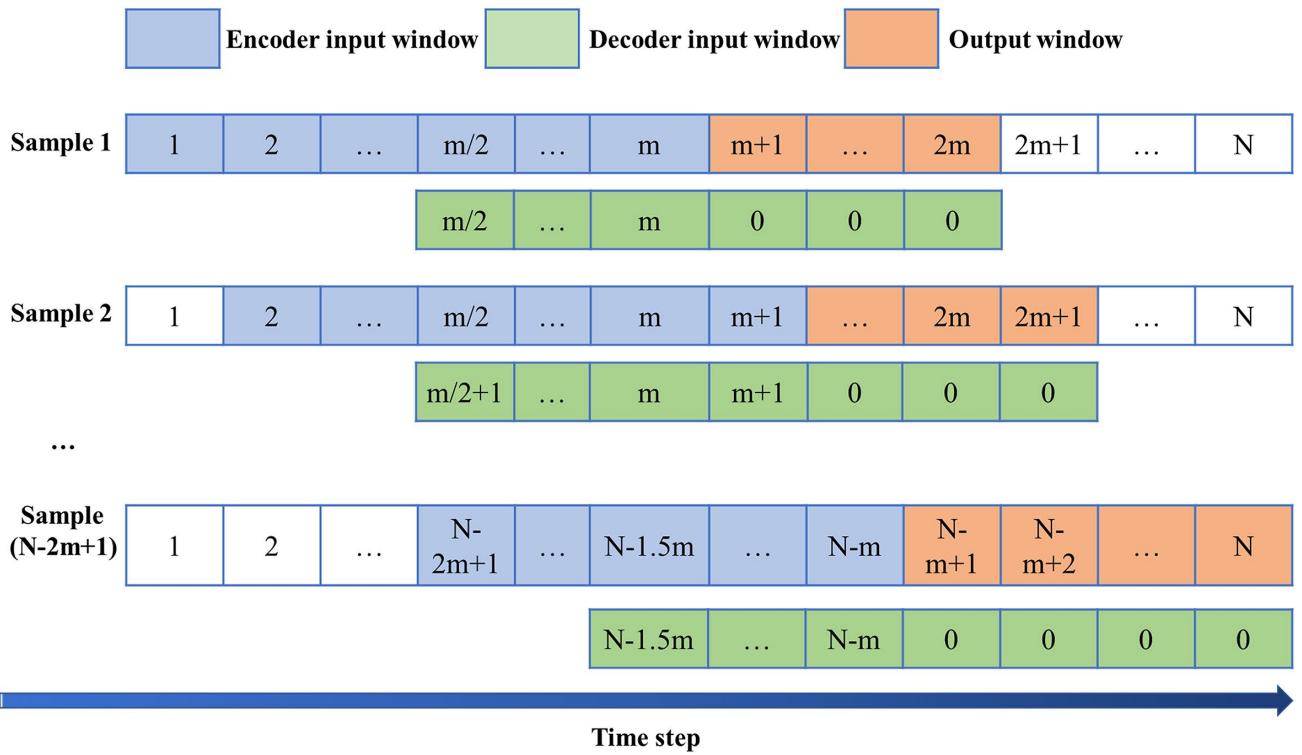


Figure S7. Sampling the encoder input, decoder input, and output from the voltage, current, and time series.

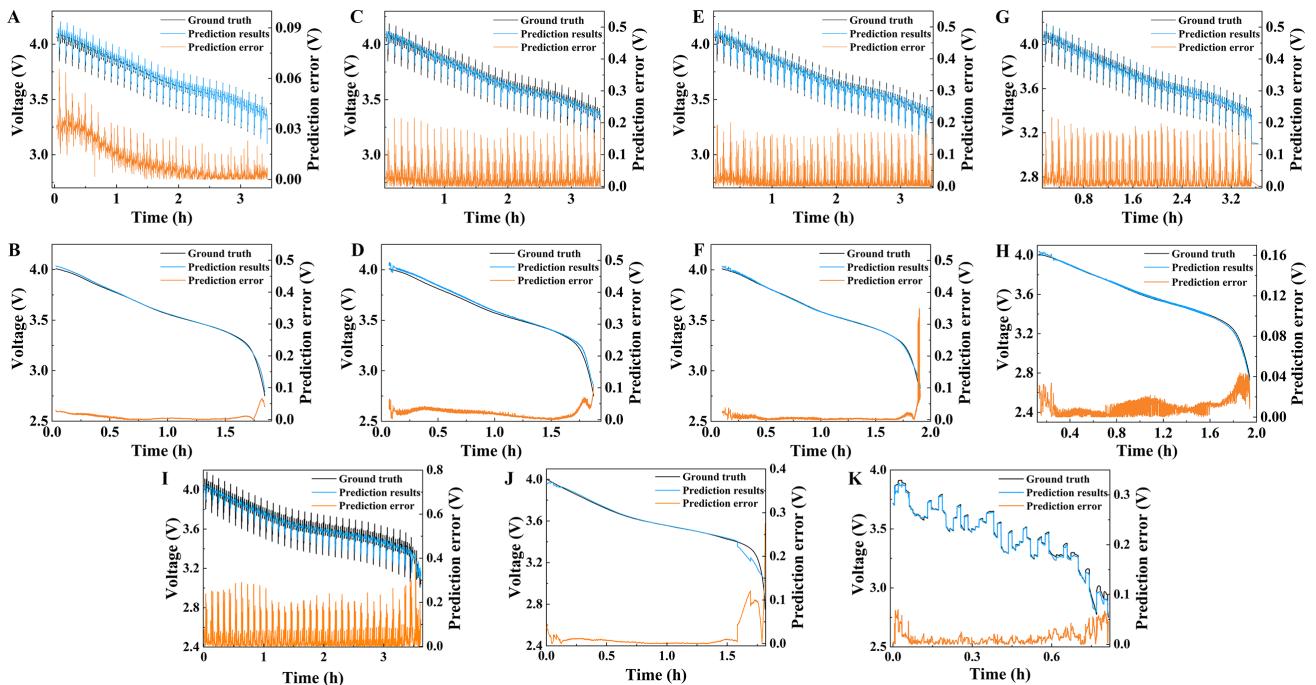


Figure S8. Examples of discharge voltage prediction results in which the cycle number is 1 for the No. 37 battery

with the output of 2 min (A), 4 min (C), 6 min (E), and 8 min (G) voltage sequences. Examples of discharge

voltage prediction results in which the cycle number is 1 for the No. 1 battery with the output of 2 min (B), 4 min (D), 6 min (F), and (H) voltage sequences. Examples of discharge voltage prediction results with the output of 8 min voltage sequences in which the cycle number is 1 for the No. 53 (I), No. 33 (J), and No. 57 (K) battery.

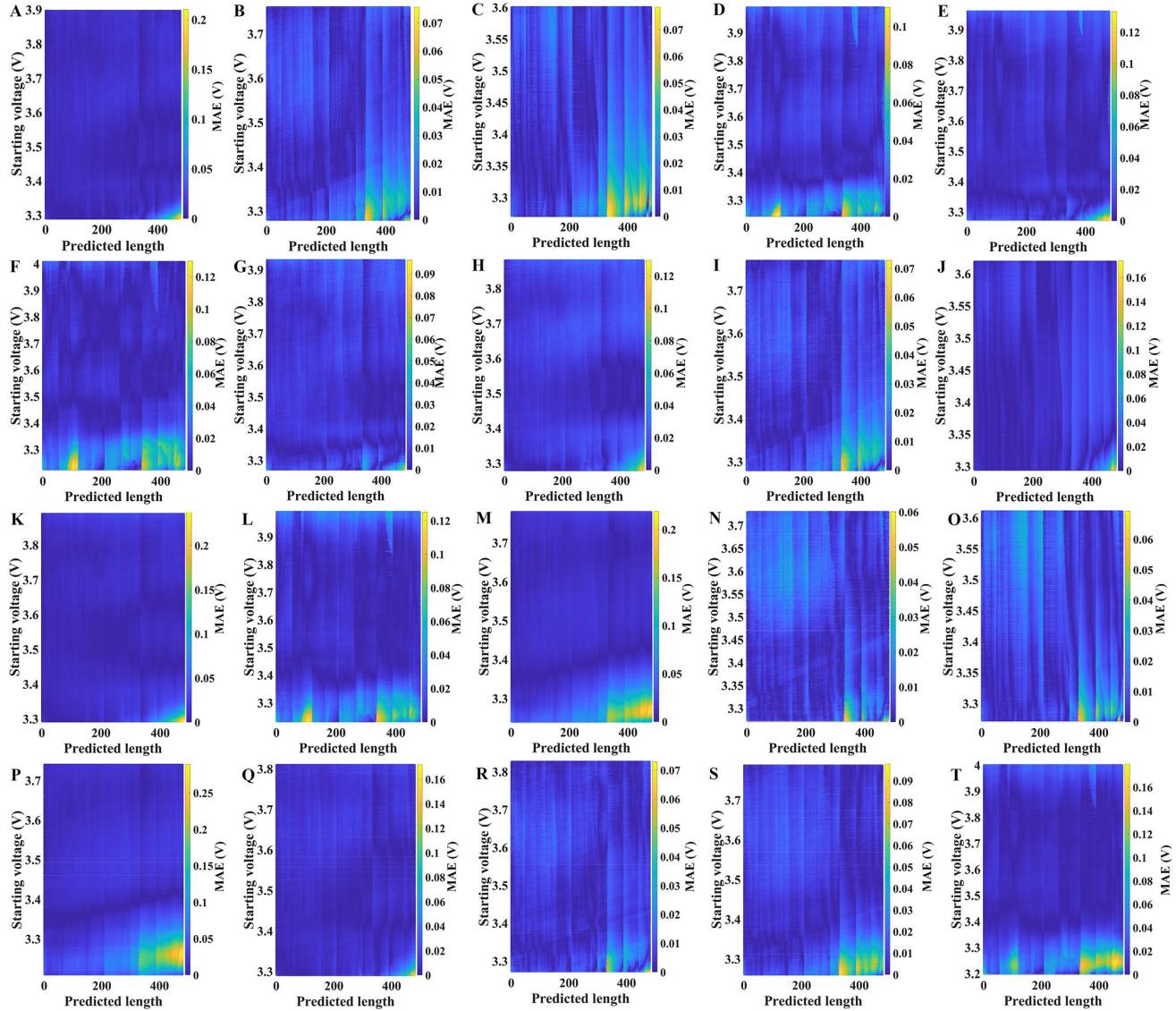


Figure S9. Voltage prediction MAE of normal test with the output of 8 min voltage sequences for No. 2 battery (A), No. 3 battery (B), No. 4 battery (C), No. 5 battery (D), No. 6 battery (E), No. 7 battery (F), No. 8 battery (G), No. 9 battery (H), No. 10 battery (I), No. 11 battery (J), No. 12 battery (K), No. 13 battery (L), No. 14 battery (M), No. 15 battery (N), No. 16 battery (O), No. 17 battery (P), No. 18 battery (Q), No. 19 battery (R), No. 20 battery (S), and

No. 21 battery (T).

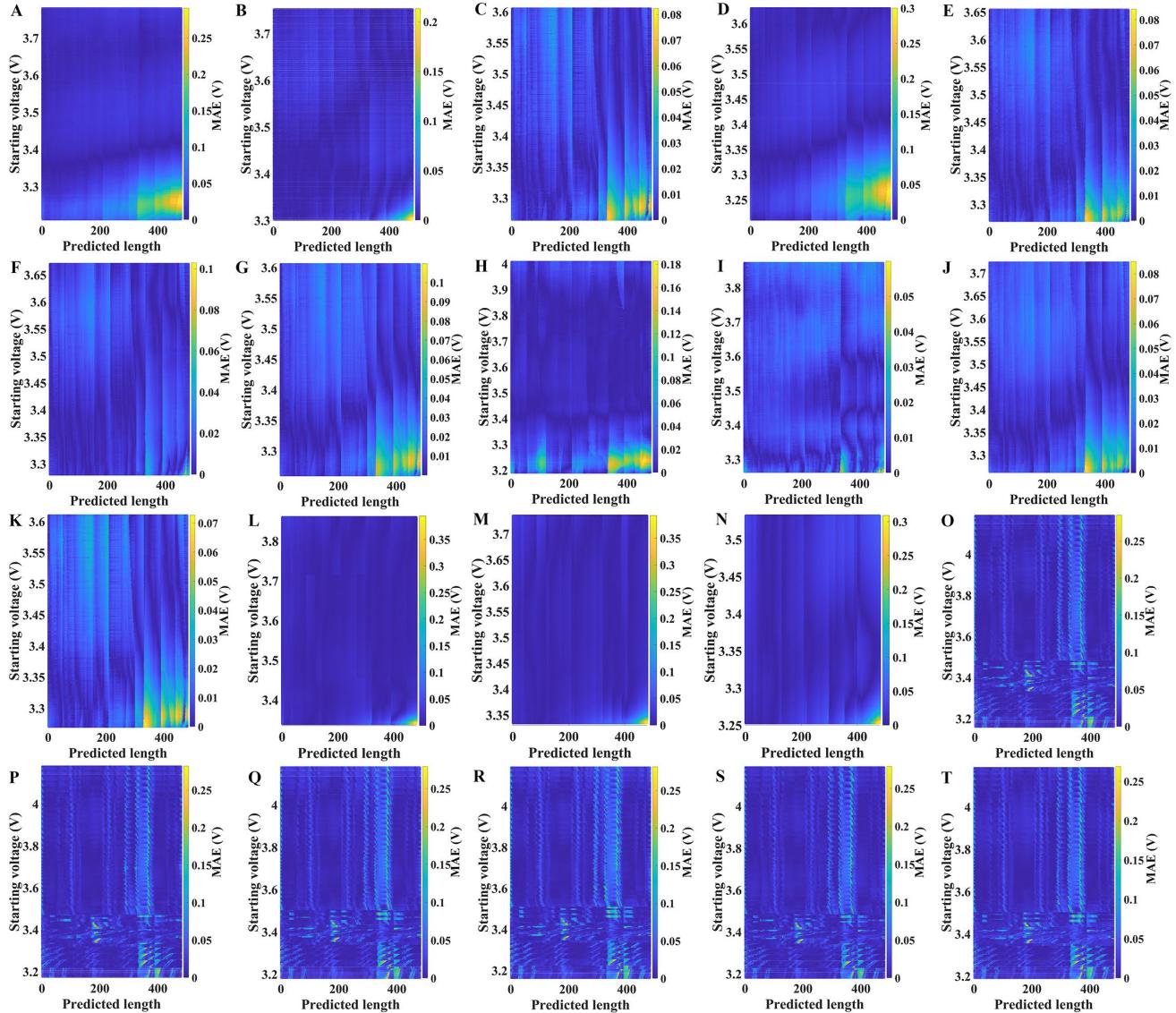


Figure S10. Voltage prediction MAE of normal test with the output of 8 min voltage sequences for No. 22 battery  
 (A), No. 23 battery (B), No. 24 battery (C), No. 25 battery (D), No. 26 battery (E), No. 27 battery (F), No. 28  
 battery (G), No. 29 battery (H), No. 30 battery (I), No. 31 battery (J), No. 32 battery (K), No. 34 battery (L), No. 35  
 battery (M), No. 36 battery (N), No. 37 battery (O), No. 38 battery (P), No. 39 battery (Q), No. 40 battery (R), No.  
 41 battery (S), and No. 42 battery (T).

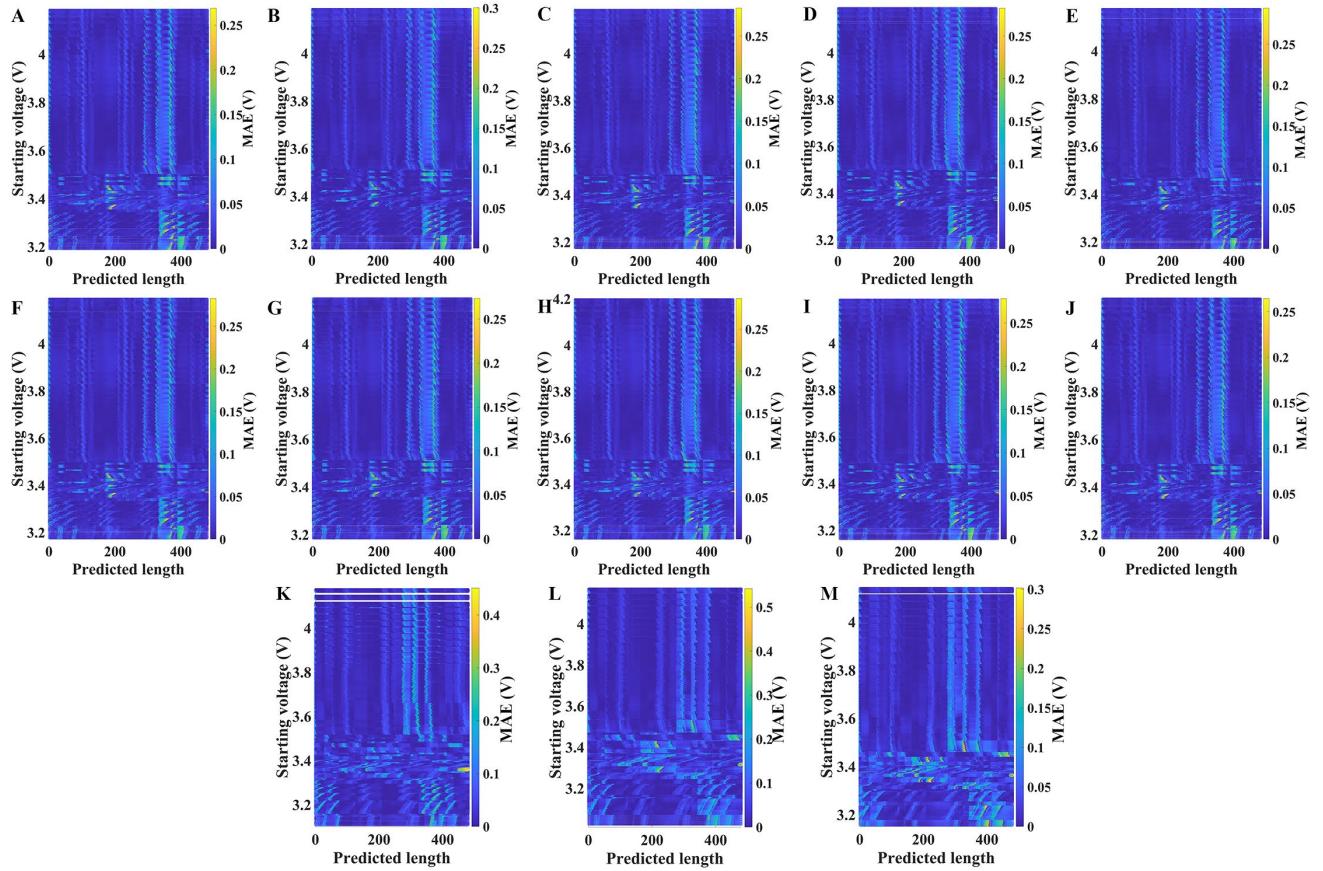


Figure S11. Voltage prediction MAE of normal test with the output of 8 min voltage sequences for No. 43 battery  
 (A), No. 44 battery (B), No. 45 battery (C), No. 46 battery (D), No. 47 battery (E), No. 48 battery (F), No. 49  
 battery (G), No. 50 battery (H), No. 51 battery (I), No. 52 battery (J), No. 54 battery (K), No. 55 battery (L), and  
 No. 56 battery (M).

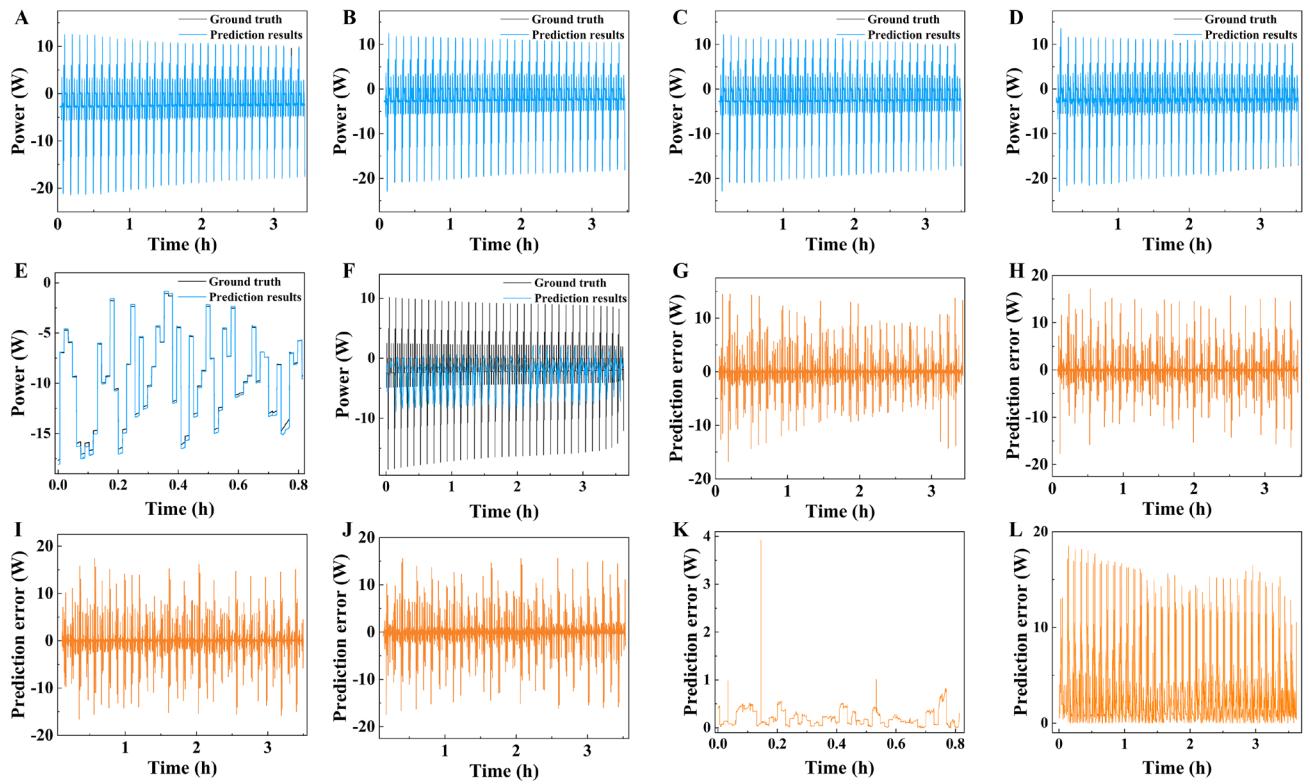


Figure S12. Examples of discharge power prediction results for the entire voltage range of the No. 37 battery when the cycle number is 1 with the output of 2 min (A), 4 min (B), 6 min (C), and 8 min (D) voltage and current sequences. Examples of discharge power prediction results when the cycle number is 1 with the output of 8 min voltage and current sequences for the entire voltage range of the No. 57 battery (E) and No. 53 battery (F). Power prediction error for No. 37 battery when the cycle number is 1 with the output of 2 min (G), 4 min (H), 6 min (I), and 8 min (J) voltage and current sequences. Power prediction error when the cycle number is 1 with the output of 8 min voltage and current sequences for No. 57 battery (K) and No. 53 battery (L).

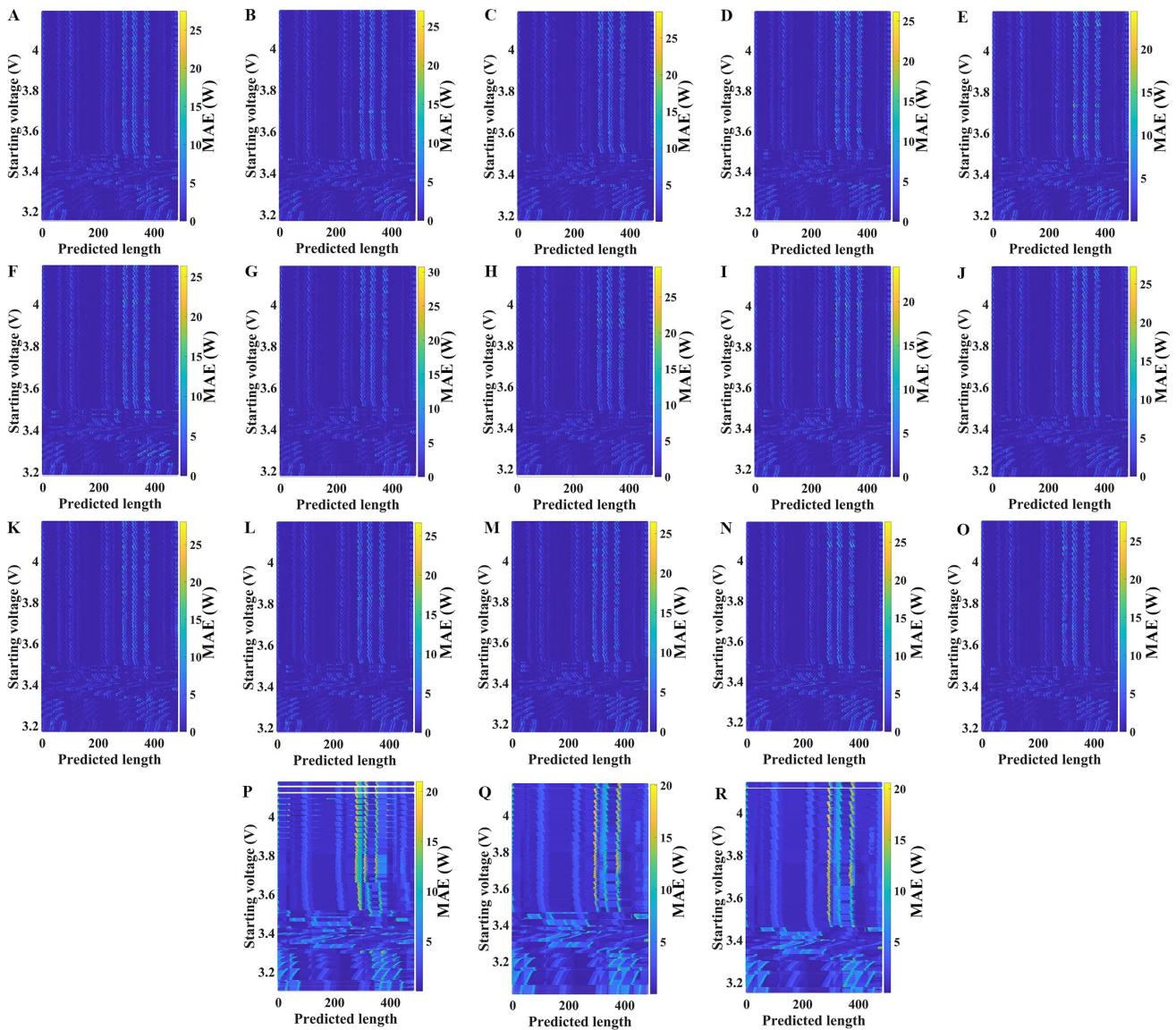


Figure S13. Power prediction MAE of normal test with the output of 8 min voltage and current sequences in the scenario of power prediction for No. 38 battery (A), No. 39 battery (B), No. 40 battery (C), No. 41 battery (D), No. 42 battery (E), No. 43 battery (F), No. 44 battery (G), No. 45 battery (H), No. 46 battery (I), No. 47 battery (J), No. 48 battery (K), No. 49 battery (L), No. 50 battery (M), No. 51 battery (N), No. 52 battery (O), No. 54 battery (P), No. 55 battery (Q), and No. 56 battery (R).

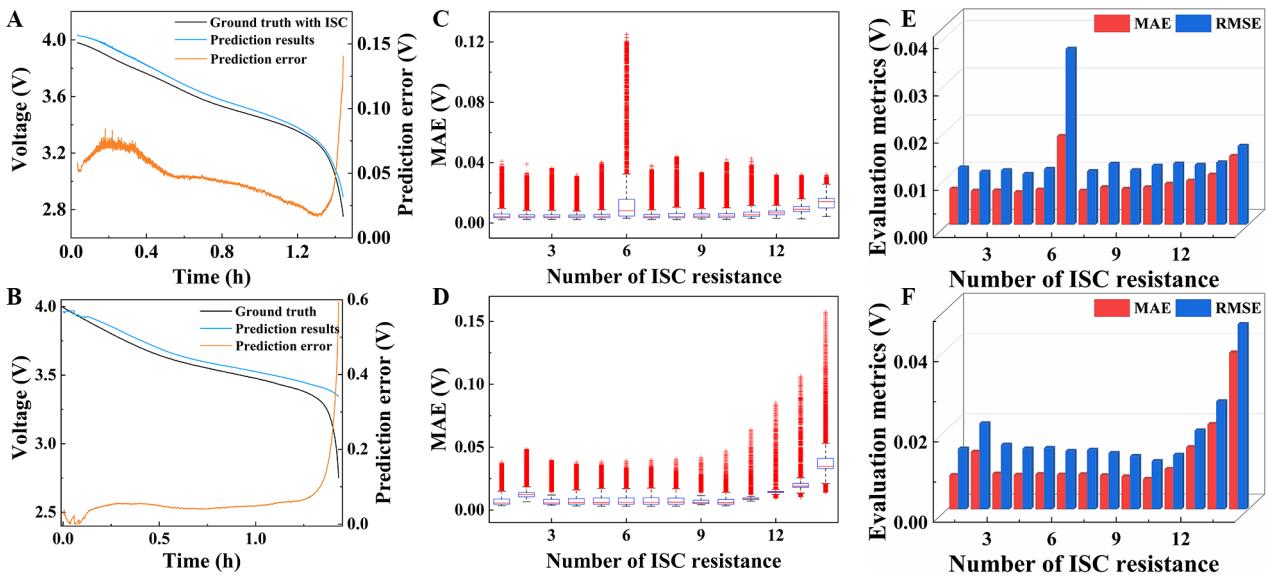


Figure S14. Examples of voltage prediction results for the No. 1 (A) and No. 33 battery (B) with the equivalent resistance for ISC is  $10\ \Omega$ . Error distribution of the prediction results with the equivalent resistance for ISC from  $1000\ \Omega$  to  $10\ \Omega$  for the No. 1 (C) and No. 33 battery (D). Voltage prediction MAE and RMSE of ISC test for the No. 1 (E) and No. 33 battery (F). In (C)-(F), the x axis is the number of ISC equivalent resistance, where 1 to 14 correspond to the equivalent resistances for ISC of  $1000\ \Omega$  to  $10\ \Omega$ , respectively.

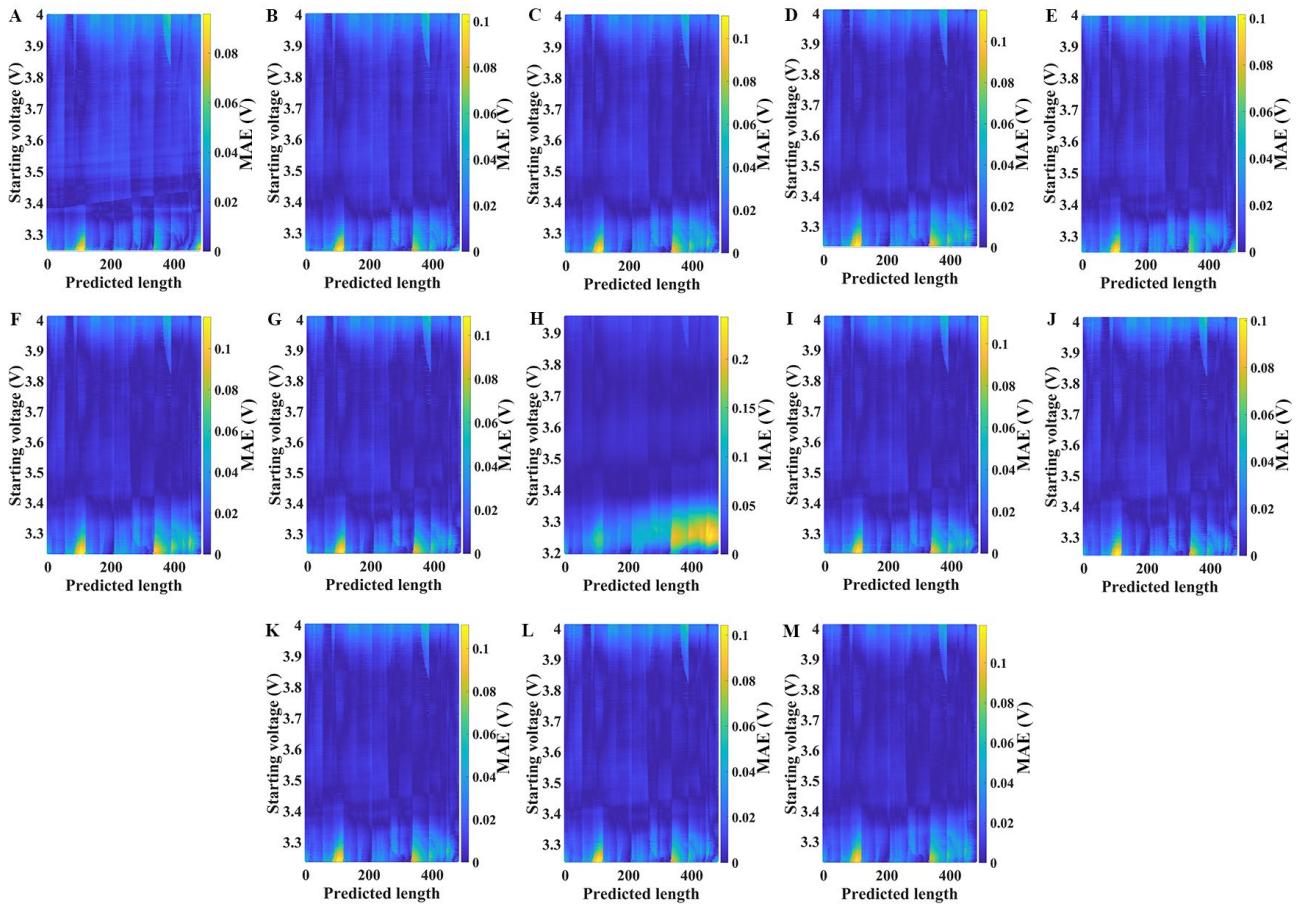


Figure S15. Voltage prediction MAE of ISC test for No. 1 battery with the equivalent resistance for ISC is  $20\ \Omega$

(A),  $30\ \Omega$  (B),  $50\ \Omega$  (C),  $100\ \Omega$  (D),  $200\ \Omega$  (E),  $300\ \Omega$  (F),  $400\ \Omega$  (G),  $500\ \Omega$  (H),  $600\ \Omega$  (I),  $700\ \Omega$  (J),  $800\ \Omega$  (K),  
 $900\ \Omega$  (L), and  $1000\ \Omega$  (M).

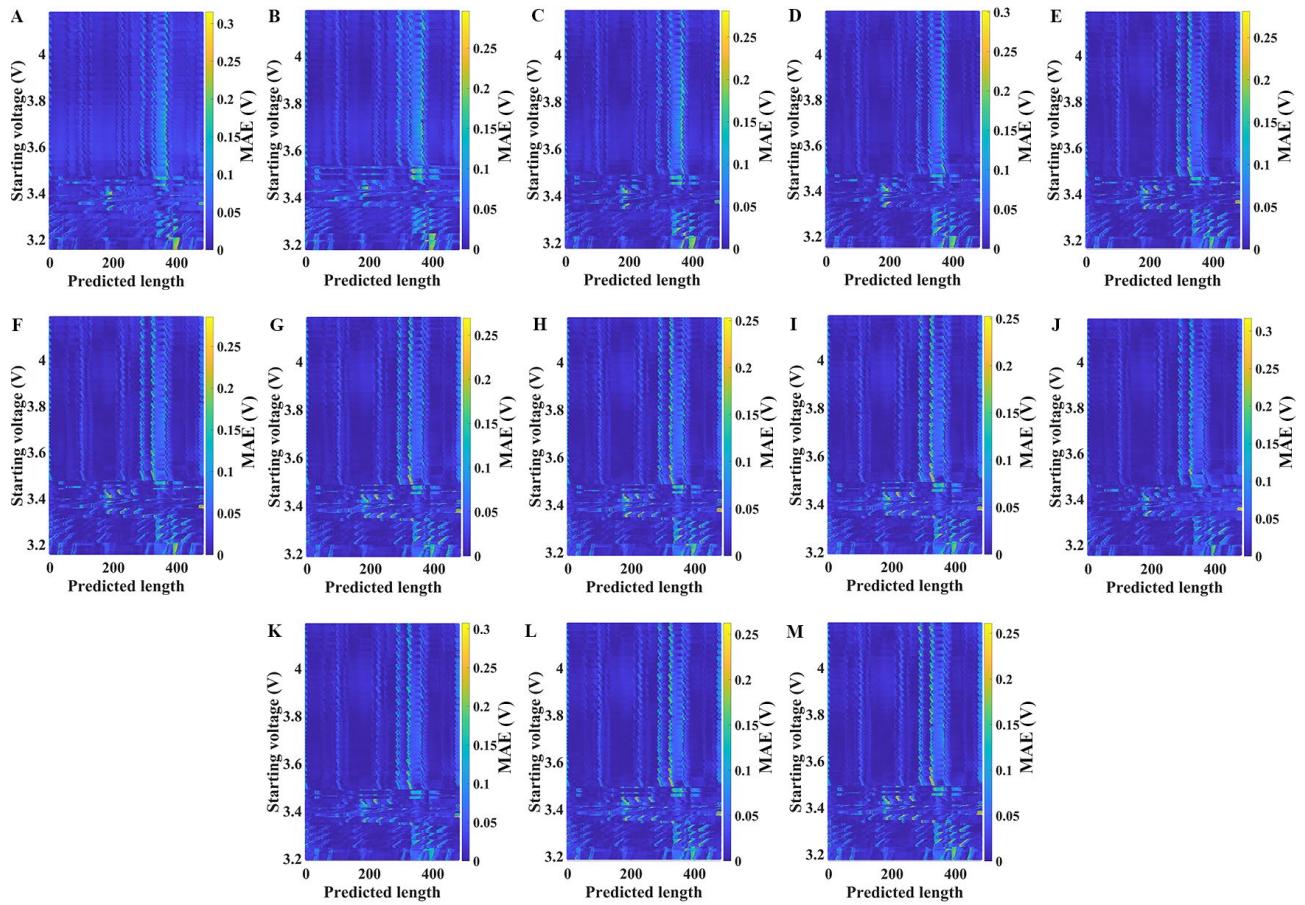


Figure S16. Voltage prediction MAE of ISC test for No. 37 battery with the equivalent resistance for ISC is  $20\ \Omega$

(A),  $30\ \Omega$  (B),  $50\ \Omega$  (C),  $100\ \Omega$  (D),  $200\ \Omega$  (E),  $300\ \Omega$  (F),  $400\ \Omega$  (G),  $500\ \Omega$  (H),  $600\ \Omega$  (I),  $700\ \Omega$  (J),  $800\ \Omega$  (K),  
 900  $\Omega$  (L), and 1000  $\Omega$  (M).

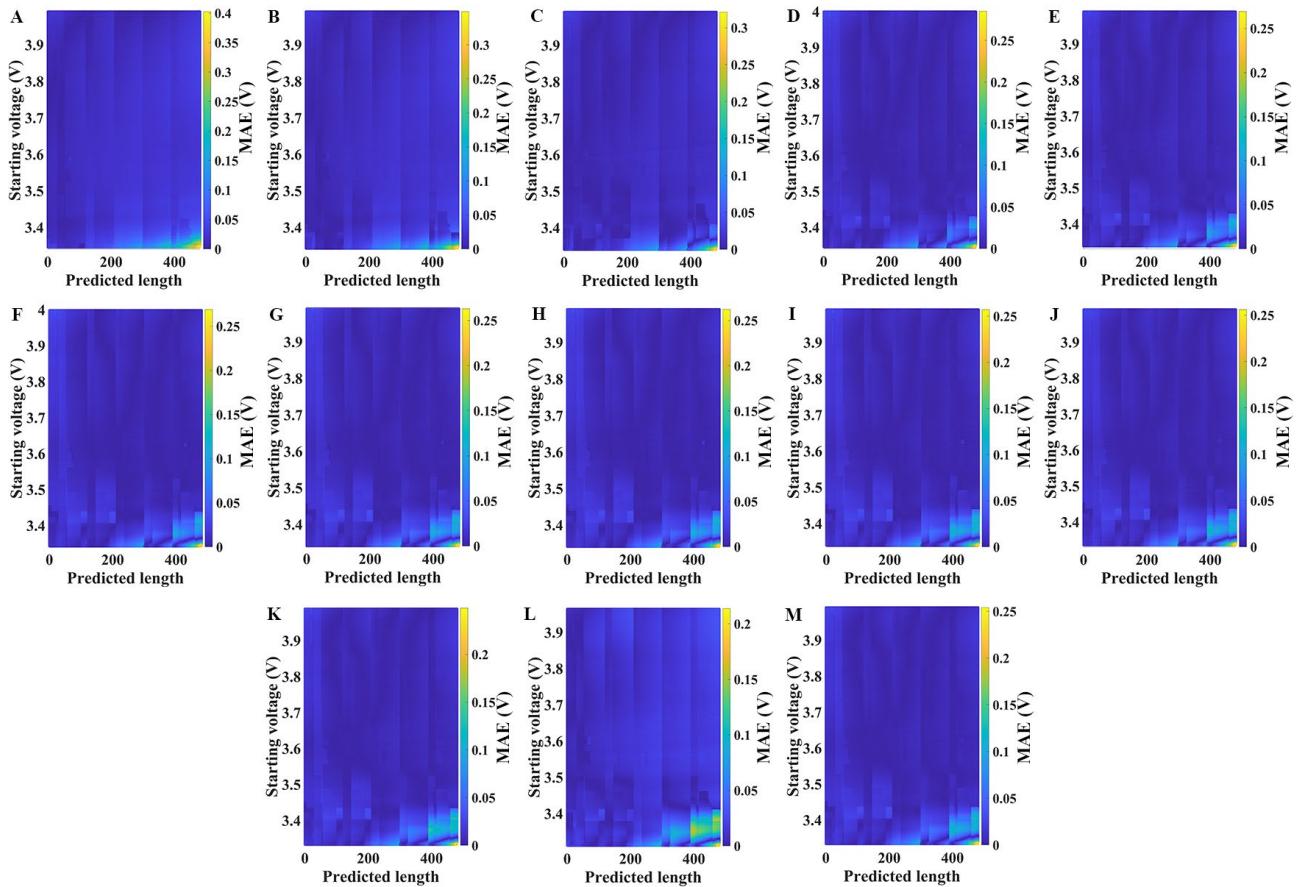


Figure S17. Voltage prediction MAE of ISC test for No. 33 battery with the equivalent resistance for ISC is  $20\ \Omega$

(A),  $30\ \Omega$  (B),  $50\ \Omega$  (C),  $100\ \Omega$  (D),  $200\ \Omega$  (E),  $300\ \Omega$  (F),  $400\ \Omega$  (G),  $500\ \Omega$  (H),  $600\ \Omega$  (I),  $700\ \Omega$  (J),  $800\ \Omega$  (K),  
 $900\ \Omega$  (L), and  $1000\ \Omega$  (M).

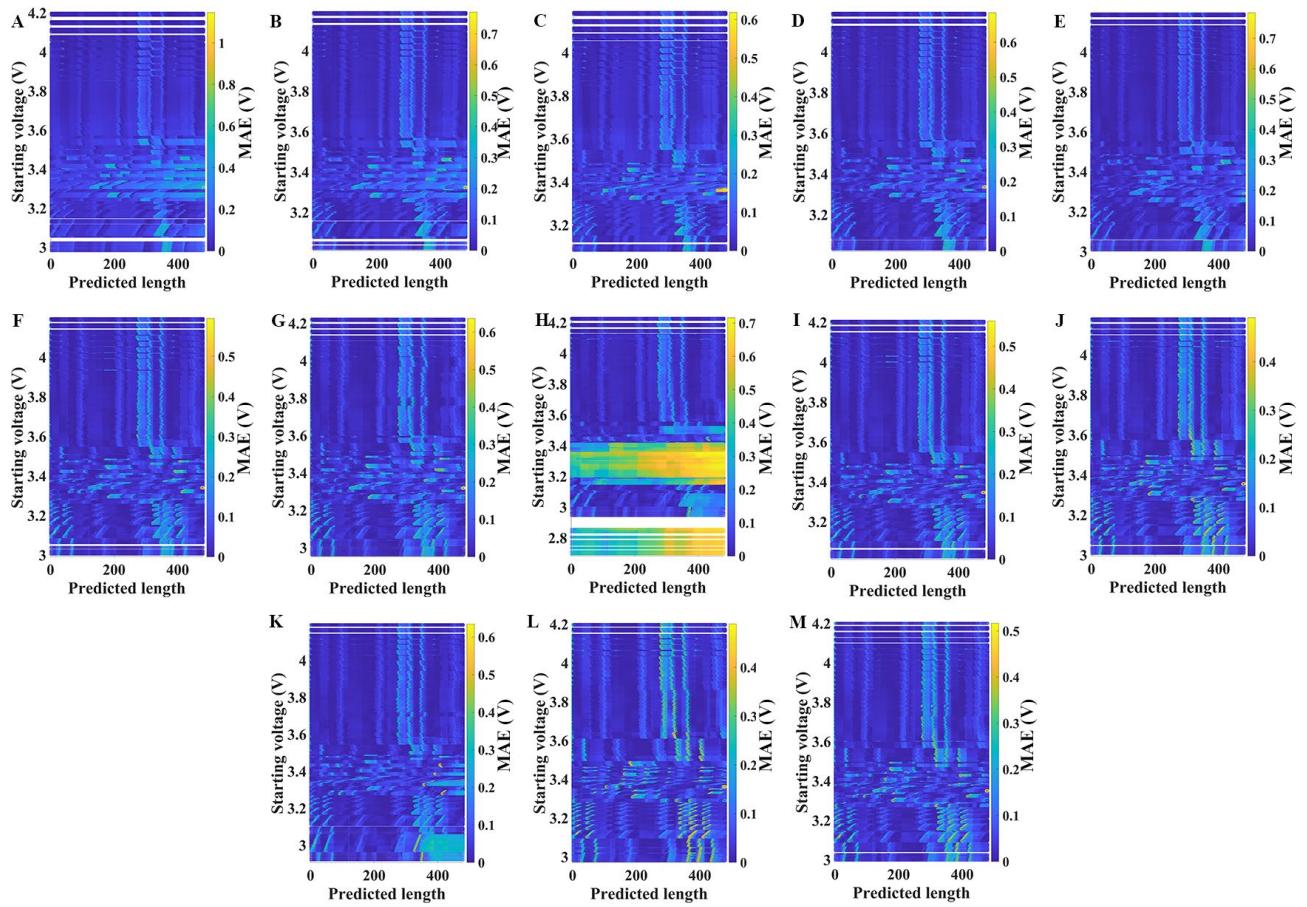


Figure S18. Voltage prediction MAE of ISC test for No. 53 battery with the equivalent resistance for ISC is  $20\ \Omega$

(A),  $30\ \Omega$  (B),  $50\ \Omega$  (C),  $100\ \Omega$  (D),  $200\ \Omega$  (E),  $300\ \Omega$  (F),  $400\ \Omega$  (G),  $500\ \Omega$  (H),  $600\ \Omega$  (I),  $700\ \Omega$  (J),  $800\ \Omega$  (K),  
 900  $\Omega$  (L), and 1000  $\Omega$  (M).

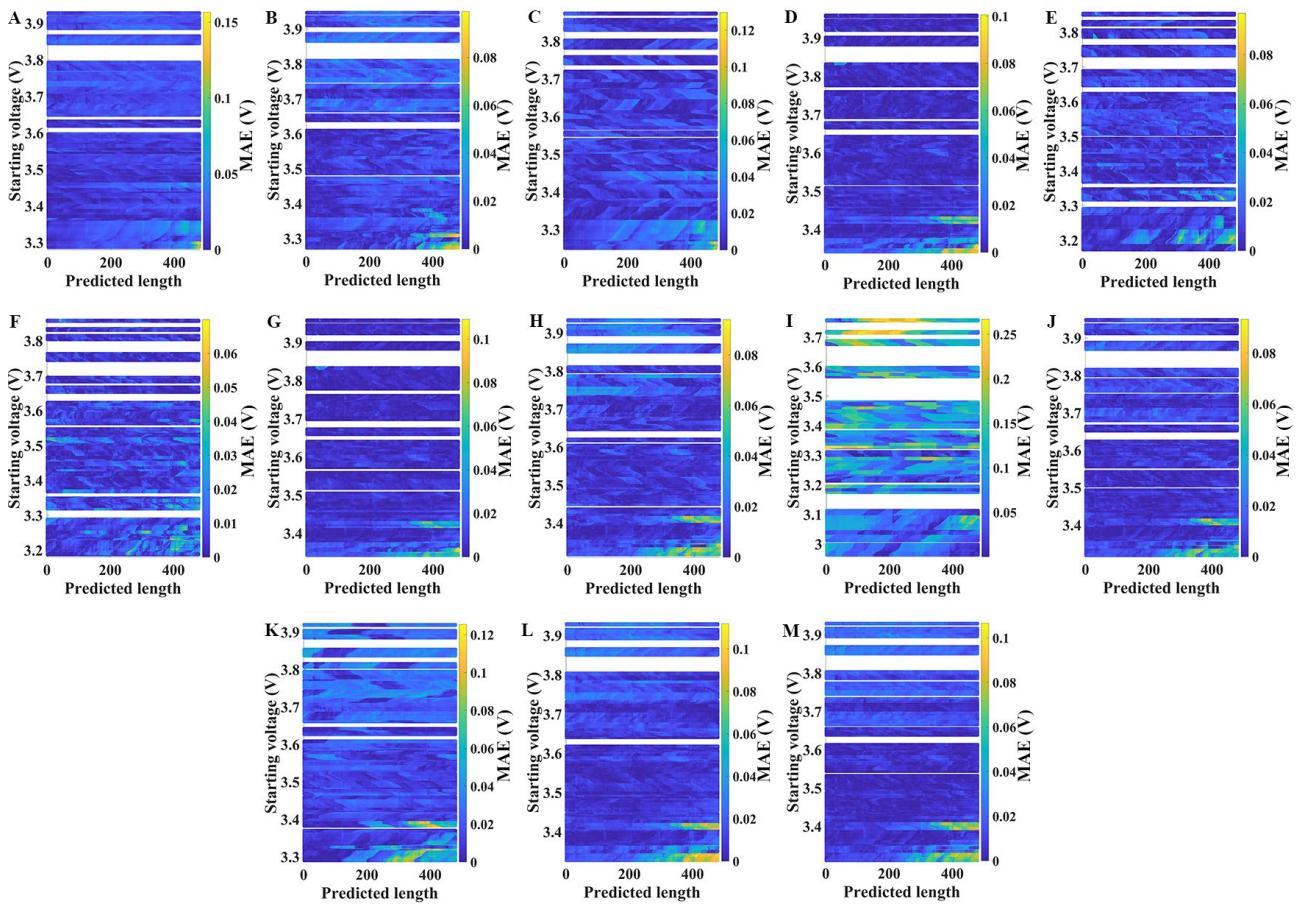


Figure S19. Voltage prediction MAE of ISC test for No. 57 battery with the equivalent resistance for ISC is  $20\ \Omega$

(A),  $30\ \Omega$  (B),  $50\ \Omega$  (C),  $100\ \Omega$  (D),  $200\ \Omega$  (E),  $300\ \Omega$  (F),  $400\ \Omega$  (G),  $500\ \Omega$  (H),  $600\ \Omega$  (I),  $700\ \Omega$  (J),  $800\ \Omega$  (K),  
 $900\ \Omega$  (L), and  $1000\ \Omega$  (M).

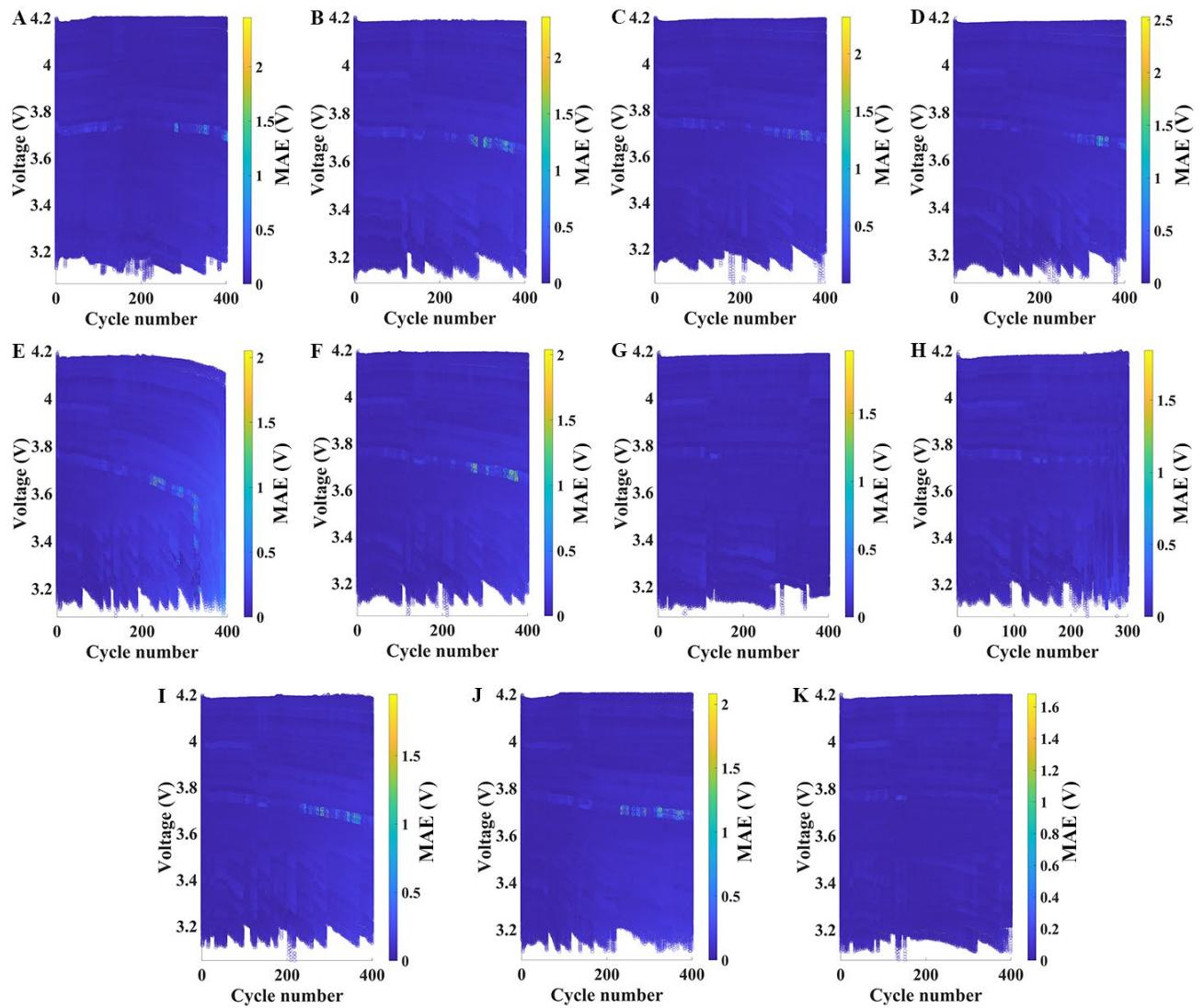


Figure S20. Voltage prediction MAE of ISC test for No. 60 battery (A), No. 61 battery (B), No. 62 battery (C), No. 63 battery (D), No. 64 battery (E), No. 65 battery (F), No. 66 battery (G), No. 67 battery (H), No. 68 battery (I), No. 69 battery (J), and No. 70 battery (K).

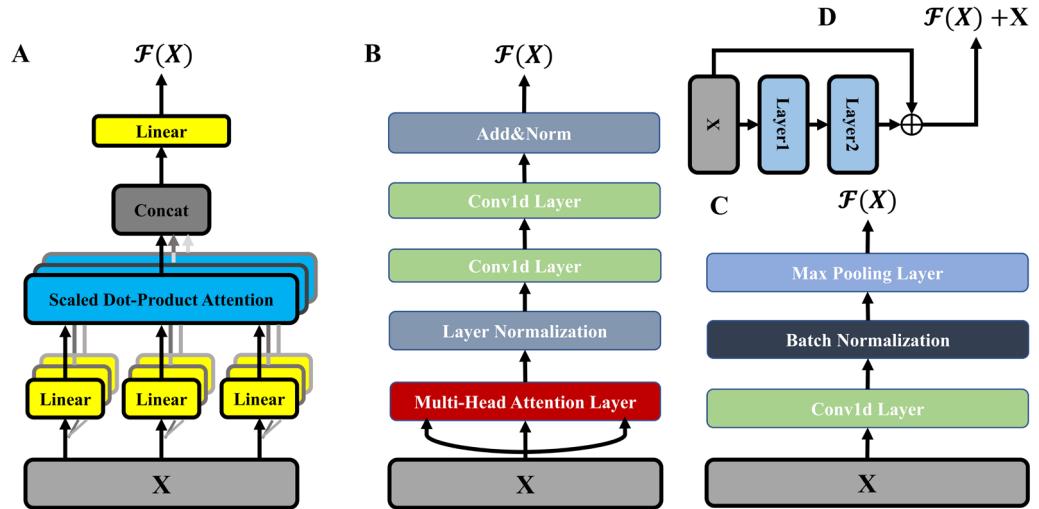


Figure S21. The architecture of the multi-head attention layer (A), encoder layer (B), conv layer (C), and residual connections (D).

## Supplementary Tables

Table S1. Description of all batteries included in the dataset. ‘Charging-discharging policy’ refers to the currents under charging and discharging and are formatted as C1-C2, where C1 and C2 represent the C rates applied in the charging process and discharging processes. TRS and VLS represent the training set and test set.

Battery Number	Dataset of ISC detection	Battery	Charging-discharging policy	Battery Number	Dataset of ISC detection	Battery	Charging-discharging policy
1	VLS	NCM811	0.5C-0.5C	36	TRS	NCM523	2.0C-2.0C
2	TRS	NCM811	0.5C-1.0C	37	TRS	NCM811	0.5C-DST
3	TRS	NCM811	0.5C-1.5C	38	TRS	NCM811	0.6C-DST
4	TRS	NCM811	0.5C-2.0C	39	TRS	NCM811	0.7C-DST
5	TRS	NCM811	0.6C-0.6C	40	TRS	NCM811	0.8C-DST
6	VLS	NCM811	0.7C-0.7C	41	TRS	NCM811	0.9C-DST
7	TRS	NCM811	0.8C-0.5C	42	TRS	NCM811	1.0C-DST
8	TRS	NCM811	0.8C-0.8C	43	VLS	NCM811	1.1C-DST
9	VLS	NCM811	0.8C-1.0C	44	VLS	NCM811	1.2C-DST
10	VLS	NCM811	0.8C-1.5C	45	TRS	NCM811	1.3C-DST
11	TRS	NCM811	0.8C-2.0C	46	VLS	NCM811	1.4C-DST
12	VLS	NCM811	0.9C-0.9C	47	TRS	NCM811	1.5C-DST
13	VLS	NCM811	1.0C-0.5C	48	TRS	NCM811	1.6C-DST
14	TRS	NCM811	1.0C-1.0C	49	TRS	NCM811	1.7C-DST
15	TRS	NCM811	1.0C-1.5C	50	VLS	NCM811	1.8C-DST
16	VLS	NCM811	1.0C-2.0C	51	TRS	NCM811	1.9C-DST
17	VLS	NCM811	1.1C-1.1C	52	VLS	NCM811	2.0C-DST
18	TRS	NCM811	1.2C-1.2C	53	VLS	NCM523	0.5C-DST
19	TRS	NCM811	1.3C-1.3C	54	TRS	NCM523	1.0C-DST
20	TRS	NCM811	1.4C-1.4C	55	TRS	NCM523	1.5C-DST
21	TRS	NCM811	1.5C-0.5C	56	TRS	NCM523	2.0C-DST
22	TRS	NCM811	1.5C-1.0C	57	TRS and VLS	NCM811	1.0C-RT
23	TRS	NCM811	1.5C-1.5C	58	TRS and VLS	NCM811	1.0C-DST
24	TRS	NCM811	1.5C-2.0C	59	-	NCM811	1.0C-DST
25	TRS	NCM811	1.6C-1.6C	60	-	NCM811	1.0C-DST
26	TRS	NCM811	1.7C-1.7C	61	-	NCM811	1.0C-DST
27	TRS	NCM811	1.8C-1.8C	62	-	NCM811	1.0C-DST
28	TRS	NCM811	1.9C-1.9C	63	-	NCM811	1.0C-DST
29	TRS	NCM811	2.0C-0.5C	64	-	NCM811	1.0C-DST
30	TRS	NCM811	2.0C-1.0C	65	-	NCM811	1.0C-DST

31	VLS	NCM811	2.0C-1.5C	66	-	NCM811	1.0C-DST
32	VLS	NCM811	2.0C-2.0C	67	-	NCM811	1.0C-DST
33	VLS	NCM523	0.5C-0.5C	68	-	NCM811	1.0C-DST
34	TRS	NCM523	1.0C-1.0C	69	-	NCM811	1.0C-DST
35	TRS	NCM523	1.5C-1.5C	70	-	NCM811	1.0C-DST

Table S2. Specifications of NCM811 and NCM523 batteries.

Battery type	Anode material	Cathode material	Nominal voltage	Cutoff voltage	Nominal capacity
18650	Graphite	$\text{LiNi}_{0.8}\text{Co}_{0.1}\text{Mn}_{0.1}\text{O}_2$	3.7V	2.75V – 4.2V	2.7Ah
18650	Graphite	$\text{LiNi}_{0.5}\text{Co}_{0.2}\text{Mn}_{0.3}\text{O}_2$	3.7V	2.75V – 4.2V	2.4Ah

Table S3. Detailed information on the random test (RT) working condition. The RT working condition includes 70 steps, and each step lasts for 50 seconds.

Step	Rate/C								
1	0.12	15	1.74	29	0.56	43	1.13	57	1.11
2	0.32	16	1.63	30	1.30	44	0.95	58	0.77
3	1.62	17	1.72	31	1.08	45	0.25	59	0.84
4	0.62	18	1.51	32	0.88	46	0.82	60	1.53
5	1.73	19	0.60	33	0.45	47	1.65	61	1.55
6	0.20	20	1.00	34	0.62	48	1.40	62	0.95
7	1.11	21	0.94	35	0.11	49	0.50	63	1.82
8	0.87	22	0.17	36	0.13	50	0.80	64	1.81
9	1.76	23	1.08	37	1.26	51	0.27	65	0.86
10	0.66	24	1.74	38	0.47	52	1.25	66	1.01
11	0.44	25	1.55	39	1.78	53	1.24	67	0.72
12	0.57	26	0.83	40	1.71	54	1.11	68	1.24
13	0.90	27	0.24	41	0.57	55	1.04	69	1.39
14	1.60	28	1.37	42	1.44	56	0.49	70	1.40

Table S4. The equivalent ISC resistances paralleled to No. 59 to No. 70 batteries.

Battery Number	59	60	61	62	63	64	65	66	67	68	69	70
ISC resistances/ $\Omega$	10	20	30	50	100	200	300	400	500	600	700	800

Table S5. Cycle policy description of all batteries included in the dataset. ‘Normal test’ refers to the battery charged and discharged with no external resistances, and ‘ISC test’ refers to simulating the electrical performance of a battery with ISC through parallel resistors during battery charge and discharge.

Cycle number	Test type	Equivalent ISC resistance	Cycle number	Test type	Equivalent ISC resistance
1	Normal test	-	15	Normal test	-
2	ISC test	1000Ω	16	ISC test	300Ω
3	Normal test	-	17	Normal test	-
4	ISC test	900Ω	18	ISC test	200Ω
5	Normal test	-	19	Normal test	-
6	ISC test	800Ω	20	ISC test	100Ω
7	Normal test	-	21	Normal test	-
8	ISC test	700Ω	22	ISC test	50Ω
9	Normal test	-	23	Normal test	-
10	ISC test	600Ω	24	ISC test	30Ω
11	Normal test	-	25	Normal test	-
12	ISC test	500Ω	26	ISC test	20Ω
13	Normal test	-	27	Normal test	-
14	ISC test	400Ω	28	ISC test	10Ω

Table S6. Model metrics for the normal battery voltage prediction results with different input/output sequence lengths and working operations. CC and DST refer to the 32 and 16 different working operations listed in Table S1, respectively.

Metrics (mV)	Input sequence length/output sequence length (s)								
	480/120		480/240		480/360		480/480		900/900
	CC	DST	CC	DST	CC	DST	CC	DST	DST
MAE	9.5	10.5	23.8	19.8	10.7	15.2	11.7	19.3	17.9
RMSE	12.7	13.3	28.8	31.0	16.1	25.4	17.7	34.4	26.9

Table S7. Model metrics for the normal battery voltage prediction results with different methods when the input and output sequences length are equal to 480.

Metrics (mV)	DNN	CNN	GRU	LSTM	Encoder-Decoder

MAE	37.5	37.8	34.2	20.9	16.8
RMSE	40.1	40.1	37.8	22.3	19.1

Table S9. Model metrics for the normal battery power prediction results under all DST working conditions with different input/output sequence lengths.

Metrics (W)	Input sequence length/output sequence length (s)				
	480/120	480/240	480/360	480/480	900/900
MAE	0.4	0.4	0.6	0.7	0.7
RMSE	1.1	1.2	1.6	1.7	1.9

Table S10. Model metrics for the normal battery power prediction results with different methods when the input and output sequences length are equal to 480.

Metrics (W)	DNN	CNN	GRU	LSTM	Encoder-Decoder
MAE	2.7	2.7	2.7	0.8	0.7
RMSE	6.3	6.5	6.2	2.2	1.9

Table S11. Model metrics for the ISC detection with different sequence lengths of No. 1-32 and No. 37-52 battery. “120, 240, 360, 480” refer to the output sequence lengths of voltage with the input of 8 min voltage sequences.

ISC resistance	Predicted sequence length (s)			
	120	240	360	480
10	94.26	86.01	87.97	95.92
20	91.16	78.45	76.78	90.57
30	85.81	70.25	76.10	88.97
50	77.94	67.77	68.94	79.50
100	63.43	67.65	70.73	78.25
200	64.26	70.70	77.10	88.29
300	64.55	71.47	72.69	84.50
400	70.60	64.11	74.69	85.96
500	59.89	63.10	75.61	85.42
600	61.65	69.39	74.63	84.05
700	59.91	66.75	70.67	84.08
800	61.77	72.00	73.43	86.71
900	60.63	67.78	73.25	83.36

1000	69.30	73.98	77.87	89.50
------	-------	-------	-------	-------

Table S12. Model metrics for the ISC detection of the dataset of NCM523 battery, the dataset over the full life cycle of NCM811 battery, and the dataset of NCM811 battery under RT working conditions with the input of 8 min sequence. NCM811-RT refers to the dataset of NCM811 batteries under RT working conditions. NCM811-FL refers to the dataset over the full life cycle of the NCM811 battery.

Dataset	ISC resistance													
	10	20	30	50	100	200	300	400	500	600	700	800	900	1000
<b>NCM523</b>	99.9	95.1	92.2	80.8	85.6	75.1	88.5	86.3	65.9	71.1	68.7	65.5	75.7	80.9
<b>NCM811-RT</b>	98.8	97.4	82.4	71.5	72.5	71.5	73.9	71.0	69.3	98.9	80.3	73.3	80	79.5
<b>NCM811-FL</b>	99.8	97.5	95.7	92.8	78.8	87.7	79.4	71.2	75.9	84.7	90.2	68.5	-	-

Table S13. Model metrics for the ISC detection with different methods when the input is 8 min sequence.

Methods	ISC resistance													
	10	20	30	50	100	200	300	400	500	600	700	800	900	1000
<b>CNN</b>	90.3	81.8	75.5	68.9	54.9	52	54.7	53.1	53.9	53.3	55.1	55.5	53.1	58.2
<b>DNN</b>	90.2	81.4	75.3	69	54.7	52.2	54.4	53.4	54	53.1	55.1	55.5	51.2	58.6
<b>GRU</b>	90.3	81.2	75.2	68.6	54.6	50.3	54.2	52	53.6	51.3	54.7	54.8	50.7	57.9
<b>LSTM</b>	82.1	87	64.9	70.1	52.5	50.6	50.2	50.9	51.5	50.2	58	50.7	50.6	53.2
<b>Encoder-Decoder</b>	95.9	90.6	89.0	79.5	78.3	88.3	84.5	86.0	85.4	84.1	84.1	86.7	83.4	89.5

Table S14. The 1D convolutional layers are in detail.

Encoder			Decoder		
network components	kernel size	number of filters	network components	kernel size	number of filters
inputs embedding layer	3	512	Inputs embedding layer	3	512
The first 1D convolutional layer of the encoder layer	1	2048	The first Conv1d layer	1	2048
The second 1D convolutional layer of the encoder layer	1	512	The second Conv1d layer	1	512
conv layer	3	512			

Table S15. The neural numbers and activation function of each layer.

Model layer		The neural numbers of each layer	The activation function used in each layer
Input embedding		3072	None
Encoder	Multi-head attention layer	1048576	None
	Conv layer	786432	Elu (see Note S3)
	Layer normalization	512	None
Decoder	Multi-head attention layer	1048576	None
	Layer normalization	512	None
	1D conv layer	786432	Elu (see Note S3)
Final dense layer		512	None