States estimation of lithium ion battery using data-driven methods

Master of Science in Aerospace Science and Technologies
Mid-term Report Defense
School of Civil Aviation
Northwestern Polytechnical University
Xi'an, Shaanxi, China

Amrit Dhakal

2025.12.11

Summary of Thesis Research Progress

Chapter	Title	Progress
1	Introduction, background and literature Review	Completed
2	Brief Overview of lithium ion battery's states estimation methods	Completed
3	CNN-BiGRU with temporal attention for SOH and RUL estimation on CS2-Cells	Completed
4	IPEformer: a time series transformer for SOH estimation on public and proprietary dataset	Not yet completed
5	Conclusion and future outlook	After the completion of all

CONTENTS

1 Introduction and significance

Introduction: Lithium ion battery

WHAT?

- LIBs: Energy Storage Devices
- $LiCoO_2$ (Cathode)
- Graphite(Anode)

WHERE?

- Portable Electronic Devices,
- Electric Vehicles,
- Grid Energy Storage System,
- More Electric Aircraft (MEA)
- UAV and Space Vehicles

WHY?

- High Energy Density
- High Charge/Discharge Cycles
- Highly Efficient

HOW?

- Battery Modelling, Data Training/Testing, Experimentations
- Prognostics and health Management, and remaining useful life (RUL) prediction
- Safety and Reliability
- Battery Management System(BMS)

The Critical Role of Battery Management Systems (BMS)

- **Energy Transition & Electricity's Role:**
  - Global shift to renewables; electricity is the key energy carrier.
- **AC/DC Conversion & Modern Tech:**
  - EVs/renewables require constant AC-DC conversion (grid → battery → motor).
  - System efficiency hinges on battery health and safety.
- **Li-ion Batteries: High Performance, High Risk:**
  - Preferred for high energy density (EVs, electronics, storage).
  - Susceptible to stress, degradation, and thermal runaway if misused.
- **Core BMS Function:**
  - Essential for monitoring and safeguarding Li-ion batteries.
  - Manages strict operating limits (voltage, temperature).
  - Balances cells to maintain capacity and safety.
  - Evaluates complex, non-linear degradation for reliability.

Battery Chemistry	Nominal voltage	Life cycles	Energy density	Cost	Safety
Lead-Acid	2.0 V	500 - 1,000	30 - 50	Low	Highly Safe
$LiFePO_4$ (LFP)	3.2 V	3,000 - 6,000	90 - 160	Medium	Safest
$LiCoO_2$ (LCO)	3.6 V	500 - 1,000	150 - 240	High	Low
$Li(NiMnCo)O_2$ (NMC)	3.6 V	1,000 - 2,000	150 - 250	Medium	Medium
$Li(NiCoAl)O_2$ (NCA)	3.6 V	500 - 1,500	200 - 280	High	Low
$LiMnO_4$ (LMO)	3.7 V	500 - 1,500	100 - 150	Medium	Good

Problem Statement: Accurate Battery Health Monitoring

- **SOH & RUL Are Key:** SOH (capacity) and RUL are vital for PHM.
- **SOH is Foundational:** SOC, SOP, SOE, and RUL depend on accurate SOH.
- **Safety & Environmental Threshold:** Critical EOL mark is 70% SOH.
- **Inherent Risk of Degradation:** Aging leads to internal damage (SEI/dendrites), risking internal short circuits and thermal runaway.
- **Real-World Safety Hazard:** EV and aviation fires highlight the consequences of Li-ion battery failure.
- **Core Challenge:** Internal degradation processes cannot be directly measured and are highly variable.
- **Research Imperative:** Advanced ML/AI algorithms in the BMS are needed for **proactive safety**, predicting failure long before a critical hazard occurs.

Figure 1 : Air Busan in South Korea caught fire due to LiBs failure

Figure 2 : Five cars were destroyed in the fire started from one car's battery failure

2

Battery States estimation methods

Battery States estimation Methods Overview

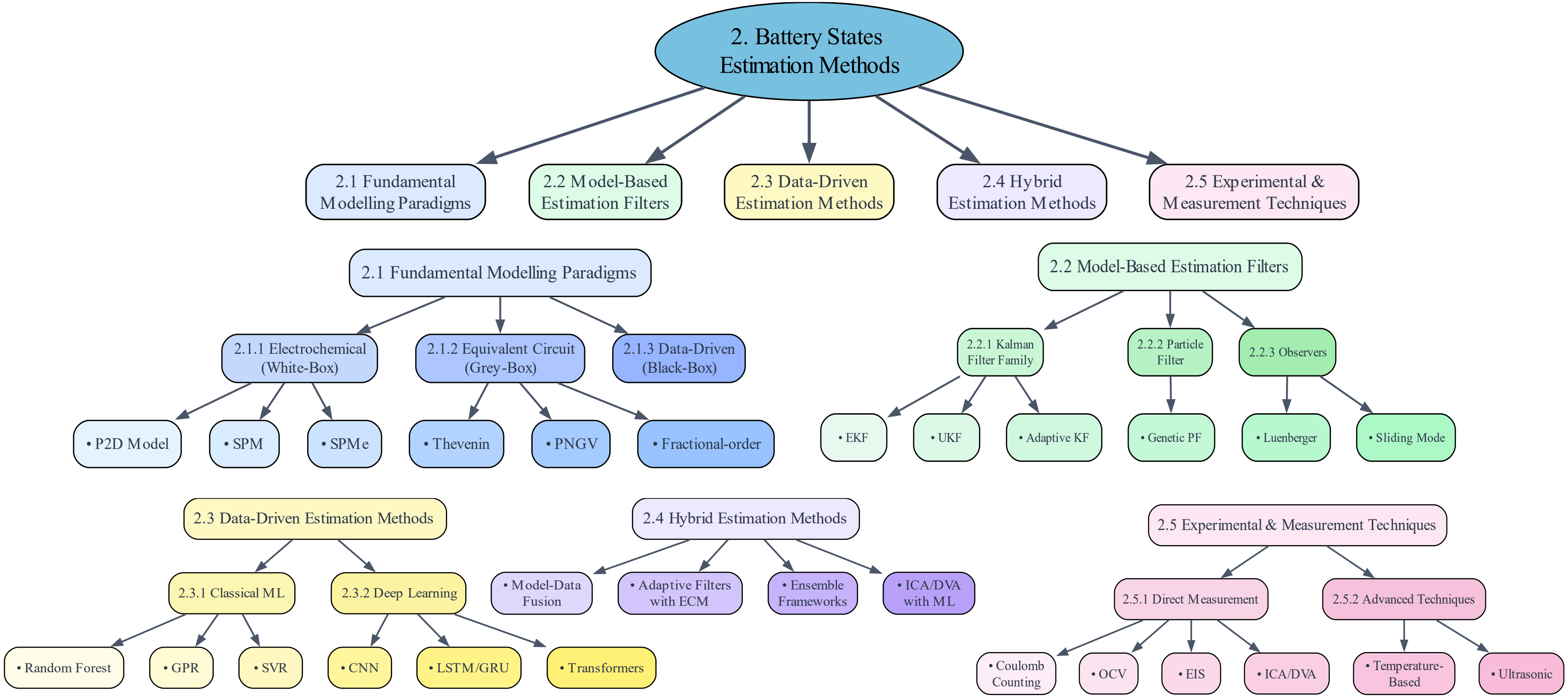


Figure. Battery States estimation methods conventional and machine learning and Deep learning methods

2.1 Physics-Based Models (PBMs)

- Also known as Electrochemical (White-box) Modelling
- **Core Strength:** Captures internal physical/chemical phenomena (potentials, concentrations, fluxes).
- **Model Hierarchy:** Ranges from molecular-scale PDEs (impractical) to the Doyle-Fuller-Newman (DFN) model.
- **Computational Challenge:** The high-fidelity DFN model is too complex for direct BMS implementation.
- **Solution Path:** Algorithms (e.g., Discrete-Time Realization) can simplify DFN to a low-order state-space model for BMS use.
- **Governing Principles (DFN Core):** A system of coupled Partial Differential Equations (PDEs) modeling:
- Charge conservation in Solid (Ohm's law with reaction current)
- Mass conservation in Solid (Fick's law for lithium diffusion): Models Li^{+} diffusion *inside* electrode particle

$$\frac{\partial c_s^r}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(D_s^r r^2 \frac{\partial c_s^r}{\partial r}\right),$$

- Charge conservation in Electrolyte (ion migration & diffusion): Governs ion transport *between* electrodes.

$$\frac{\partial}{\partial x}\left[\kappa_{\text{eff}}^r\left(\frac{\partial \Phi_e^r}{\partial x} + \frac{2RT(t_+^0 - 1)}{F}\left(1 + \frac{\partial \ln f_\pm}{\partial \ln c_e^r}\right)\frac{\partial \ln c_e}{\partial x}\right)\right] + a_s^r F j^r = 0,$$

- Mass conservation in electrolyte
- Kinetics (Butler-Volmer equation for reaction rates)

$$j^r = j_0^r\left[exp\left(\frac{(1-\alpha^r)F\eta^r}{RT}\right) - exp\left(\frac{-\alpha^r F\eta^r}{RT}\right)\right]$$

2.2 Equivalent Circuit (Grey-Box) Modelling

- Uses electrical components (R, C, voltage sources) to model battery dynamics.
- Combines a **physically-measured OCV** with **empirically-fitted R/C parameters**.
- Computationally efficient, ideal for BMS state estimation and real-time use.
- General n-RC ECM Structure:
  - Terminal Voltage:

$$V(t) = U_{\mathrm{OCV}}\big(z(t)\big) + I(t)R_0(z,T) + \sum_{k=1}^{n} V_k(t)$$

  - RC Branch Dynamics:

$$\dot{V}_k(t) = -\frac{V_k(t)}{R_k(z,T)C_k(z,T)} + I(t), \quad k = 1, \dots, n$$

  - SOC Evolution:

$$\dot{z}(t) = -\frac{\eta I(t)}{Q_{\mathrm{nom}}},$$

- **Common Model Variants:**
  - Rint (0-RC): $V(t) = U_{\mathrm{OCV}}(z) + IR_0$. (Simplest, low accuracy)
  - Thevenin (1-RC): Captures fast polarization (~10-30s).
  - 2-RC Model: Industry standard. **Fast RC** (charge transfer) & **Slow RC** (diffusion effects).
  - n-RC / Randles w. Warburg: For high accuracy; models specific diffusion impedance.

2.3 Model based estimation filters

- Estimate battery states (SOC, SOH) from noisy measurements using a model.
- **Kalman Filter (KF) & Variants:** Optimal linear estimator for Gaussian noise.
  - Predict-Correct cycle: Time update → Measurement update.
  - State vector: $x_k = [z_k \quad V_{1,k} \quad V_{2,k} \quad \cdots]^T$
- **Extended KF (EKF):** Industry standard. Linearizes nonlinear models (e.g., OCV-SOC) using Jacobians.
- **Unscented KF (UKF):** More accurate for strong nonlinearities. Uses deterministic sampling (sigma points), no Jacobians.
- **Dual EKF (DEKF):** Runs two filters in parallel for simultaneous **state (SOC)** and **parameter (SOH)** estimation.
- **Adaptive KF (AKF):** Adjusts noise covariances online for robustness against aging/conditions.
- **Particle Filter (PF):** Sequential Monte Carlo; uses weighted particles to represent state probability.
- Excels with strong nonlinearities & non-Gaussian noise (e.g., LFP flat OCV).
- Computationally heavy (500-5000 particles); often used offline/research.
- **Genetic PF (GPF):** Enhances diversity via crossover/mutation, improving long-term SOH tracking.
- **Luenberger Observer (LBO):** Deterministic observer using gain LL to drive error to zero.
- **Sliding Mode Observer (SMO):** Highly robust. Uses discontinuous control to force finite-time convergence.
  - Popular in automotive BMS for robustness against parameter variations and noise.
  - Structure: $\dot{\hat{z}} = -\frac{\eta I}{Q} + k_1 \mathrm{sign}\big(V - \hat{V}\big)$,

2.4 Data-Driven Estimation Methods

- Learns complex degradation patterns directly from operational data.
- Does not require detailed physics-based models of degradation.

Classical Machine Learning (ML) Methods

- Extract health features (voltage, current, impedance) → Train regression models to predict capacity/resistance (SOH).
- **Support Vector Regression (SVR):**
  - Finds a function that fits data within an error margin (ε).
  - Uses kernels (e.g., RBF) to handle non-linear relationships.
  - Well-suited for high-dimensional feature spaces (e.g., ICA/DVA features).
- **Random Forest Regression (RFR):**
  - Ensemble of decision trees trained on random data subsets.
  - **Prediction:** y^(x)=1T∑t=1Tht(x)*y*^(*x*)=*T*1∑*t*=1*Tht*(*x*)
  - Robust to overfitting; provides feature importance scores.
- **Gaussian Process Regression (GPR):**
  - Bayesian, non-parametric method. Provides full **predictive distribution** (mean + uncertainty).
  - Ideal for limited datasets and safety-critical applications requiring uncertainty bounds.
  - Computationally heavy (O(N3)*O*(*N*3) scaling).

3

Experiments on CALCE dataset

3.1 Dataset description

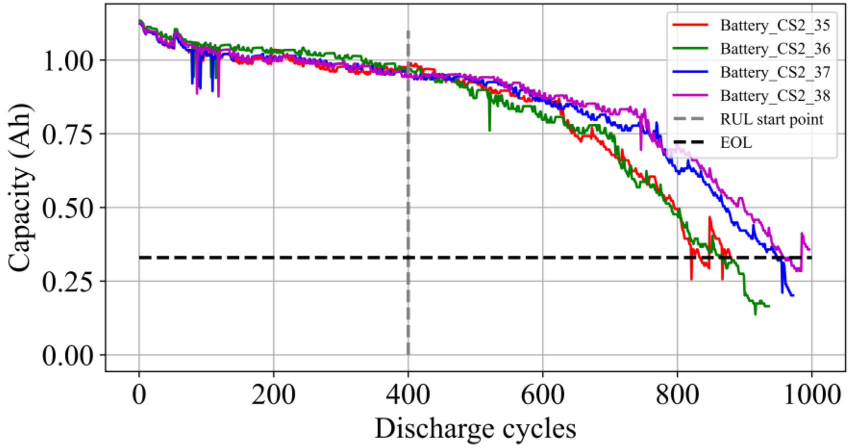


Figure. Battery capacity degradation on four cells of CALCE dataset

Parameters	Values
Dataset	CALCE CS2 (Pouch Cells)
Cells	CS2_35, CS2_36, CS2_37, CS2_38
Cell Chemistry	$LiCoO_2$ cathode,
Dimensions	5.4 x 33.6 x 50.4 mm
Nominal Voltage	4.2 V
Energy Capacity	1.1 Ah
Charging Current	0.55 A or 0.5 C-rate
Discharging Current	1.1 A or 1 C-rate
Voltage Range	4.2 V (charge) to 2.7 V(discharge)

Table. CALCE dataset CS2 cells battery parameters

$$\textbf{Remaining Useful Life (RUL)} =$$
$$RUL_{true} = N_{EOL}^{true} - N_{current},$$
$$RUL_{pred} = N_{EOL}^{pred} - N_{current},$$

where N_{EOL}^{true} and N_{EOL}^{pred} are the true and predicted end of-life cycles, respectively, and $N_{current}$ is the RUL monitoring cycle

$$\textbf{State of Health (SOH)} = \frac{C_{current}}{C_{rated}} \times 100\%,$$

where $C_{current}$ is the measured capacity of the battery and C_{rated} is the rated or initial capacity

3.2 Evaluation Metrics

- Root Mean Square Error(RMSE) $= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(SOH_{true,i} - SOH_{pred,i}\right)^2}$,
- Mean Absolute Error(MAE) $= \frac{1}{n}\sum_{i=1}^{n}\left|SOH_{true,i} - SOH_{pred,i}\right|$,
- Coefficient of determination $\mathrm{R}^2_{\mathrm{SOH}} = 1 - \frac{\sum_{i=1}^{n}\left(SOH_{true,i} - SOH_{pred,i}\right)^2}{\sum_{i=1}^{n}\left(SOH_{true,i} - SOH_{true}\right)^2}$,
- Remaining Useful Life (RUL) Error $= \left|RUL_{pred} - RUL_{true}\right|(in\ cycles)$

3.3 Research Methodology: Gated Recurrent Unit

- Captures long-term temporal dependencies in cycling data
- Efficient for time-series sensor inputs (voltage, current, temperature)
- Hidden state updates directly (no separate memory cell)
- Reset & update gates : prevent gradient dispersion
- Maintains long-term memory capacity with reduced computational load

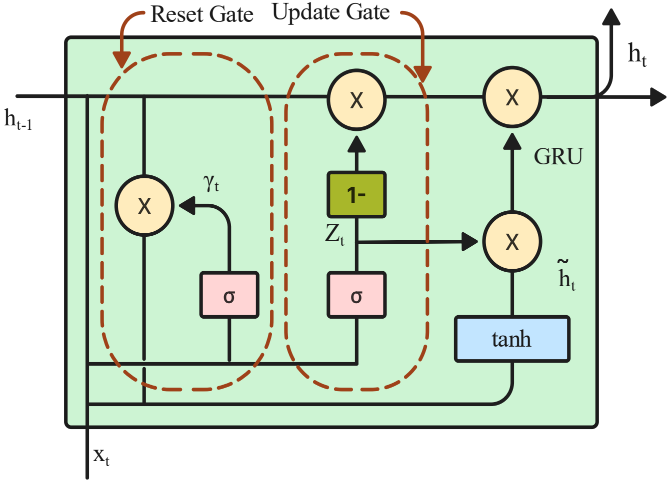


Figure. Gated Recurrent Unit (GRU)

$$Z_t = \sigma(W_z.[h_{t-1}, X_t]),$$

$$R_t = \sigma(W_r.[h_{t-1}, X_t]),$$

$$\hat{h}_t = \tanh(W.[R_t * h_{t-1}, X_t]),$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \hat{h}_t,$$

$$\text{where, } \sigma = \frac{1}{1+e^{-x}}, \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

3.4 Bidirectional GRU

- Processes data forward & backward
- Learns richer temporal correlations and degradation features
- Improves robustness and accuracy of SOH prediction
- Enhanced feature extraction from cycling data
- Stronger predictive performance for battery health

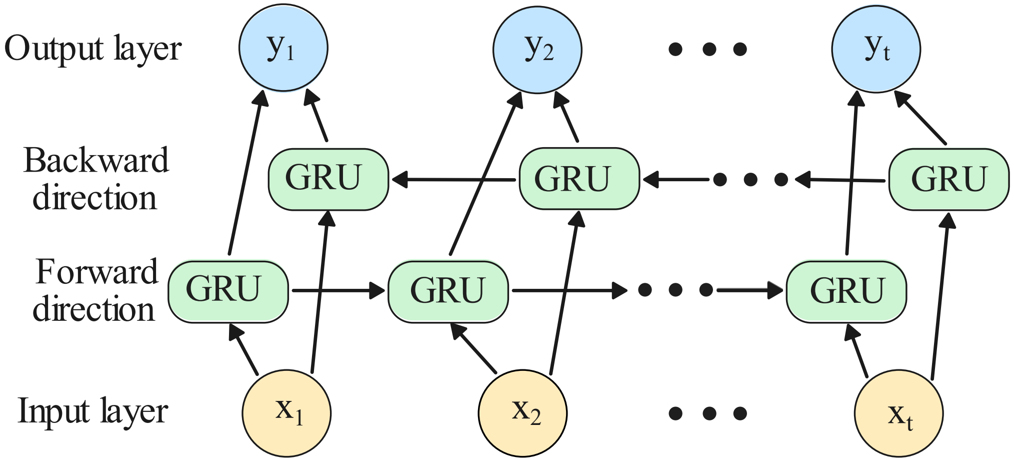


Figure. Bidirectional Gated Recurrent Unit (GRU)

3.5 Temporal Attention Mechanism

For input sequence $\boldsymbol{H} = [\boldsymbol{h_1}, \boldsymbol{h_2}, \ldots, \boldsymbol{h_T}] \epsilon \mathbb{R}^{\boldsymbol{T \times d}}$:

$$e_t = \tanh(W_a h_t + b_a)\,,\alpha_t = \frac{exp(e_t^{\mathrm{T}} u_a)}{\sum_{k=1}^{\mathrm{T}} exp(e_t^{\mathrm{T}} u_a)},$$

where, W_a and b_a are the **weight** and the **bias** matrix,

u_a is a vector,

e_t represents the relevance of the hidden state h_t with respect to overall sequence.

The final context vector *c* is :

$$c = \sum_{t=1}^{\mathrm{T}} \alpha_t h_t \text{ and } Y_{SOH} = f(c)$$

***Y* is output** of the model,

f(.) represent fully connected layer

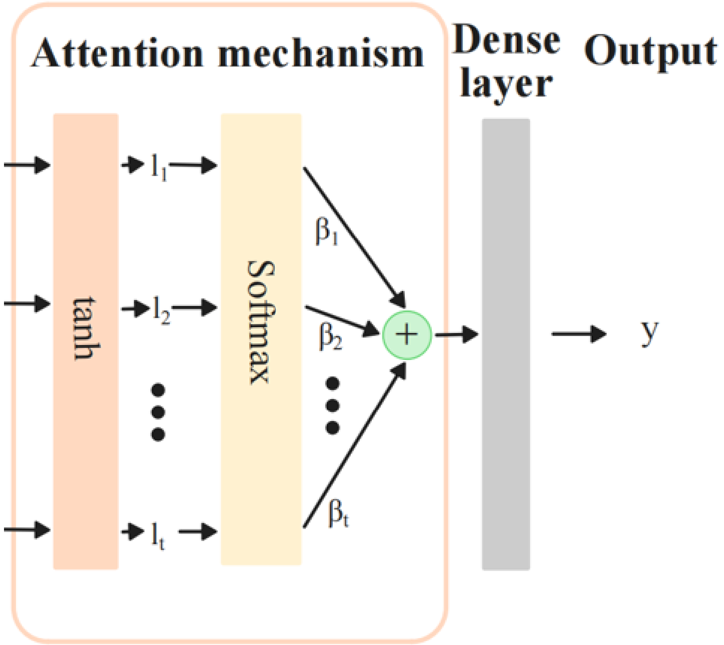


Figure. Attention mechanism with dense layer and final output

3.6 Model Execution Process

- Input → Feature Extraction: Two 1D Convolutional layers for local feature capture
- BiGRu Modeling: BiGRU layer to capture forward & backward long term dependencies
- Attention Mechanism: To emphasize critical time steps
- Dimensionality Reduction: Fully Connected (FC) layers: 1 output node

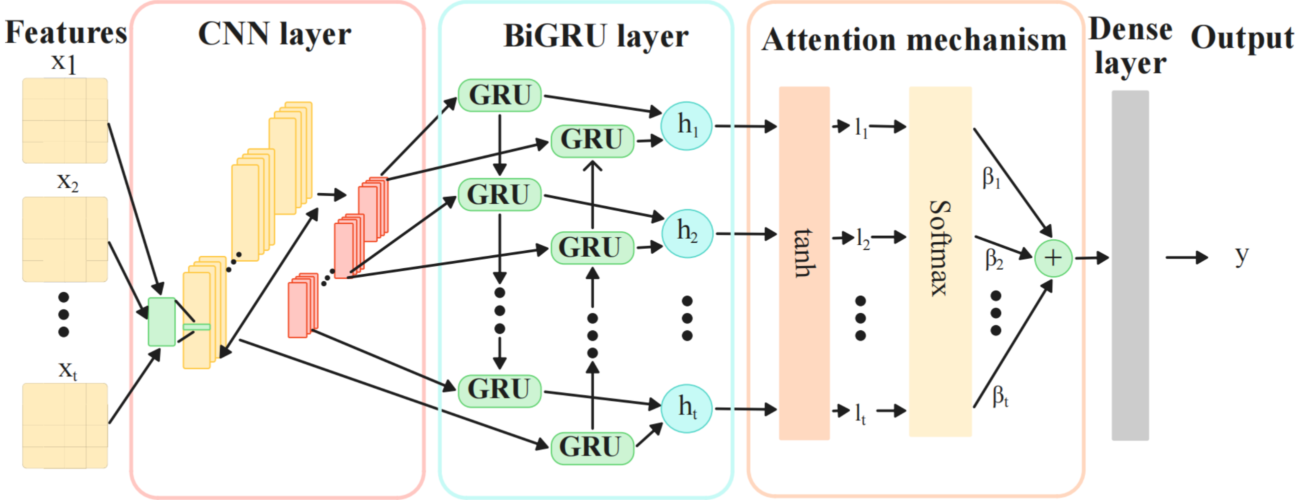


Figure. Full Process showing Execution of the Proposed Model. Feature matrices as input, CNN layer for feature extraction, Bidirectional Gated Recurrent Unit (GRU) followed by temporal attention and dense layer for final output of SOH

3.7 Training and Testing Setup

- Loss Function: Mean Squared Error (MSE)
- Optimizer: AdamW
- Learning Rate: 0.001
- Dropout: 0.25 (to prevent overfitting)
- Training: 1000 epochs, mini-batch size = 16
- Validation: Leave-One-Out Cross Validation (LOOCV)

Group	**Training Cells**	**Testing Cell**
Group I	CS2_36, CS2_37, CS2_38	CS2_35
Group II	CS2_35, CS2_37, CS2_38	CS2_36
Group III	CS2_35, CS2_36, CS2_38	CS2_37
Group IV	CS2_35, CS2_36, CS2_37	CS2_38

Table. CALCE dataset CS2 cells battery training & testing groups

3.8 Results and discussion : CS2-35

SOH Estimation Error:

MAE: 0.0072,

RMSE: 0.0121,

R^2: 0.9965

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*[a]	*Error*[b]
1D CNN	0.0269	0.0313	0.9770	404	420	16
Vanilla LSTM	0.0180	0.0216	0.9891	404	419	15
BiLSTM-GRU	0.0125	0.0172	0.9930	404	416	12
CNN-BiLSTM	0.0116	0.0171	0.9931	404	414	10
Proposed	**0.0072**	**0.0121**	**0.9965**	**404**	**406**	**2**

Table. Error table of CALCE dataset CS2-35 cell as testing cell

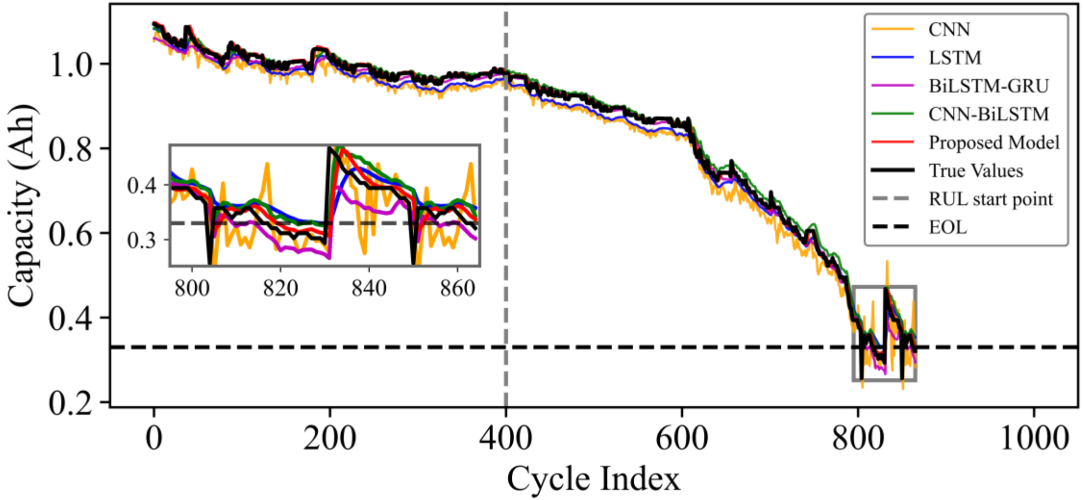


Figure. SOH predicition comparision of different models on CS2-35 cell with failure threshold

RUL Estimation Error:

Cycles: 2

3.8 Results and discussion : CS2-36

RUL Estimation Error:

Cycles: 3

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	*R^2*	*True*	*Pred.*[a]	*Error*[b]
1D CNN	0.0264	0.0310	0.9855	450	463	13
Vanilla LSTM	0.0208	0.0258	0.9900	450	457	7
BiLSTM-GRU	0.0131	0.02448	0.9910	450	455	5
CNN-BiLSTM	0.0199	0.0305	0.9859	450	442	8
Proposed	**0.0082**	**0.0145**	**0.9968**	450	**453**	**3**

Table. Error table of CALCE dataset CS2-36 cell as testing cell

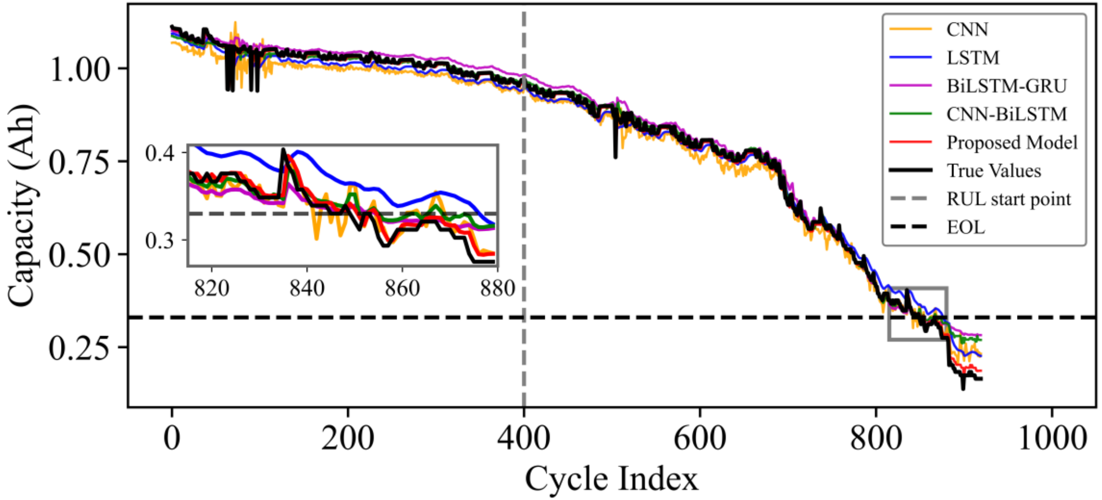


Figure. SOH prediction comparision of different models on CS2-36 cell with failure threshold

SOH Estimation Error:

MAE: 0.0082,

RMSE: 0.0145,

R^2: 0.9968

3.8 Results and discussion : CS2-37

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*[a]	*Error*[b]
1D CNN	0.0297	0.0321	0.9759	532	551	19
Vanilla LSTM	0.0112	0.0216	0.9941	532	537	5
BiLSTM-GRU	0.0110	0.0172	0.9944	532	536	4
CNN-BiLSTM	0.0173	0.0215	0.9891	532	541	9
Proposed	**0.0065**	**0.0124**	**0.9963**	532	**533**	**1**

Table. Error table of CALCE dataset CS2-35 cell as testing cell

RUL Estimation Error:

Cycles: 1

SOH Estimation Error:

MAE: 0.0065,

RMSE: 0.0124,

R^2: 0.9963

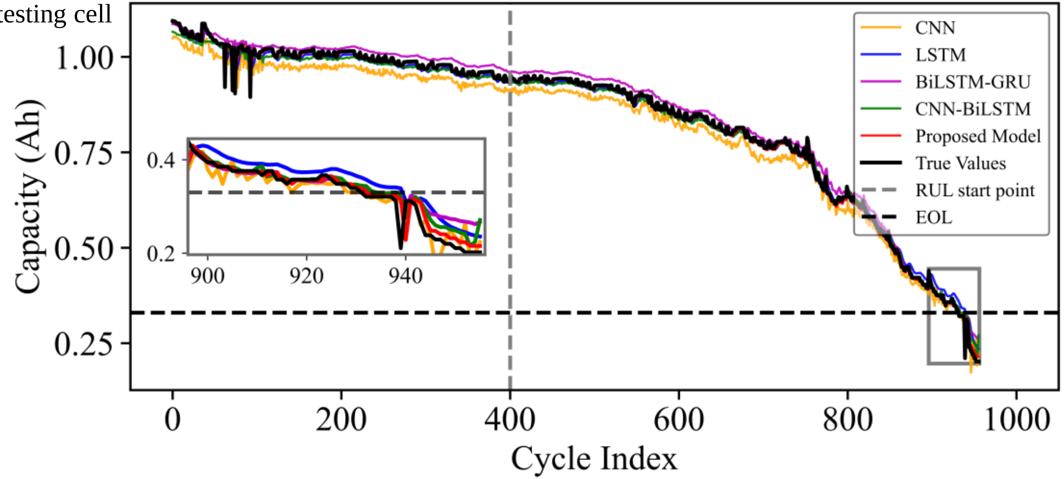


Figure. SOH prediction comparision of different models on CS2-37 cell with failure threshold

3.8 Results and discussion : CS2-38

SOH Estimation Error:

MAE: 0.0075,

RMSE: 0.0131,

R^2: 0.9958

Model name	SOH Prediction			RUL Comparison		
	MAE	*RMSE*	R^2	*True*	*Pred.*[a]	*Error*[b]
1D CNN	0.0264	0.0368	0.9671	540	553	13
Vanilla LSTM	0.0095	0.0149	0.9946	540	547	7
BiLSTM-GRU	0.0083	0.0133	0.9956	540	550	10
CNN-BiLSTM	0.0084	0.0134	0.9856	540	548	8
Proposed	**0.0075**	**0.0131**	**0.9958**	540	**540**	**0**

Table. Error table of CALCE dataset CS2-38 cell as testing cell

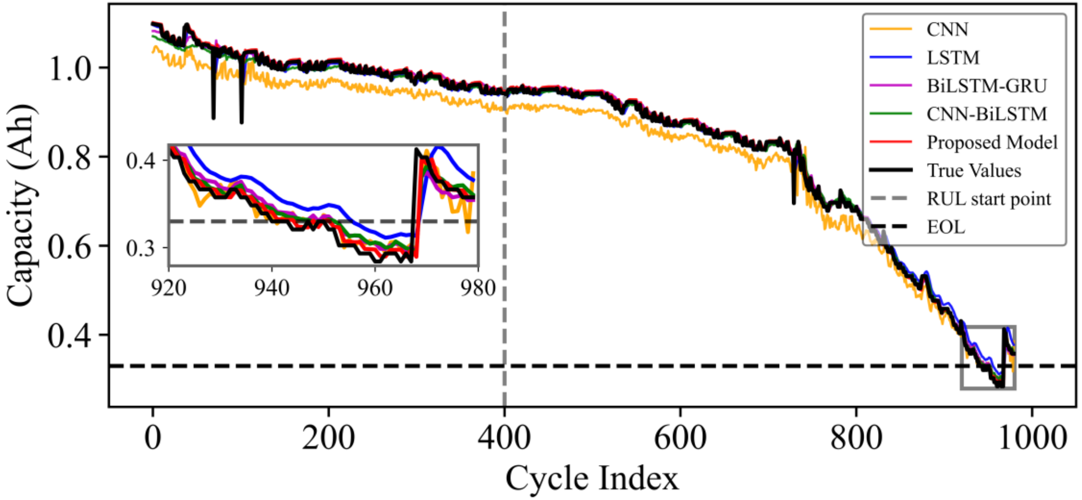


Figure. SOH prediction comparision of different models on CS2-38 cell with failure threshold

RUL Estimation Error:

Cycles: 0

3.9 Discussion on results

Accurate SOH & RUL Estimation

- Essential for assessing battery health, degradation, and lifespan
- Optimizes usage, improves safety, and reduces cost

Proposed Framework

- Hybrid CNN-BiGRU with Temporal Attention
- Captures short & long term dependencies
- Sliding window strategy for sequence generation
- Regularization via dropout and weight decay

Performance

- MAE:0.0072, RMSE: 0.0121, R^2: 0.9965
- Robust, reliable, and strong generalization

Impact & Future Work

- Enhances accuracy of SOH & RUL prediction
- Extends data-driven battery prognostics
- Develop lightweight model, real-time model for online BMS in Electric vehicles, Aerospace, Renewable energy storage systems

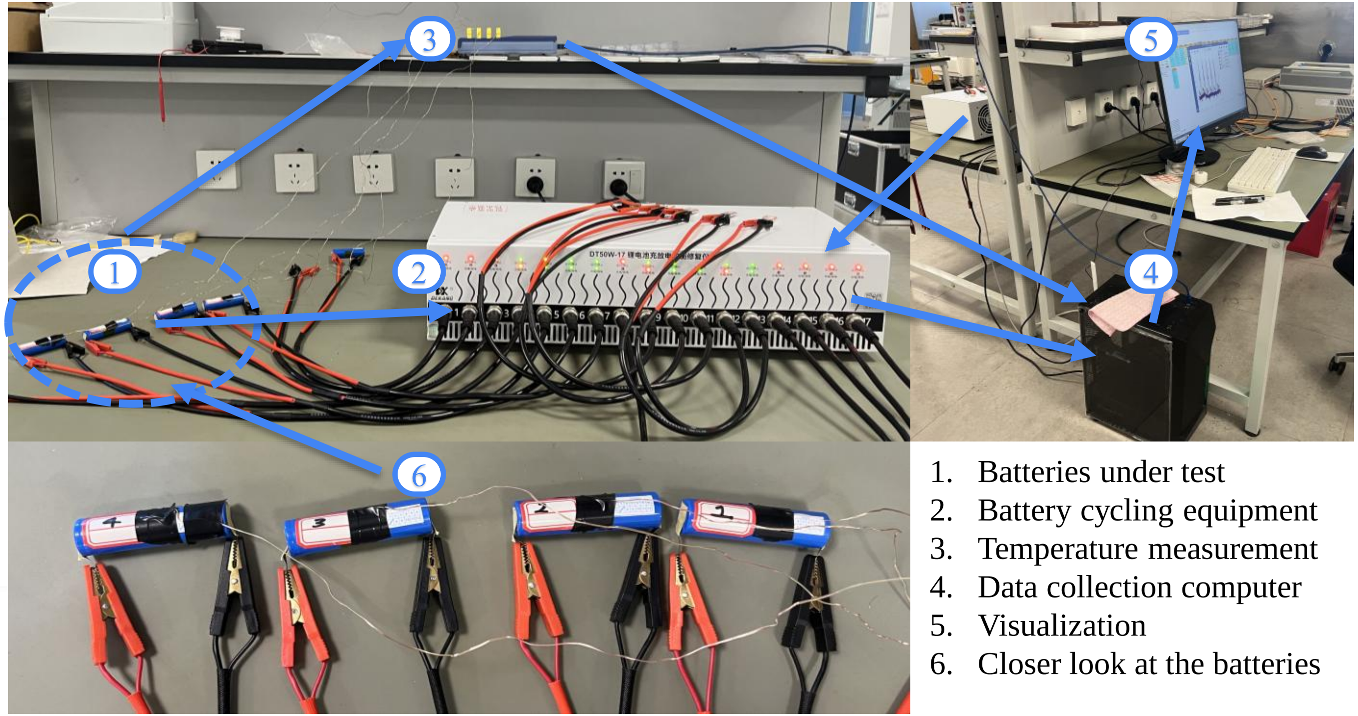
4

Experiments on Six datasets

4.1 Proprietary Dataset description

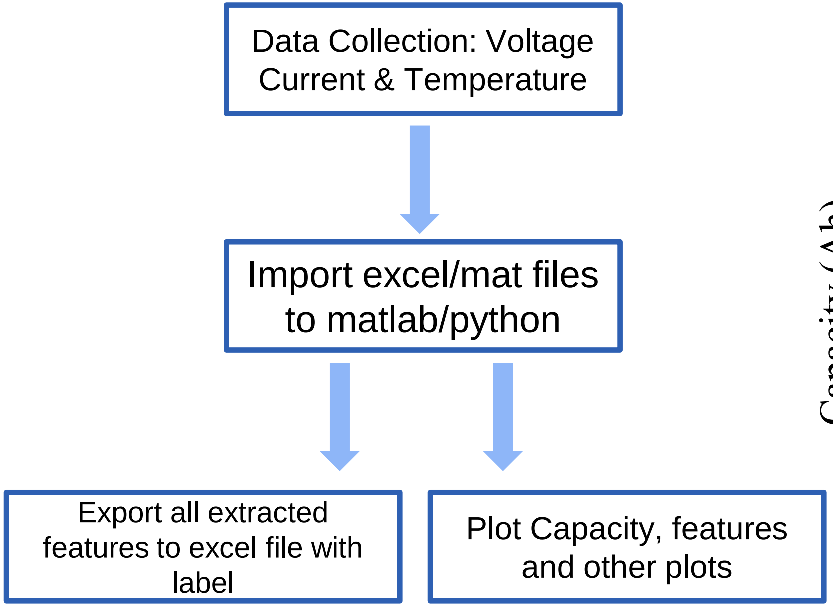
Parameter	350 mAh Pouch Cells	2200 mAh Cylindrical (18650) Cells
Cells used	Cell1, Cell2, Cell3, Cell4, Cell5	Cell1, Cell2, Cell3, Cell4
Form factor	Pouch	Cylindrical (18×65 mm)
Dimensions	35×20×5 mm (± tolerances)	Ø 18.3 ± 2 mm, Height 67 ± 2 mm
Weight	6 g	44 g
Rated capacity	350 mAh	2200 mAh
EOL capacity (30 % fade)	245 mAh	1540 mAh
Charging current	0.5C (175 mA) → CV at 4.2 V until 5 mA	0.5C (1.1 A) → CV at 4.2 V until 22 mA
Discharging current	1C (350 mA)	1C (2.2 A)
Charging protocol	CC-CV	CC-CV
Discharging protocol	CC	CC
Nominal voltage	3.7 V	3.7 V
Charge cut-off voltage	4.2 V	4.2 V
Discharge cut-off voltage	2.75 V	2.75 V
Rated charging temperature	10–45 °C	10–45 °C
Rated discharging temperature	–10–80 °C	–10–55 °C
Rated cycle life	500–800 cycles	300–500 cycles

4.2 Experiment setup and data collection for proprietary dataset

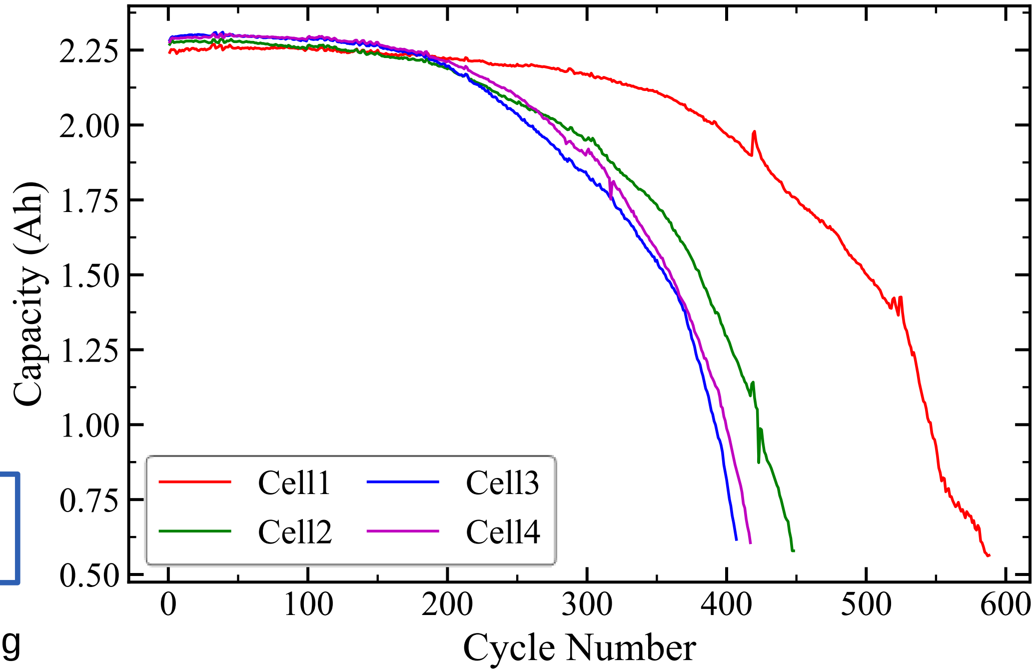


1. Batteries under test
2. Battery cycling equipment
3. Temperature measurement
4. Data collection computer
5. Visualization
6. Closer look at the batteries

4.3 Data processing and analysis of battery datasets

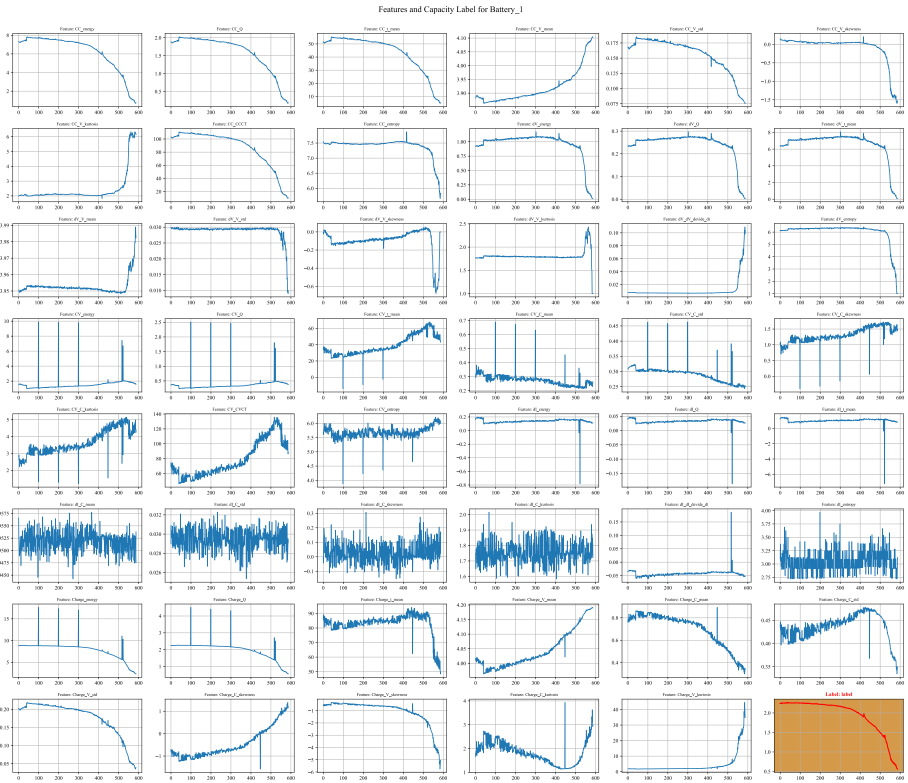


All six dataset undergo the same data processing and features analysis

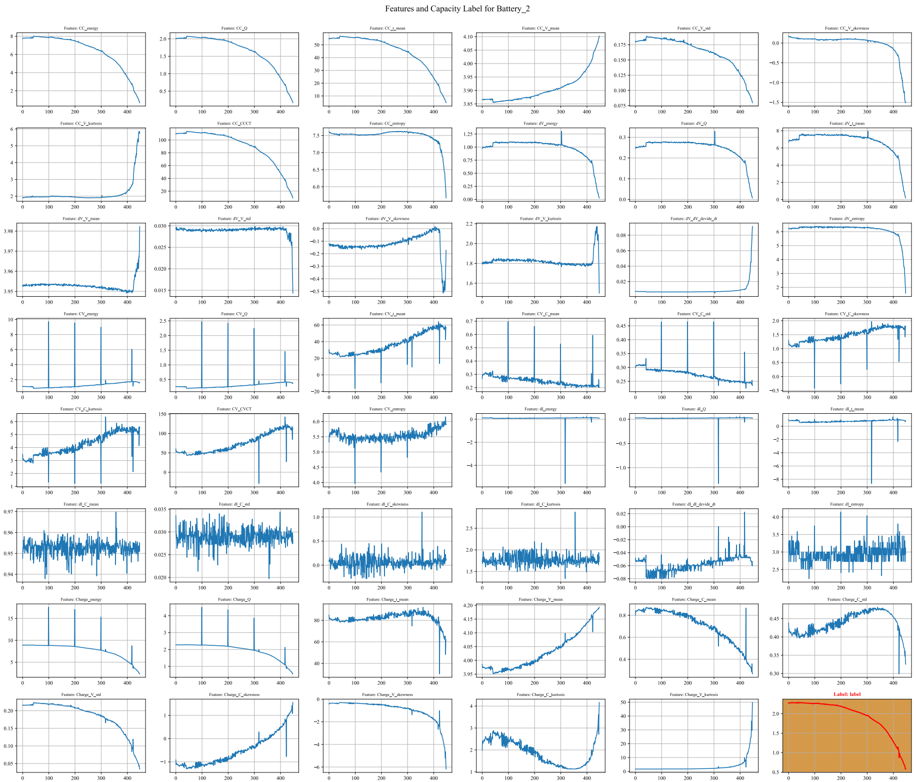


Capacity degradation plot of 2200 mAh dataset

4.3.1 Data processing and analysis of Cylindrical Cell 2200 mAh

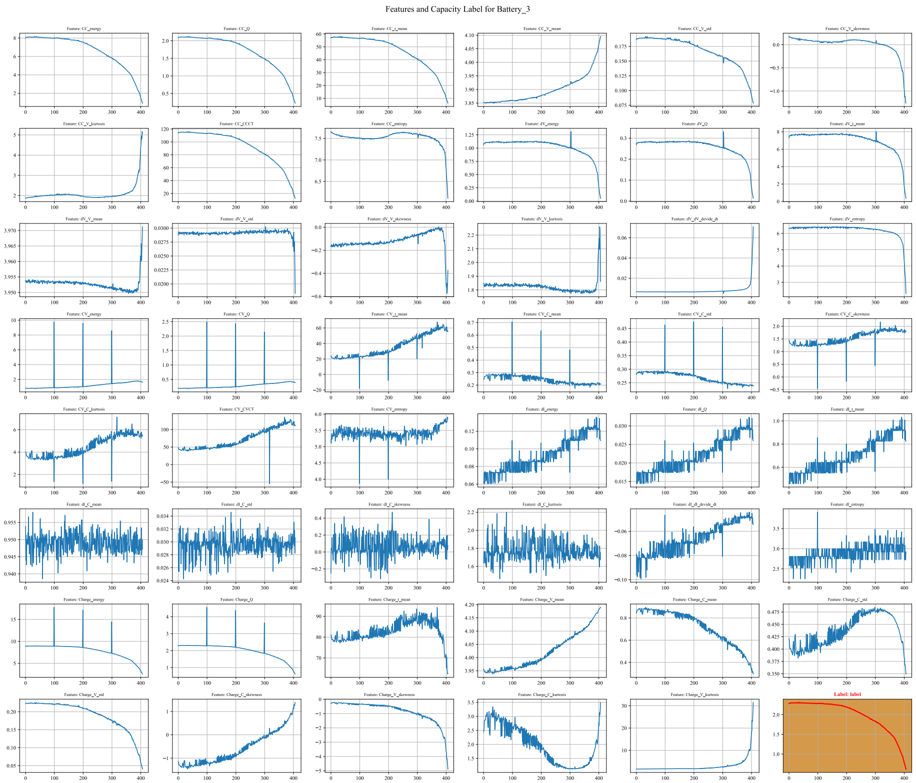


Lab dataset Handcrafted features of Cell1

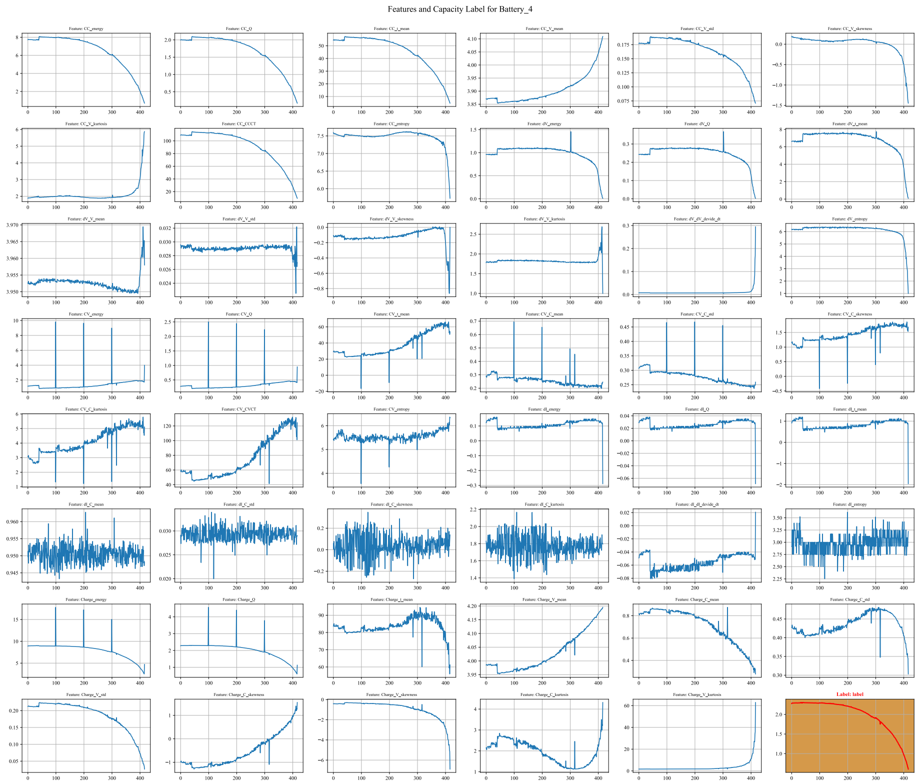


Lab dataset Handcrafted features of Cell2

4.3.1 Data processing and analysis of Cylindrical Cell 2200 mAh

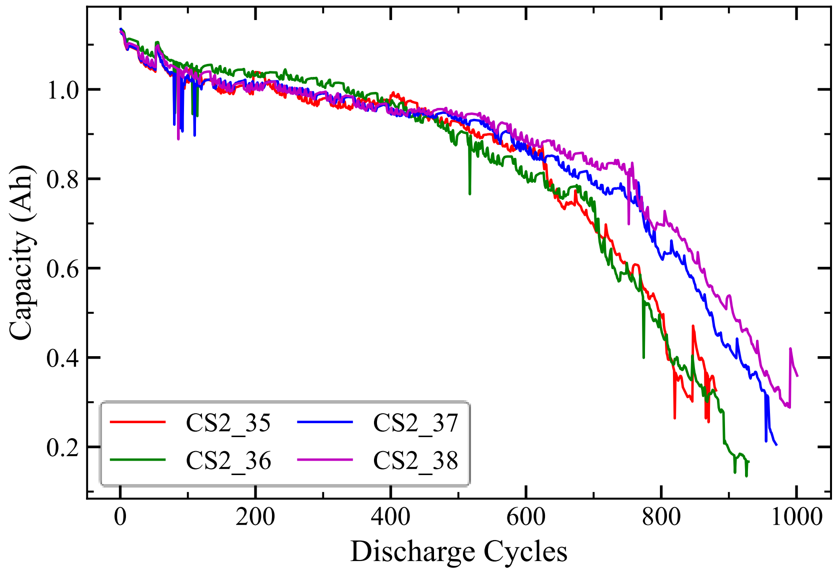


Lab dataset Handcrafted features of Cell3

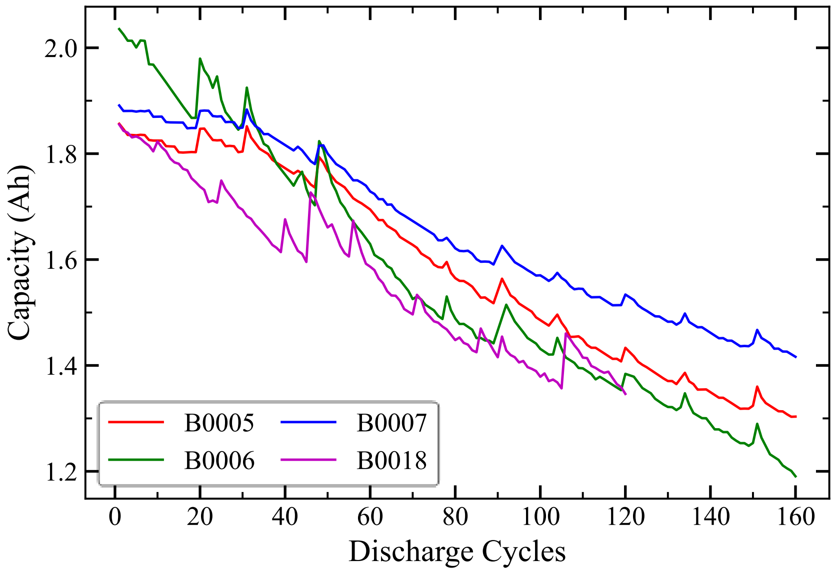


Lab dataset Handcrafted features of Cell4

4.5.1 Public dataset : CALCE and NASA capacity plots

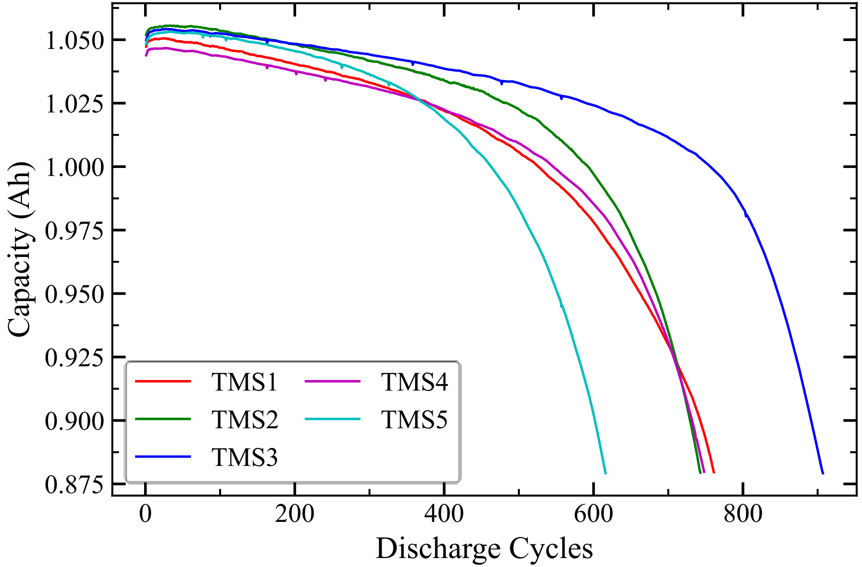


Capacity degradation plot of CALCE dataset

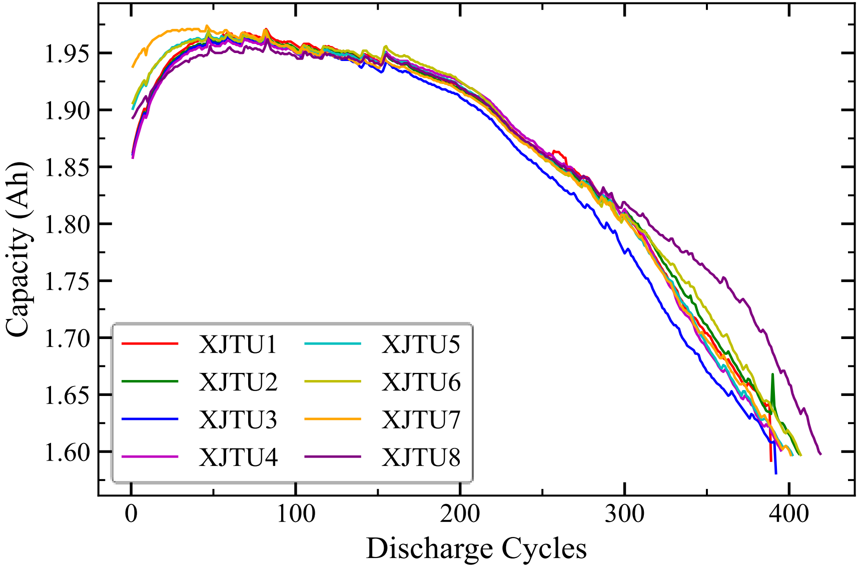


Capacity degradation plot of first batch of NASA dataset

4.5.2 Public dataset : MIT and XJTU capacity plots

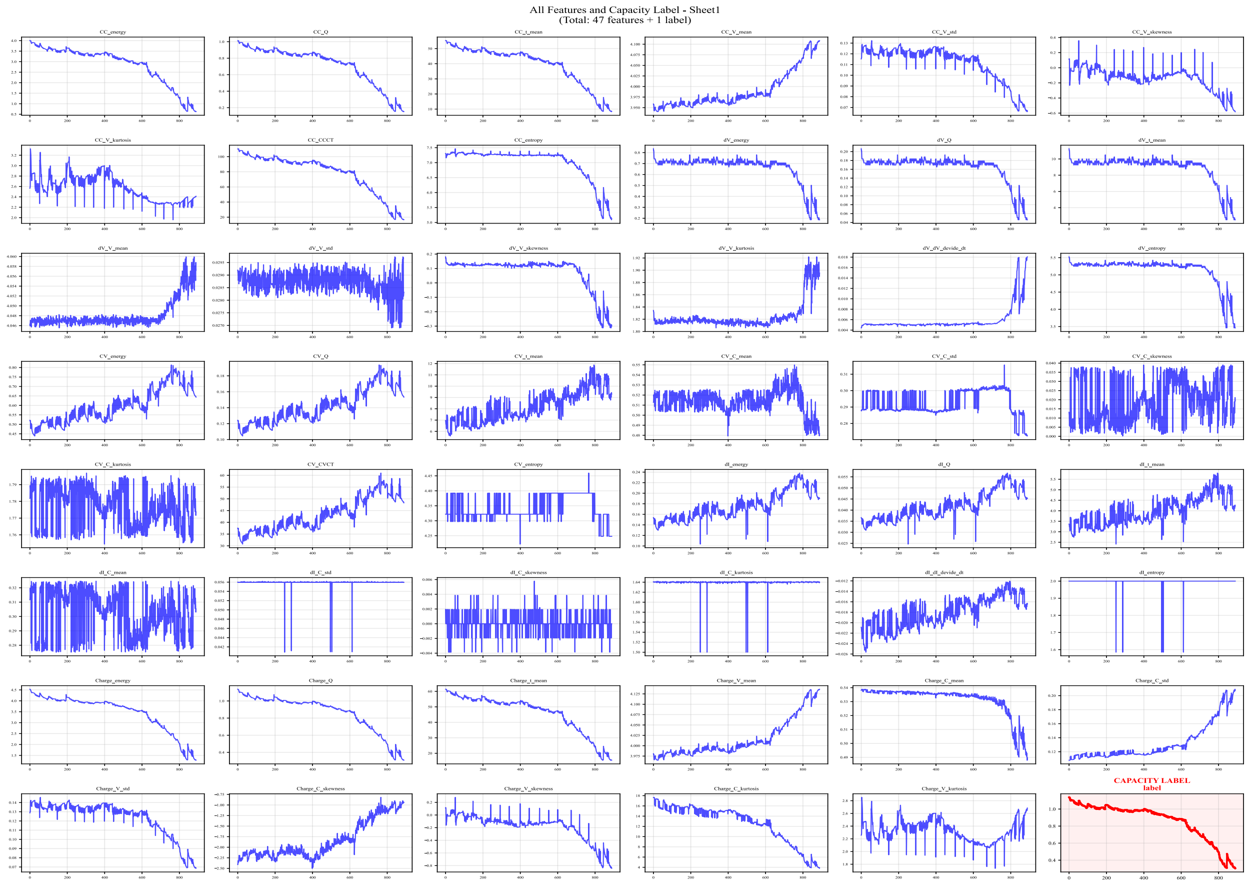


Capacity degradation plot of first batch of Toyota-MIT-Stanford dataset



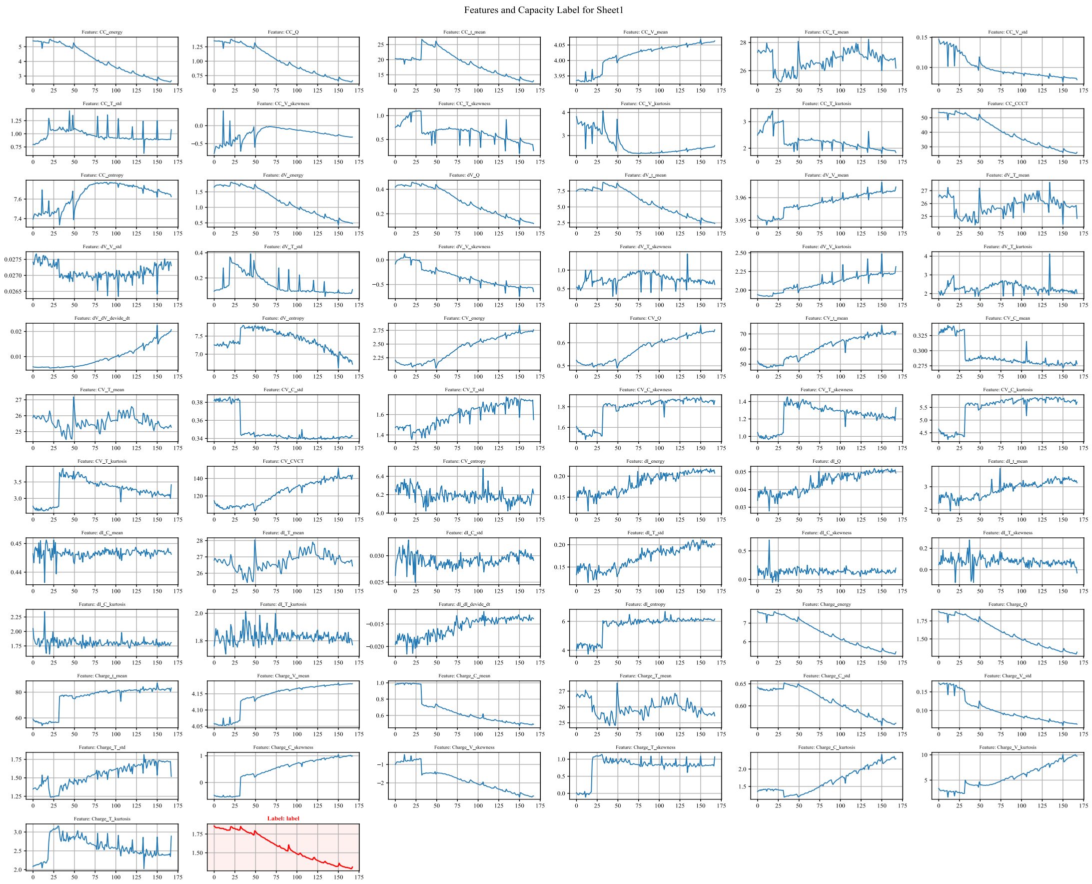
Capacity degradation plot of first batch of XJTU dataset

4.6.1 Public dataset : CALCE handcrafted features



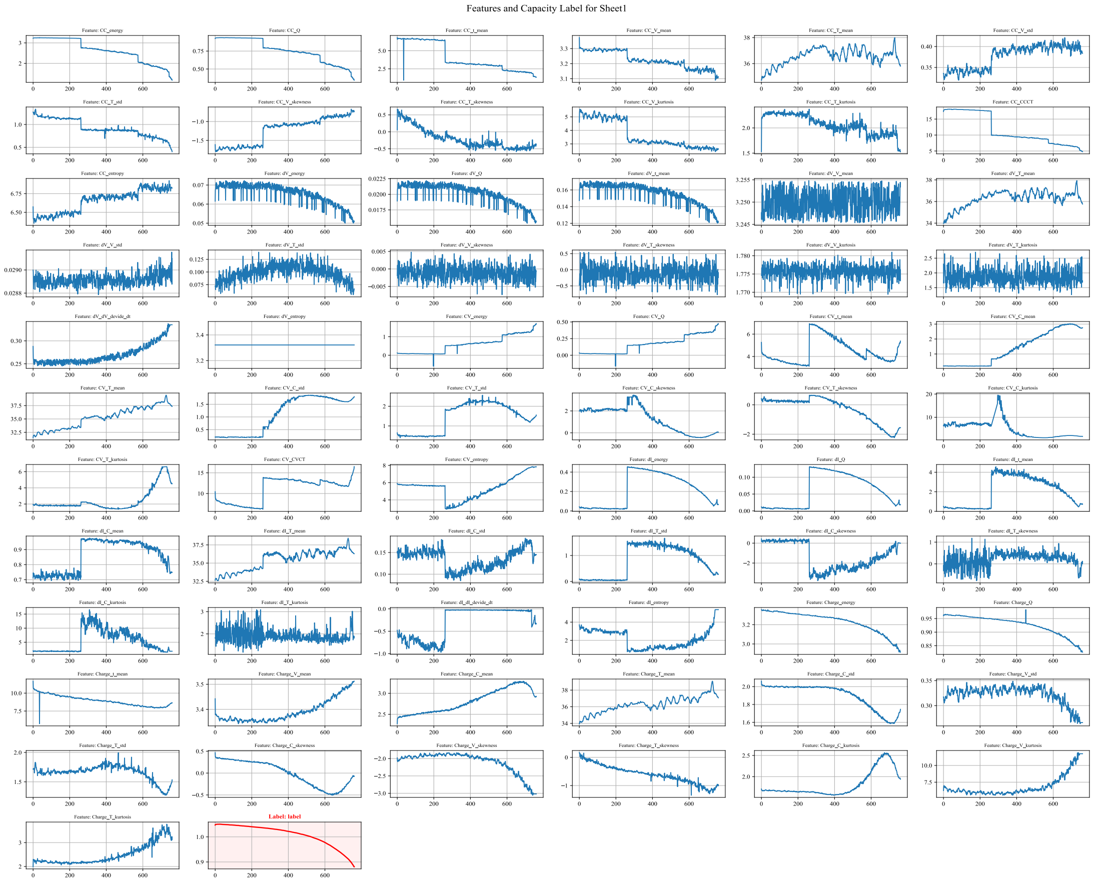
CALCE dataset Handcrafted features of CS2_35

4.6.2 Public dataset : NASA handcrafted features



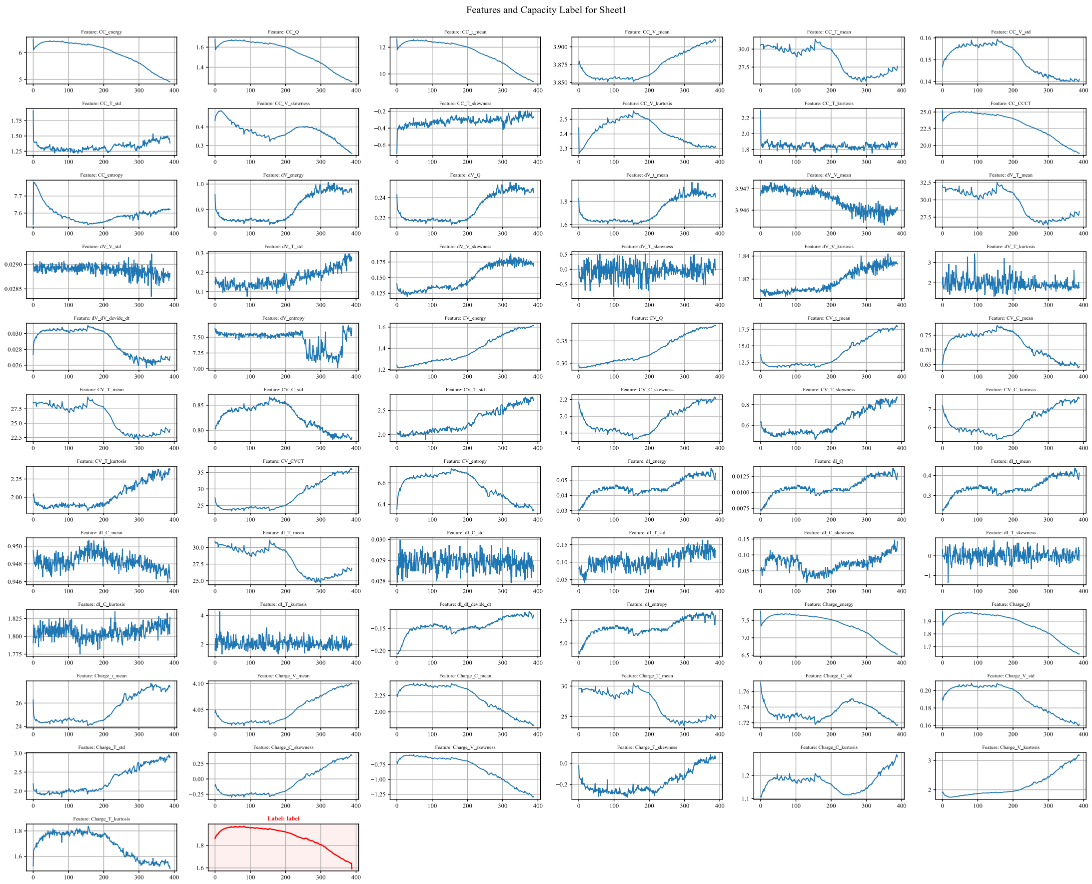
NASA dataset Handcrafted features of B0005

4.6.3 Public dataset : MIT handcrafted features



MIT dataset Handcrafted features of Cell1

4.6.4 Public dataset : XJTU handcrafted features



Handcrafted features of XJTU dataset Cell1 batch1

4.7 IPEformer : Inverted patch embedding time series transformer architecture

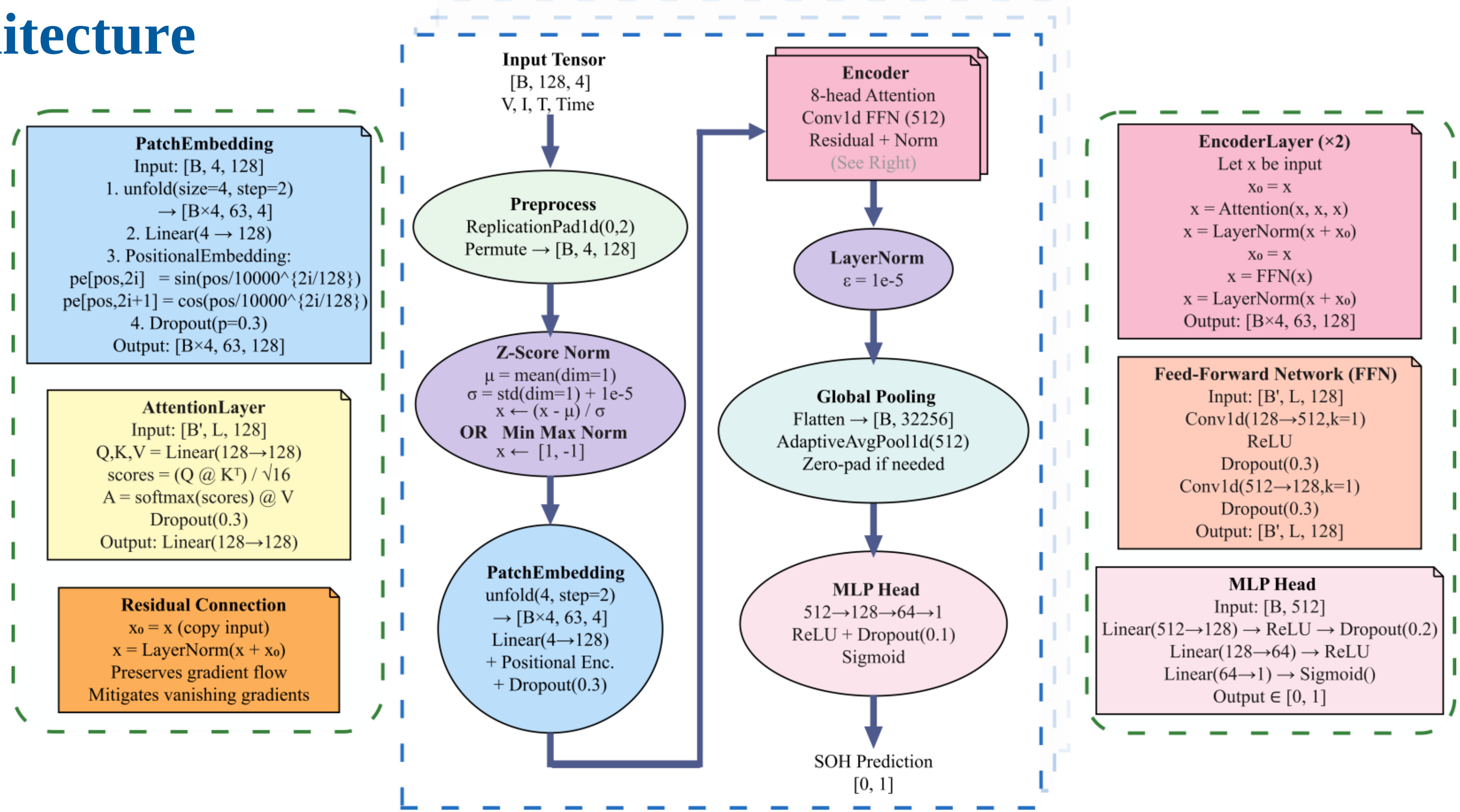
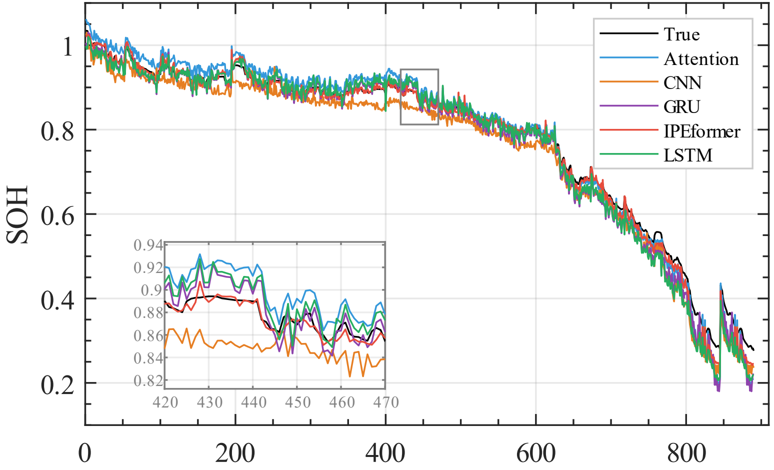
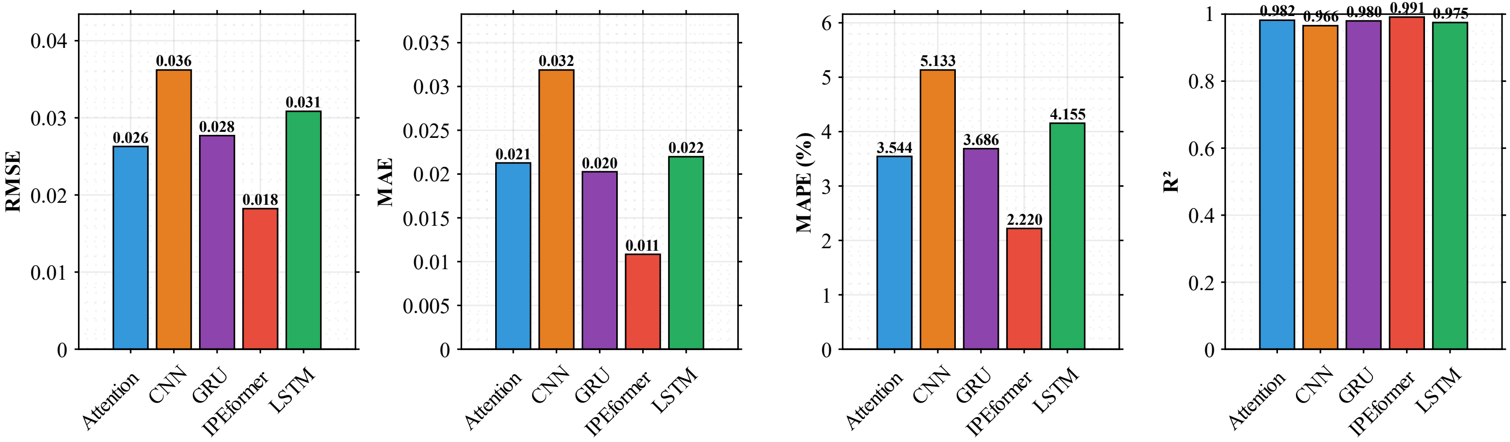


Illustration of IPEformer architecture

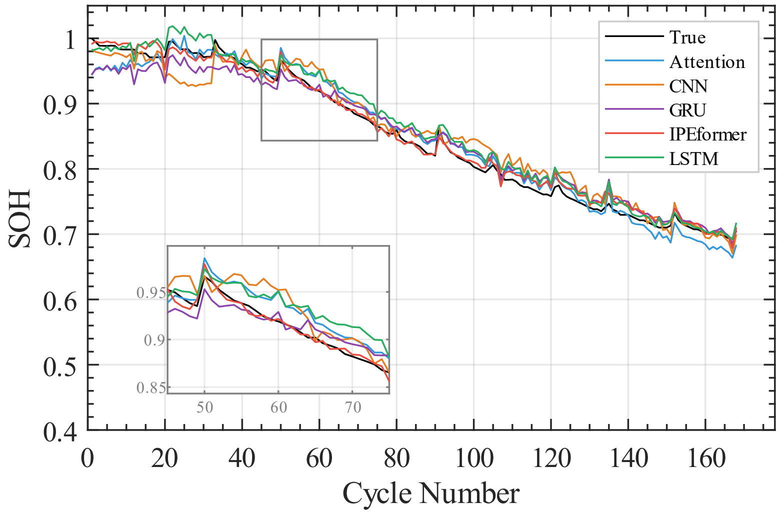
4.7 Results on CALCE CS2 35 battery

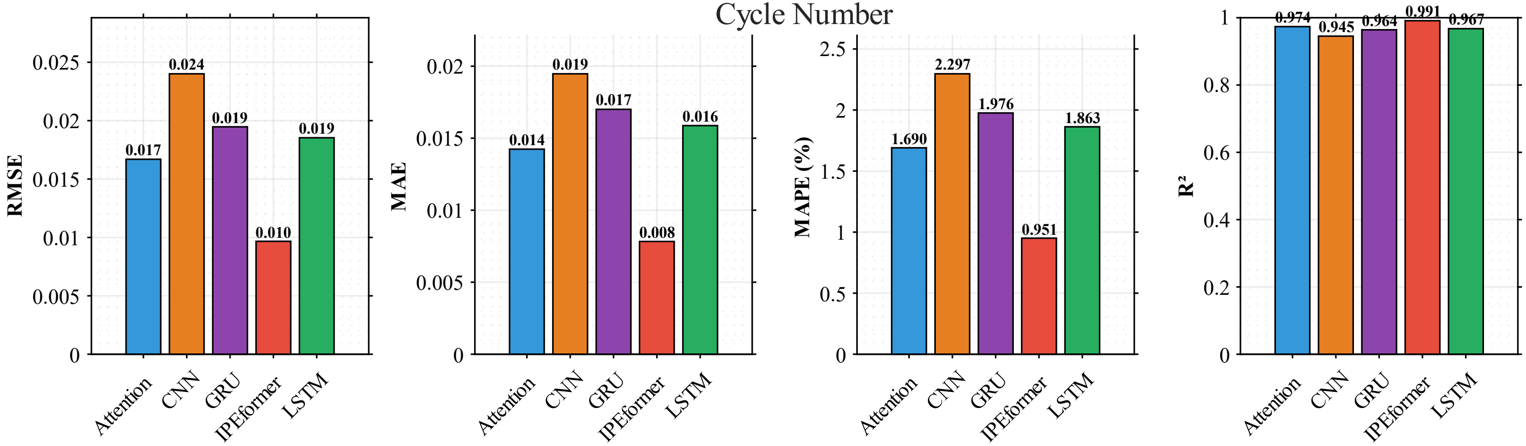
西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY



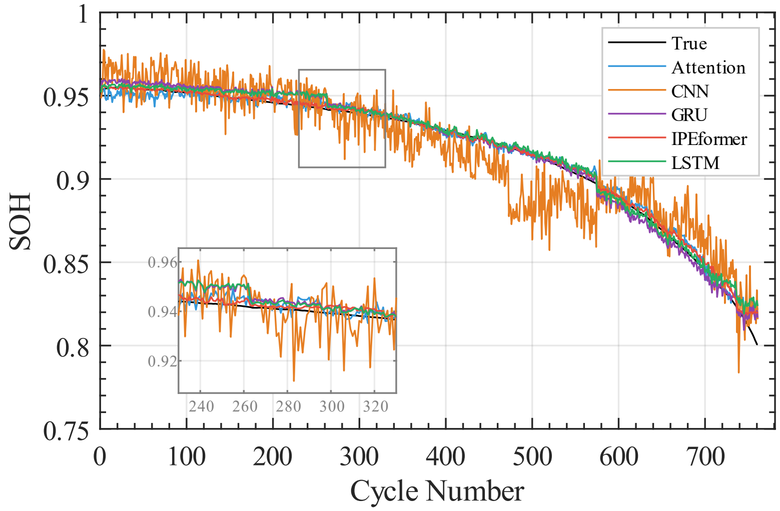


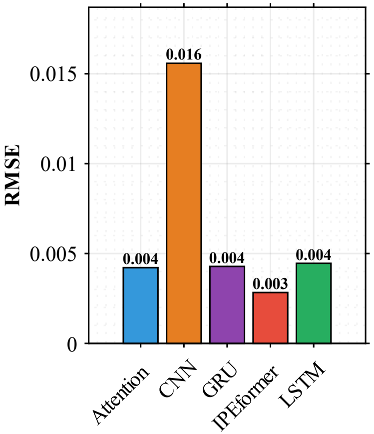
4.7 Results on NASA B0005 battery

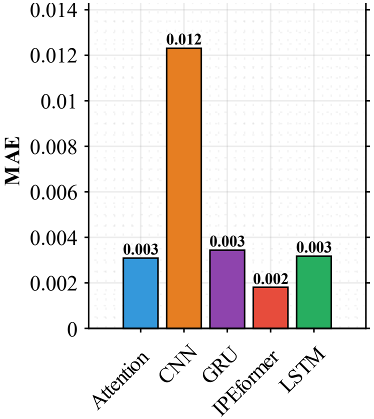


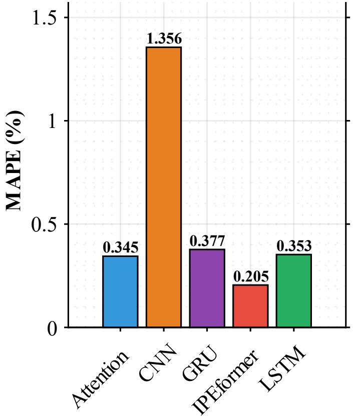


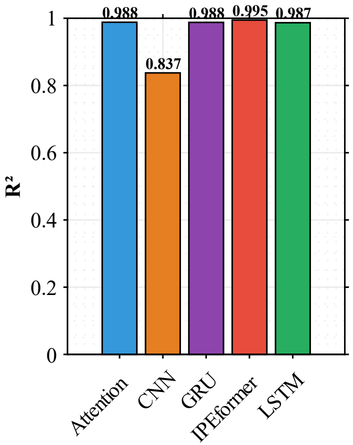
4.7 Results on MIT Cell1 battery



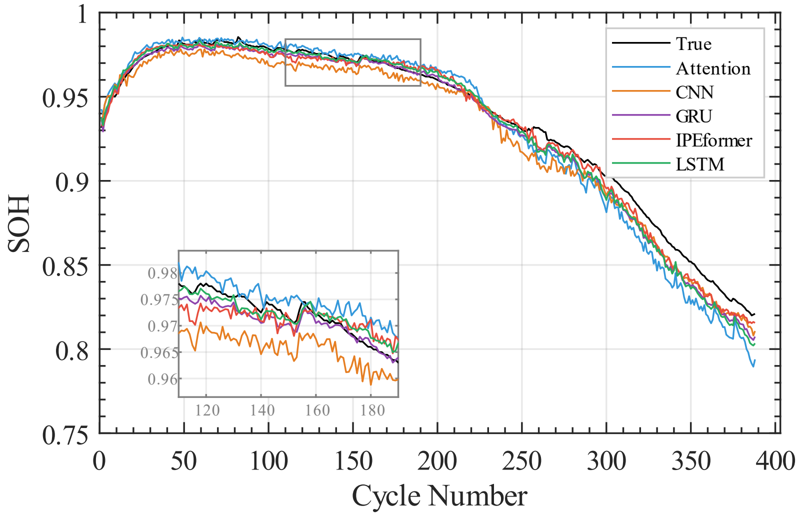


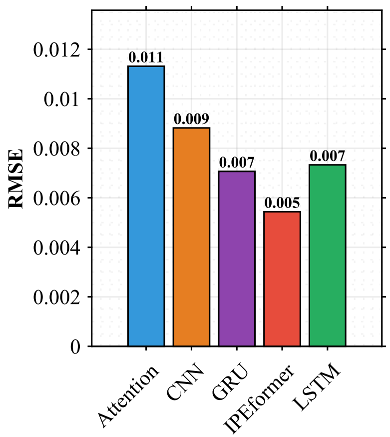


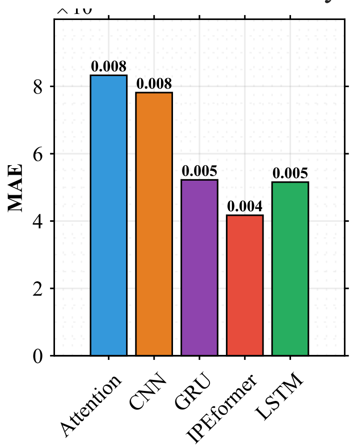


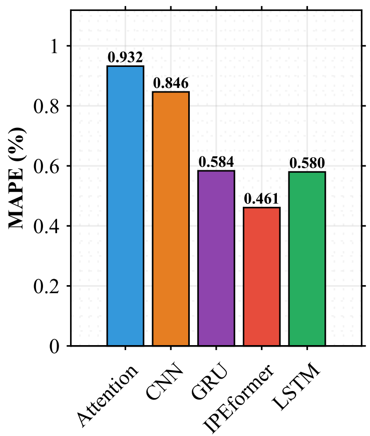


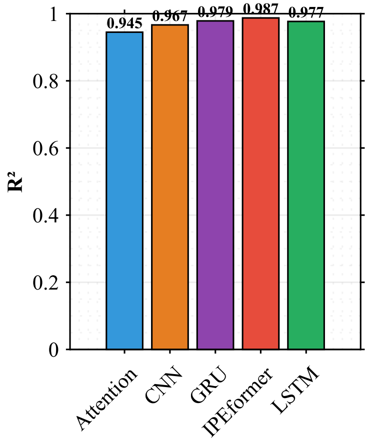
4.7 Results on XJTU Cell1 battery











5

Remaining work and Planning

4.1 The research content has not completed as planned, and the reason

1. Achieving SOH estimation performance clearly superior to the current state-of-the-art on all four public datasets (NASA, MIT-Stanford, CALCE, XJTU).
    - Reason: Greater-than-expected difficulty in surpassing literature baselines; extensive time was required for data preprocessing, health-indicator selection/construction, and experimentation with numerous deep-learning architectures.
2. Thorough analysis and performance optimization specifically on the NASA randomized dataset.
    - Reason: The NASA dataset exhibits higher complexity and noise; the current data loader, preprocessing pipeline, and training strategy need significant redesign and additional in-depth analysis to yield competitive results.
3. Completion of the laboratory battery cycling experiment for collecting real-world degradation data.
    - Reason: The physical cycling tests (especially long-term degradation tests under multiple operating conditions) have progressed slower than anticipated due to equipment scheduling, battery cell variability, and the need for repeated validation measurements.
    - Current progress: Approximately 80% completed. The remaining 20% (final cycles under extreme conditions and verification tests) is expected to require an additional 1 month (completion targeted for late December 2025). After data collection, post-processing, analysis, and integration into the model validation will be carried out immediately, and a dedicated journal paper summarizing the experimental dataset and preliminary findings will be prepared in parallel.

4.1 The existing problems, and the situation needed to explain

The main reasons certain research tasks were not completed according to the original plan are as follows:

- Encountered greater-than-expected difficulty in obtaining results that clearly outperform the current state-of-the-art methods reported in the literature.
- A significant amount of time was spent on data preprocessing, health-indicator construction/selection, and extensive experiments with various deep-learning architectures for SOH estimation on four public datasets (NASA, MIT-Stanford, CALCE, and XJTU).
- The NASA dataset proved particularly challenging; the current data loader and training pipeline need substantial improvement and more thorough analysis to achieve competitive performance.

Due to the above issues, the following adjustments to the original research plan are proposed:

- Slightly extend the experimental validation phase for the NASA dataset (additional 1–2 months).
- Shift partial focus from developing entirely new architectures to enhanced feature engineering and model optimization of the existing IPEformer framework, while still aiming to surpass literature baselines on all datasets.
- No major change to the overall research objectives, but the timeline for final comparative results and paper submission will be postponed accordingly.

4.7 The next research work arrangement

The follow-up research work for the remaining period is planned as follows:

- Perform targeted feature extraction and health-indicator optimization for the NASA, MIT-Stanford, CALCE, and XJTU datasets before final model training.
- Further improve the IPEformer model (architecture refinement, loss function design, attention mechanism optimization, etc.) to achieve higher accuracy and ensure it outperforms existing state-of-the-art models on at least three of the four datasets.
- Conduct systematic hyperparameter tuning (learning rate, batch size, dropout, transformer layers/depth, etc.) and ablation studies.
- Complete all remaining experiments, organize comparative results, and finish writing the journal manuscript.
- Expected completion timeline: core experiments concluded by the end of January 2026; full paper submitted by March 2026.

End of the Presentation !

Thank you for your attention!

Presenter : Amrit Dhakal

Northwestern Polytechnical University

Xi’an Shaanxi China

Mail: amritdhakal@mail.nwpu.edu.cn