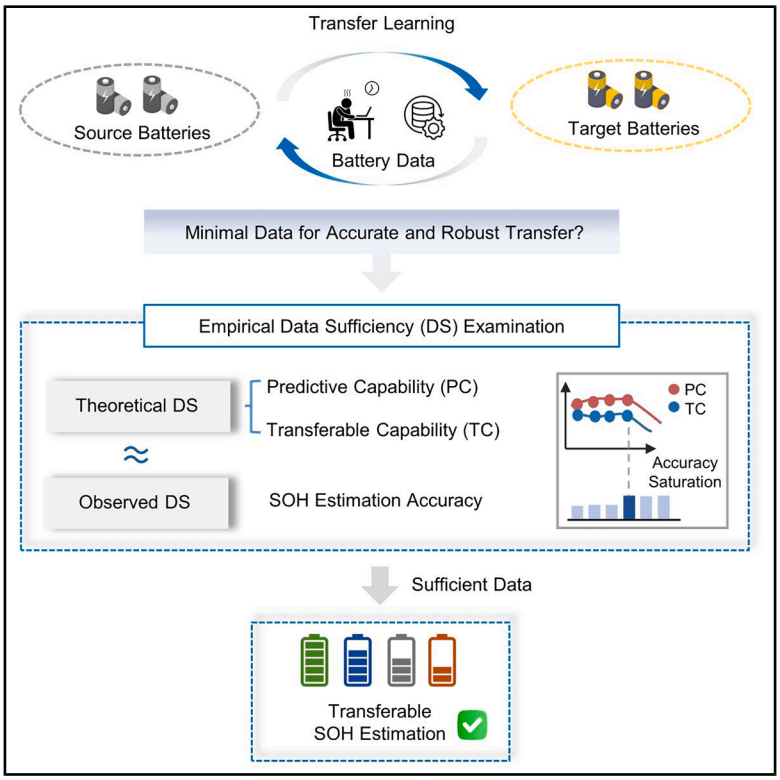
Cell Reports
Physical Science

Article

Data sufficiency for transferable lithium-ion battery periodical SOH estimation under resource constraints

Graphical abstract



Highlights

- Empirical data sufficiency identifies minimal data for accurate, transferable SOH estimation
- On average, 8% of lifetime data ensures accurate and transferable SOH estimation
- The results are validated on 310 batteries, 7 datasets, 6 chemistries, and 7 transfer cases
- Data sufficiency evaluation saves time and cost before developing battery management algorithms

Authors

Lin Su (苏琳), Shengyu Tao (陶晟宇), Weihan Li (栗伟涵), Dirk Uwe Sauer, Guangmin Zhou (周光敏), Xuan Zhang (张璇)

Correspondence

sytao@berkeley.edu (S.T.), guangminzhou@sz.tsinghua.edu.cn (G.Z.), xuanzhang@sz.tsinghua.edu.cn (X.Z.)

In brief

Su, Tao, et al. present an empirical data sufficiency (DS) framework for data amount examination in transferable battery SOH estimation. With examined DS, battery management algorithm developers may stop data curation processes when sufficient data have been collected, saving data curation costs, model development time, and model iteration periods.

Su et al., 2025, Cell Reports Physical Science 6, 102901
October 15, 2025
https://doi.org/10.1016/j.xcrp.2025.102901

CellPress

Cell Reports
Physical Science

CellPress
OPEN ACCESS

Article

Data sufficiency for transferable lithium-ion battery periodical SOH estimation under resource constraints

Lin Su (苏琳),[1] Shengyu Tao (陶晟宇),[1,2,5,*] Weihan Li (栗伟涵),[3,4] Dirk Uwe Sauer,[3,4] Guangmin Zhou (周光敏),[1,*] and Xuan Zhang (张璇)[1,*]

[1]Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, Guangdong 518055, China
[2]Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720, USA
[3]Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, Campus-Boulevard 89, Aachen, Germany
[4]Center for Ageing, Reliability and Lifetime Prediction of Electrochemical and Power Electronic Systems (CARL), RWTH Aachen University, Campus-Boulevard 89, Aachen, Germany
[5]Lead contact
*Correspondence: sytao@berkeley.edu (S.T.), guangminzhou@sz.tsinghua.edu.cn (G.Z.), xuanzhang@sz.tsinghua.edu.cn (X.Z.)
https://doi.org/10.1016/j.xcrp.2025.102901

SUMMARY

Battery state-of-health (SOH) estimation is vital for the safety of energy storage systems, yet full-life-cycle data curation remains resource intensive. Here, we present an empirical examination of data sufficiency (DS) to identify the data amount needed for SOH estimation algorithms with anticipated predictability and transferability. DS is defined as a linear function of normalized predictability and transferability of physical features. The effectiveness of DS is validated on 7 datasets (involving 310 batteries over 300,000 cycles), encompassing 6 materials (i.e., LFP, NCM811, NCM333, NCM523, NCA, and LCO) and 7 transfer scenarios (i.e., temperature, charging rate, discharging rate, and cutoff voltage). Results show that, on average, no more than 8% of lifetime data can achieve a median mean absolute percentage error (MAPE) of 1% under the investigated transfer scenarios, with the calculated DS aligning with post hoc DS. This work suggests the careful evaluation of DS for building data-driven battery algorithms before massive, expensive, and time-consuming data curation.

INTRODUCTION

Lithium-ion batteries (LIBs) have been widely adopted in extended industries, including electric vehicles, renewable energy storage, and portable electronics.[1–5] A common challenge across these applications is to estimate the remaining capacity of batteries under given and stochastic operation conditions, which ensures efficient and safety-informed use.[6,7] Data-driven state-of-health (SOH) estimation offers significant advantages by enabling timely health monitoring and maintenance decisions. These methods can also provide insights for periodic battery checks, help optimize battery operation, and enhance safety, thereby contributing to a safer, more durable and sustainable battery energy ecosystem.

Data-driven battery SOH estimation has garnered widespread attention in both academia and industry.[8,9] The fundamental concept is to estimate the battery's remaining capacity by computing the SOH as the ratio of remaining to initial capacity.[10] However, in real-world applications, batteries operate under diverse and complex conditions.[11,12] Models developed from past data are often not directly applicable to new batteries under different operating scenarios. With advances in data science and artificial intelligence, data-driven methods have been applied for battery SOH estimation. The main idea is to obtain a statistical relationship between the already measured data and the aging state of batteries through machine learning algorithms.[13–21] However, these methods typically require large volumes of historical data, which creates challenges for data storage and computation, especially in edge devices with considerable resource constraints. To reduce data curation demands, some studies have introduced data generation or transfer learning strategies.[22] Yet, the fidelity of generated data remains challenged by a lack of understanding of underlying degradations.[23–27] In parallel, numerous studies have attempted to transfer pre-trained models to new operating conditions by retraining them using limited data from those conditions. Unfortunately, existing studies rely on subjective choices regarding the amount of training data, for example, using the first 100 cycles, 30%, or 66% of the battery life to retrain or fine-tune models.[28–31] These approaches lack justification and are inconsistent across datasets and operating conditions. As a result, a key limitation remains unresolved: there is no unified framework to determine how much early-cycle data from the target domain is sufficient to enable reliable model adaptation for accurate SOH estimation.

This gap presents a fundamental dilemma in practical applications: without a clear benchmark for early data sufficiency (DS), it

is difficult to decide when a pre-trained model can be confidently deployed to a new battery system under different operating conditions. The lack of such a criterion limits the efficiency of transfer learning and results in unnecessary data curation, repeated retraining, and high storage costs. This issue is particularly critical in resource-constrained environments, such as on-board systems or distributed controllers, where minimizing data curation, storage, and processing is key to efficient edge deployment. Therefore, it becomes essential to record and utilize only the early but informative portion of data under new operation conditions rather than collecting and storing them throughout the entire lifespan. However, the predictive capability (PC) of early battery data to map to SOH has not yet been clearly defined and quantified.[32,33] Given the sensitivities of battery degradation over varied operating conditions, the cross-domain stability of early data for SOH estimation, i.e., the transferable capability (TC), also remains unresolved.[34–38] Therefore, investigating the PC and TC of using limited early data of new operation conditions to adapt the trained model to those conditions is crucial, as it answers a fundamental question of when sufficient data can be collected. This is of great significance to the battery community, as it enables economic and reliable SOH estimation for novel battery materials research and management algorithm integration under varying conditions, particularly in determining the sufficient time point at which a reliable and adaptable SOH estimation model can be possible, given limited resource constraints.

Here, we propose an empirical study for DS to evaluate the adequacy of data from the accessible domain in predicting battery SOH, addressing both the predictability of capacity outcomes and their transferability across operating conditions, as demonstrated in Figure 1. Figure 1A illustrates the challenge of extensive data collection requirements to obtain a predictive model in battery management, compounded by the robustness challenges posed by diverse operating conditions. Determining the optimal point to adapt to target domain degradation is critical for achieving accurate SOH estimation while reducing resource and storage requirements for model deployment. Figure 1B presents the methodological workflow used in our study to address the DS. DS is defined as the minimum amount of data required to achieve anticipated estimation accuracy, beyond which additional data do not significantly improve the predictive performance, referred to as observed DS (ODS). We introduce theoretical DS (TDS), formulated as a time-evolving linear function of DS components, quantifying the PC and transferability of features. DS components, including PC and TC, assess battery degradation behavior and its adaptability across operating conditions, represented as a linear function of normalized predictability and normalized transferability. The proposed empirical study for DS has been validated in 7 widely studied battery datasets covering 6 materials, 7 operational conditions, and 7 transfer scenarios, showing that our model achieves less than 1% median mean absolute percentage error (MAPE) in predicting the capacity of target batteries using approximately 8% of the lifetime data, as a conservative guideline derived from the observed data space. Such success can be attributed to the strong alignment between the ODS and TDS, with an average error of less than 0.4%. The result also incorporates sensitivity analysis across different SOH levels, yielding an average error of 30 cycles across the datasets. These findings underscore the critical role of DS in battery SOH estimation and present a generalizable framework that enhances the reliability and adaptability of robust estimation tasks. This study provides insights for other data-driven diagnostic fields by emphasizing the importance of establishing DS assessment criteria to ensure predictive accuracy and generalizability while reducing unnecessary data storage, thus avoiding redundant data with limited informational value, especially in resource-constrained scenarios.

RESULTS

Datasets and feature engineering

To validate the feasibility and generalizability of the proposed empirical study for DS, we compile 7 datasets covering 6 battery materials, 7 operational conditions, and 7 transfer scenarios. An overview of the seven datasets is provided in Table S1. It includes a nine-stage fast-charging dataset[39] generated under four distinct temperatures, 25°C, 35°C, 45°C, and 55°C, along with six publicly available datasets.[40–45] In Figure 2A, the charging voltage curves across all seven datasets exhibit an increasing trend with the state of charge (SOC). However, the rate of change, curve shapes, and voltage ranges vary significantly due to differences in battery materials, operational conditions, and testing environments. Details of the charging and discharging conditions for the public datasets are provided in Note S1. Additionally, voltage curves across cycles reveal a clear degradation trend, with different cycles represented by color gradients to emphasize progression (Figure 2A). The color of the curve transitions from light to dark, representing the progression of battery life from early to late stages. Batteries with higher cycle counts can charge less, meaning that curves with darker colors correspond to a lower SOC. The batteries in each dataset operate under diverse conditions, such as temperature, cutoff voltage, charge rate, and discharge rate. Assuming that the red curves represent known batteries and the blue curves represent unknown batteries, the capacity of both groups decreases progressively with cycles, exhibiting distinct degradation patterns (Figure 2B). This assumption is reasonable given that battery cycling is time consuming; thus, manufacturers often perform accelerated tests and then transfer insights derived from them to the conditions of interest.

To establish this transfer relationship, we classify the batteries into two domains, the source domain and the target domain, based on operating conditions, with specific criteria provided in Table S2. From each dataset, we select the batteries with the shortest and longest lifespans to represent aggressive and moderate degradation trends, respectively (Figure 2C). This classification highlights the considerable range of battery lifespans observed, with the longest lifespan reaching 2,000 cycles, while the shortest lifespan is merely around 170 cycles. To reflect internal degradation mechanisms, interpretable features are extracted from charging voltage curves with prominent degradation trends. Although these features differ in their mathematical formulations, they retain consistent physical meanings across datasets and capture key electrochemical processes, including charging acceptance, polarization rate, resistance, and various

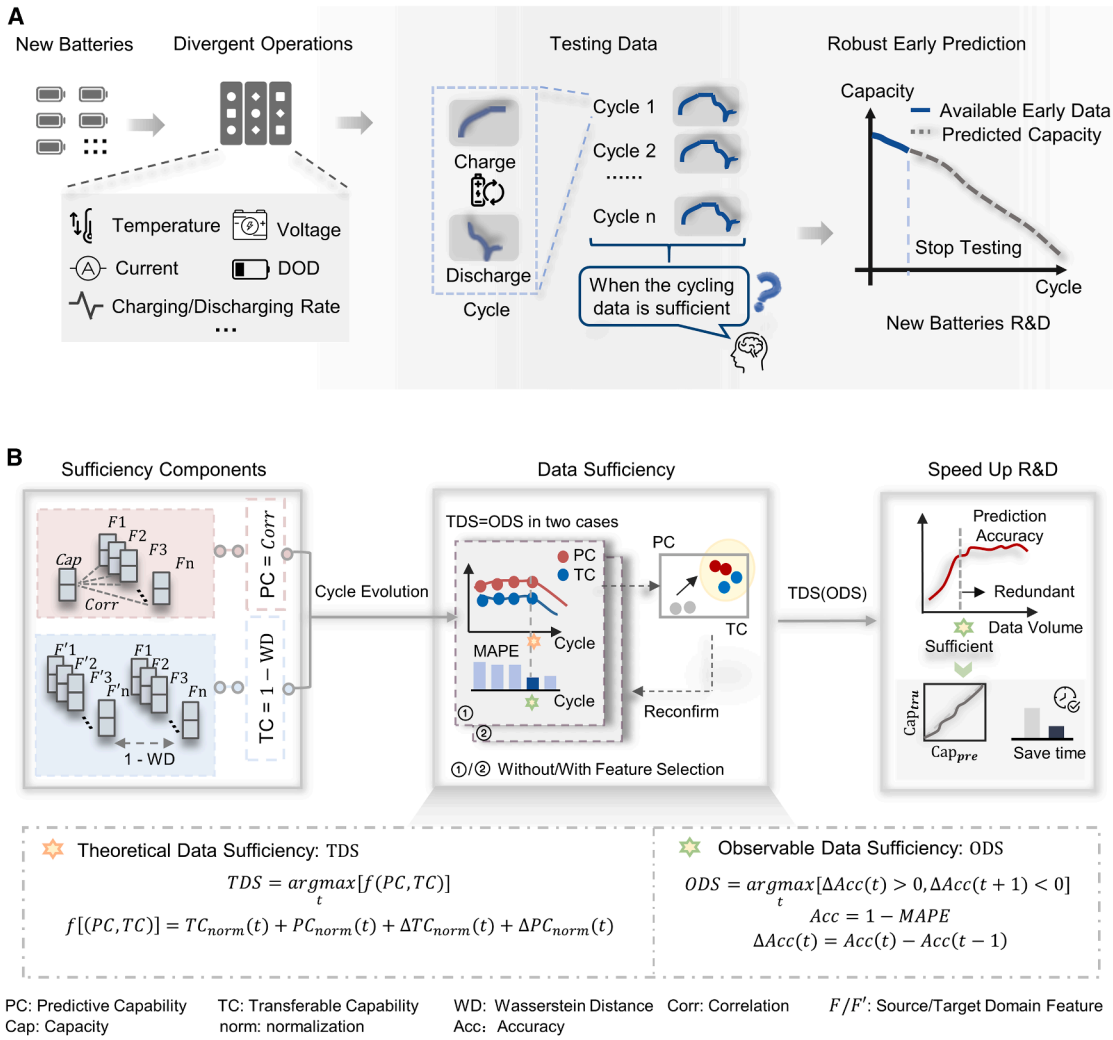


Figure 1. The empirical study of data sufficiency

(A) Illustration of the challenges associated with data testing volume, coupled with robustness challenges arising from diverse operating conditions, such as temperature, voltage, current, and depth of discharge (DOD). The process determines when cycling data become sufficient to start predicting, using already available early data to predict the subsequent battery SOH.

(B) The approach defines predictive capability (PC) and transferable capability (TC) as data sufficiency (DS) components derived from physically interpretable features. The evolution of PC and TC over cycles establishes the theoretical DS, while the relationship between estimation accuracy and the number of cycles defines the observed DS. Formulas are explained in the methods section. Features located in the top right quadrant of the PC-TC plot are selected to simplify the model and validate the identified DS through feature selection.

types of polarization (electrochemical, ohmic, and concentration). Details of feature names and physical interpretations are provided in Tables S3 and S4, with extraction processes illustrated in Figure S1. Derived from the ninth charging stage in our dataset as an example, we observe substantial variations in feature distributions across different temperature conditions (Figure 2D). Distributions of other features are shown in Figures S2–S4. Estimating SOH using limited data remains challenging because the capacity fade during initial cycles is minimal (Figure 2E), even at elevated temperatures. For instance, high temperatures accelerate capacity fade but result in only a 3.7% reduction during early cycles. Despite limited early changes, extracted features effectively capture degradation trends. Modeling capacity as the dependent variable and the extracted features as independent variables reveals an almost linear decrease in capacity as a function of features extracted from stage 1, with a Pearson correlation coefficient of −0.96 (Figure 2F). Similarly, other features have a strong correlation with early capacity, as shown in Figure S5. These results underscore the utility of early-stage features in capturing degradation trends, even if capacity changes are negligible at the early stage.[43]

Predictive capability and transferable capability

When new operating conditions are different from the existing conditions, it is crucial to focus on the PC and TC at the same

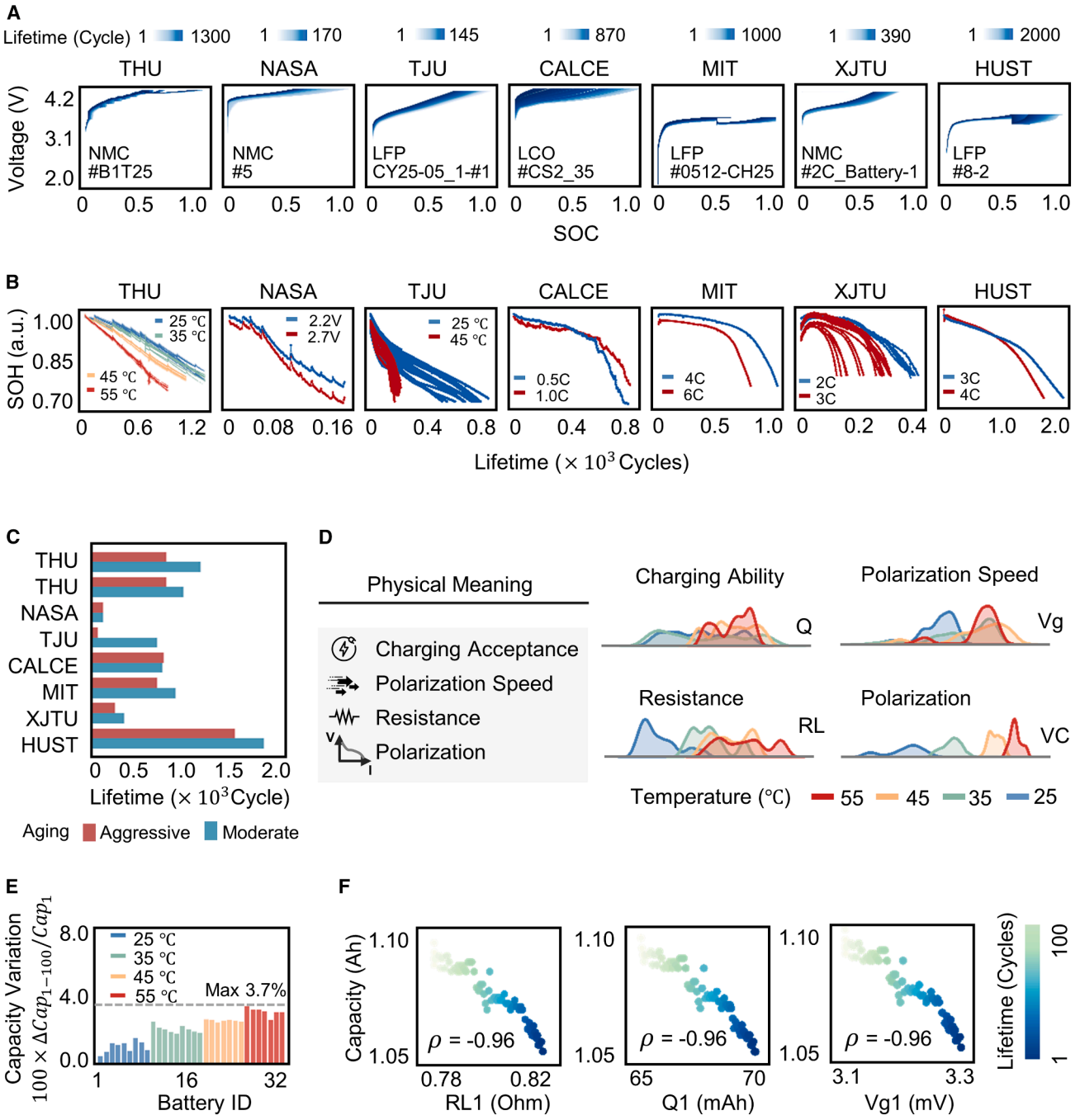


Figure 2. Visualization of datasets and features

(A) Charging voltage and state of charge profiles across seven datasets.

(B) State-of-health (SOH) degradation trends under different operating conditions in the seven datasets. Since SOH is a relative measure, representing the ratio of remaining capacity to initial capacity, it has no specific physical units. Therefore, arbitrary units (a.u.) are used to denote SOH, emphasizing that the value is a unitless relative quantity.

(C) Classification of source and target domains based on aggressive and moderate degradation, with battery lifespans for selected batteries in each domain. The THU dataset includes two domain adaptation scenarios: THU 1 (source domain at 55°C, target domain at 25°C) and THU 2 (source domain at 55°C, target domain at 35°C).

(D) Taxonomy of extracted features, categorized into four groups based on physical meaning, and their divergence across different temperature domains in our data.

(E) Capacity reduction from the 1st to the 100th cycle for 32 batteries at various temperatures in our data.

(F) Relationship between stage 1 features and capacity in our data during the first 100 cycles (taking the first battery at 25°C in the target domain as an example).

time. This helps determine whether the features can adapt to varying conditions, enabling robust performance across environments, such as temperatures, charging rates, discharging rates, etc. We select the THU dataset to showcase our approach (the results in the other 6 datasets are presented in later sections), including nine charging stages featuring high and low current conditions and their switching. The dataset's information richness allows for a proper identification of TC and PC based on the physical features that can be generalized to the other 6 widely studied datasets.

Figure 3A illustrates that for different feature groups, both the TC and PC values remain consistently high and stable during the first 80 cycles, indicating strong PCs and TCs in the early stages of degradation. This trend is consistent across target domains at three different temperatures, with detailed variation trends shown in Figure S6. However, as battery degradation progresses, the differences between the source and target domains become more pronounced, leading to a gradual decline in TC values in later cycles. After the 80th cycle, PC decreases sharply before increasing again, possibly reflecting a transition from early degradation mechanisms to an activation phase, followed by a stabilized degradation until failure.[39] Figure 3B uses variations in the amount of available data from the target domain to fine-tune the source model, and all features are used to predict the capacity in the target domain. It shows that as the amount of available data increases, there is a critical point at which the estimation error in the 25°C and 35°C target domains drops significantly. Beyond this point, adding more data does not significantly improve estimation accuracy, suggesting that an optimal data volume for early-stage estimations exists. Remarkably, the 80-cycle threshold used to achieve optimal estimation performance corresponds to the period where both the PC and TC values remain high, as shown in Figure 3A, indicating that data collected beyond this threshold become redundant.

This finding further suggests that PC and TC can guide the identification of a DS point. To explore the potential physical meanings of PC and TC, we present the distribution of these metrics across all cycles for each feature in Figure 3C (with 55°C as the source domain and 25°C as the target domain, while the distributions for 35°C and 45°C target domains are provided in Figures S7 and S8). A clear pattern emerges: features from stages 1 and 9, which involve low current, show higher PC and lower TC, while features from stages 2–8, associated with higher current, exhibit the opposite trend. This suggests a relationship between TC, PC, and features in small or large current stages. To make it clearer, we further explore the distribution of PC and TC across different feature groups in Figures 3D–3G. Stages 1 and 9 are defined as thermodynamic processes and stages 2–8 as kinetic processes based on the differing current intensities.[39,46] Thermodynamics represents the degradation of battery materials, which leads to irreversible capacity loss, and is associated with PC. Kinetics, on the other hand, represents the battery operation and is closely related to TC. The average values of PC and TC are calculated for these two degradation patterns. Thermodynamic features consistently demonstrate higher PC values but lower TC values, indicating they are more effective for direct SOH estimation. In contrast, kinetic features show higher TC values but lower PC values, suggesting they are suited for transferring models across varying operation conditions. This finding highlights that thermodynamic features may be more reliable for accurate capacity estimation given fixed operation conditions, while kinetic features play a key role in ensuring model robustness and adaptability across different operating scenarios.

Battery SOH estimation

Our dataset includes nine charging stages, resulting in a large feature set of 42 variables. To optimize the model, we employ PC and TC for feature selection, aiming to minimize the feature set while preserving predictive accuracy. First, we calculate the estimation accuracy for each feature by individually inputting it into the transfer model and predicting the target domain. Features with an estimation accuracy ranked in the top 50% during the initial one-third of their lifespan are selected (see Figure S8). These high-accuracy features demonstrate exceptional performance under varying operational conditions and should be prioritized. Assuming that we have not initially recognized these superior features, we plot the distribution of all features in terms of PC and TC, marking the selected features on the graph. Figure 4A shows that, under the 25°C target domain, these selected features are concentrated in the top right corner, indicating high estimation and transferability performance.[47]

Interestingly, features that perform well when trained and predicted individually are primarily located in the top right corner of the PC-TC plot, which corresponds to higher PC and TC values. Next, we train and fine-tune the model using only the selected features. Figure 4B compares the performance of various methods, including standard machine learning models (support vector regression [SVR] and Gaussian process regression [GPR]), non-transfer long short-term memory (LSTM) models, and our model. The results demonstrate that fine-tuning provides superior performance for predicting the capacity of new batteries compared to non-fine-tuned models, as well as outperforming traditional machine learning methods. Additionally, fine-tuning is particularly well suited for models with neural network architectures, whereas conventional machine learning methods are unable to perform fine-tuning. Consequently, other machine learning models are trained on source domain data and used to predict battery capacity beyond the available data in the target domain. SVR and GPR produce fixed initial estimations that remain stable over time. In contrast, our approach consistently achieves the lowest estimation error across varying data volumes. Notably, the estimation errors significantly decrease by the 80th cycle and remain nearly constant with further cycling. This finding not only validates the effectiveness of our method but also underscores the 80th cycle as the DS point. Furthermore, the model with selected features still exhibits a DS point, consistent with the findings shown in Figure 3B regarding the DS point without feature selection, though the degree of estimation accuracy varies.

Similarly, Figures 4C and 4D depict the results for the 35°C target domain, where features with high predictive accuracy are also clustered in the top right corner of the PC-TC plot. Once again, the estimation error does not significantly decrease after the 80th cycle, reinforcing the conclusion that 80 cycles represent the threshold for sufficient data volume. Given the

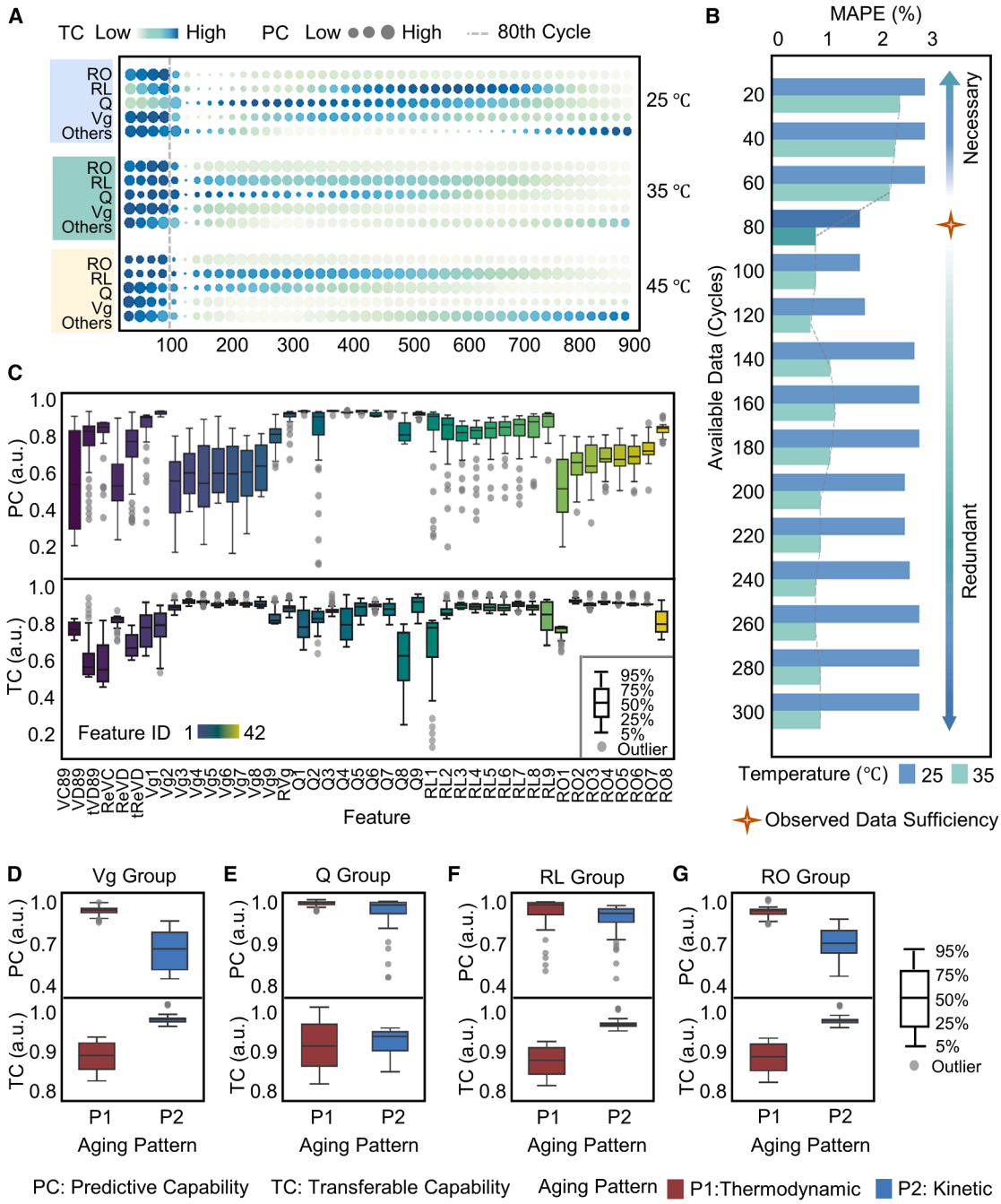


Figure 3. Predictive capability and transferable capability analysis across cycles and features

(A) Predictive capacity (PC) and transferable capability (TC) trends over degradation cycles for different feature categories, with color intensity representing TC values and the size of the bubbles representing PC values.

(B) MAPE of capacity estimations as a function of available cycles, where MAPE is computed based on the predicted capacity beyond the available cycles. Results show that 80 cycles are sufficient in the target domains (25°C and 35°C).

(C) Distribution of PC and TC across 42 features.

(D–G) Distribution of PC and TC for low-current charging features (stages 1 and 9) and high-current charging features (stages 2–8) across feature groups. For TC, Vg features averaged 0.85 ± 0.06 (P1) and 0.95 ± 0.02 (P2); Q features averaged 0.89 ± 0.06 (P1) and 0.90 ± 0.04 (P2); RL features averaged 0.80 ± 0.13 (P1) and 0.93 ± 0.02 (P2); RO features averaged 0.85 ± 0.06 (P1) and 0.93 ± 0.02 (P2); and polarization features averaged 0.72 ± 0.09 (P1) and 0.84 ± 0.03 (P2). For PC, Vg features averaged 0.92 ± 0.04 (P1) and 0.63 ± 0.15 (P2); Q features averaged 0.98 ± 0.008 (P1) and 0.95 ± 0.05 (P2); RL features averaged 0.91 ± 0.13 (P1) and 0.86 ± 0.12 (P2); RO features averaged 0.89 ± 0.04 (P1) and 0.70 ± 0.10 (P2); and polarization features averaged 0.83 ± 0.16 (P1) and 0.59 ± 0.23 (P2).

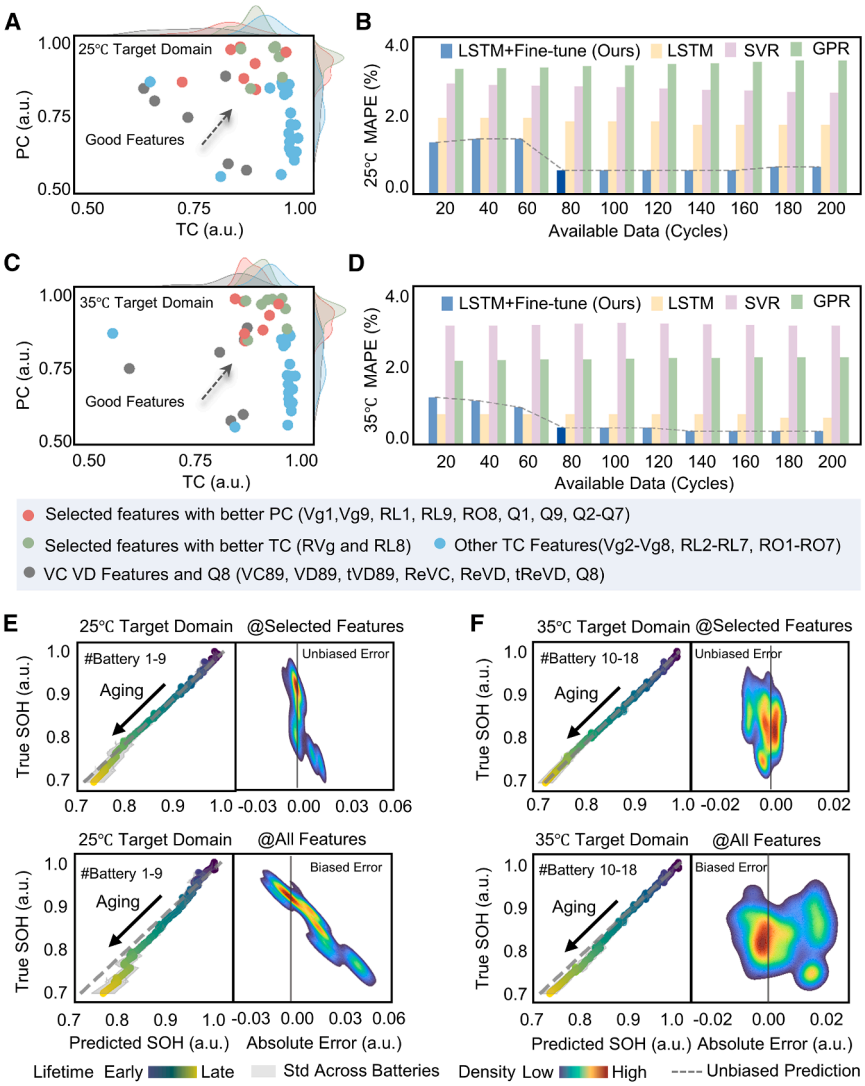


Figure 4. Model performance with feature selection in the THU dataset

(A) Feature distribution in the 25°C target domain, with transferable capability on the *x* axis and predictive capability on the *y* axis, showing high-accuracy features concentrated in the top right corner.

(B) Comparison of different methods: MAPE of estimations in the 25°C target domain (55°C is the source domain) using features selected from (A) across different available data amounts.

(C) Feature distribution in the 35°C target domain, with high-accuracy features concentrated in the top right corner.

(D) Comparison of different methods: MAPE of estimations in the 35°C target domain (55°C is the source domain) using features selected from (C) across different available data amounts.

(E) Parity plot and error distribution for SOH estimation using selected features versus all features in the 25°C target domain.

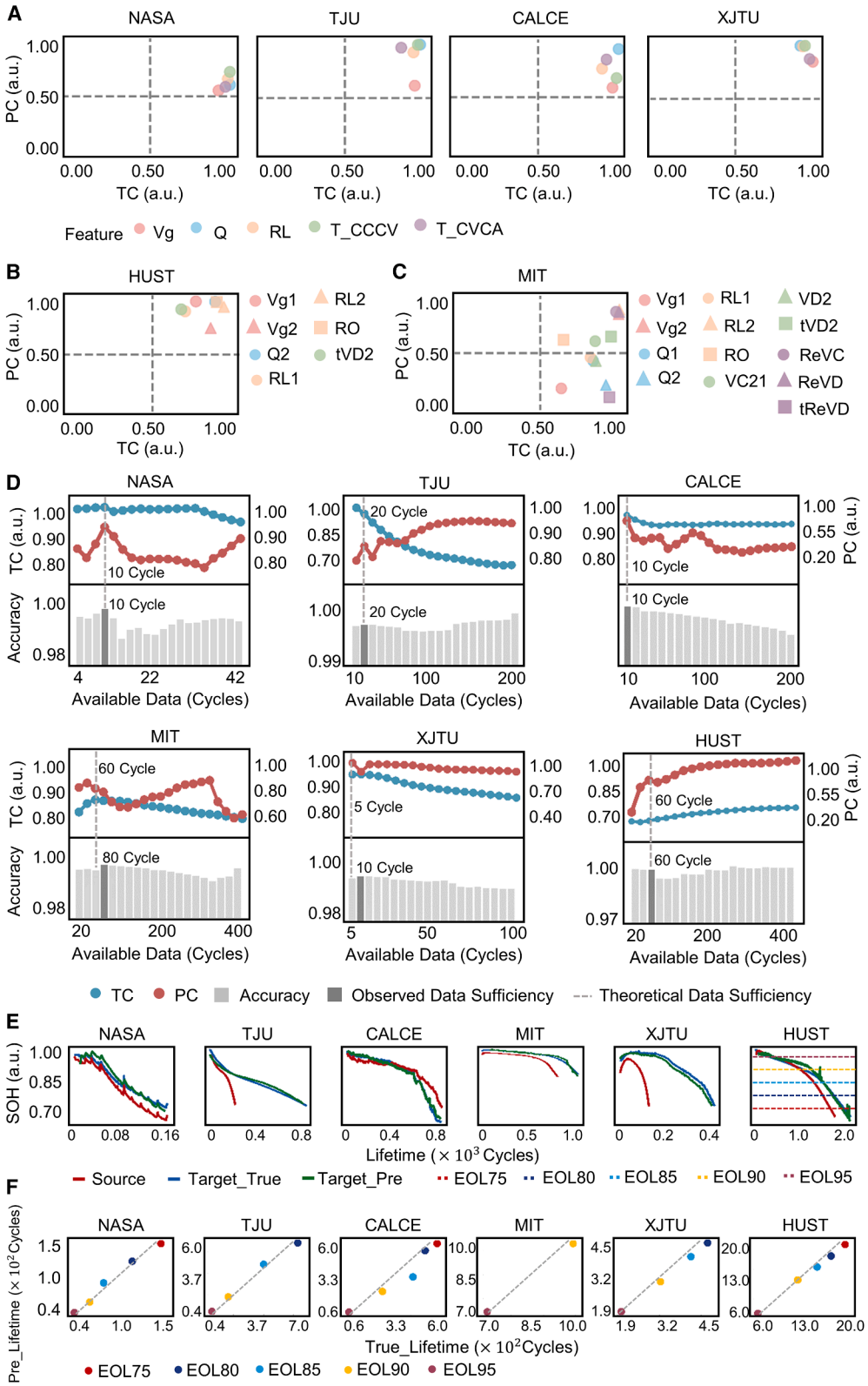
(F) Parity plot and error distribution for SOH estimation using selected features versus all features in the 35°C target domain.

large number of features, it is essential to evaluate feature performance to reduce redundancy and decrease model complexity. We compare the predicted capacities and actual capacities, as well as the error distributions, using the selected features versus all features from the former 80 available data. As shown in Figure 4E, estimations for all nine batteries in the 25°C target domain using selected features are more closely aligned along the diagonal, indicating an accurate estimation based on actual capacities, with absolute error distributions centered around zero, suggesting a minimal error bias. In contrast, estimations using all features tend to deviate from actual capacities, with most absolute errors above zero, indicating bias. Figure 4F presents similar trends for capacity parity and error distribution plots in the 35°C target domain.

Empirical DS verification across datasets

To assess the generalizability of the proposed method, the performance is validated across six additional datasets besides our dataset. The seven datasets collectively comprise 310 batteries and over 300,000 cycles, encompassing diverse conditions, including 6 distinct materials, 7 operational scenarios, and 7 transfer scenarios, as summarized in Table S1. These datasets include well-known sources, such as the NASA, CALCE, TJU, XJTU, HUST, and MIT datasets. To ensure consistency in feature extraction, the interpretations of features from these datasets are harmonized with those extracted from the THU dataset by aligning physical meanings. Features are denoised and grouped into four categories: Vg (polarization speed), Q (charging capacity), RL (resistance), and polarization group (electrochemical, ohmic, and concentration polarization). Detailed features are shown in Table S4.

To identify the optimal DS point, the temporal evolutions of PC and TC over battery cycles are analyzed. The TDS is defined as the cycle corresponding to the maximum linear combination of the mean PC and mean TC values across all features. This initial estimate is further recalculated through feature selection based on feature distributions. The three-step TDS identification process incorporating feature selection is as follows. First, using all features, the preliminary TDS is identified based on the cycle exhibiting the highest combined average PC and TC values. The initial cycle indicates a high overall feature performance but lacks detailed insight into individual feature contributions. Second, at the initial TDS, feature distributions are plotted for each dataset, with TC values on the *x* axis and PC values on the *y* axis. To simplify the model, features in the top right quadrant of the plot, where PC > 0.5 and TC > 0.5, are selected. Finally, using the selected features, the average TC and average PC



(legend on next page)

values are recalculated to identify the final TDS, following a similar approach to that in the first step.

Figures 5A–5C correspond to the first and second steps of the TDS identification process. As shown in Figure 5A, under constant current-constant voltage (CCCV) conditions, five key features (including two from the polarization group) are consistently identified across datasets. These features are clustered in the top right quadrant of the PC-TC plot, highlighting the effectiveness of the interpretable feature engineering framework. Figures 5B and 5C illustrate that for multi-stage charging and discharging profiles, additional stage-specific features can be identified when aligned with their physical interpretations. These features also predominantly cluster in the top right quadrant, further supporting the trend that features selected at the initial TDS exhibit high average PC and TC values. Figure 5D corresponds to the third step of the TDS identification process. From the top right quadrant of the PC-TC plot, critical features for each dataset are identified, leading to a recalculated TDS. The recalculated TDS represents a validation and adjustment of the initial TDS derived from all features, further optimizing the DS point. Beyond the TDS, the ODS is defined as the cycle where estimation accuracy ceases to improve, providing an intuitive explanation of the trade-off between estimation accuracy and testing duration. A more intuitive definition of ODS is the point where the derivative of accuracy is zero (see the methods).

As shown in Figure 5D, the alignment between TDS and ODS across datasets supports the hypothesis that features with high PC and TC values enable accurate and robust estimations. Both the PC-TC-defined TDS and the accuracy-defined ODS effectively identify DS points. The observed trends in PC and TC values further validate our hypothesis. PC values initially increase, decline, and then rise again, while TC values typically show a high initial stage followed by a gradual decline. These trends are consistent across datasets and align with earlier observations in the THU dataset. For the NASA, TJU, CALCE, and HUST datasets, the TDS and ODS correspond perfectly, while for the MIT and XJTU datasets, the discrepancies are minimal, only 1.5% and 1.3%, respectively. Importantly, these sufficiency points are observed early in the battery life cycle, typically within about the first 8% of the lifespan.

We stress that the predictive and transferable performance by the DS-guided feature selection and model reduction is consistent across different degradation mechanisms, as evidenced by different datasets. As shown in Figure 5E, the predicted SOH closely matches the actual SOH even if the source domain and target domain have distinct patterns. Furthermore, beyond SOH estimation, performing sensitivity analysis at different SOH levels is of particular interest for practical applications such as rapid residual value evaluation.[43] The predicted ends of life (EOLs) corresponding to capacity thresholds of 75%, 80%, 85%, 90%, and 95% of the initial capacity (denoted as EOL75, EOL80, EOL85, EOL90, and EOL95) are compared with the true EOL across six datasets. Monitoring the EOL is essential to ensure that the predicted SOH does not deviate from the actual SOH, even when the overall estimation error is low. The EOL is determined by first obtaining the predicted SOH and then drawing a horizontal line at the specified capacity threshold and identifying the cycle at which this line intersects the curve. This comparison enables a comprehensive assessment of the model's robustness in predicting not only the overall SOH but also the exact cycle when critical capacity losses can occur. The results, summarized in Figure 5F, demonstrate the model's high generalizability in estimating EOL across different degradation thresholds, which is often associated with different degradation mechanism dominance. The estimation errors are 12, 42, 34, 38, 24, and 32 cycles for the NASA, TJU, CALCE, MIT, XJTU, and HUST datasets, respectively.

DISCUSSION

This empirical study proposes an examination of battery DS to determine the minimum data volume required for fine-tuning to achieve accurate battery SOH estimation, a challenge for which prior work has not yet provided a generalizable solution. The ODS calculated from the estimation accuracy aligns closely with the TDS derived from PC-TC metrics. This consistency is validated across seven datasets with 6 varying materials, 7 operating conditions, and 7 transfer scenarios. It demonstrates that an average of the first 8% of battery life cycle data (THU 6.0%, NASA 5.9%, TJU 2.5%, CALCE 1.2%, MIT 8.0%, XJTU 2.5%, and HUST 3.0%), as a conservative guideline derived from our multi-dataset analysis, is sufficient to achieve anticipated estimation accuracies and cross-operation-condition transferability. Beyond this critical DS point, additional data do not necessarily improve estimation performance under transferable settings. Building on Severson et al., which utilized data from the first 100 cycles for lifetime estimations, this study demonstrates that early-stage data are sufficient to adapt source domain models to target domains, thereby enabling accurate SOH estimation.[43] This work provides a reliable foundation based on an empirical study of DS that explains why and to what extent early-life data can forecast future performance under single and across multiple operation conditions. It also underscores the practical benefits of reducing experimental costs

Figure 5. Validation of data sufficiency across datasets

(A) Feature distribution for transferable capability (TC) and predictive capability (PC) under CCCV conditions, where the physical meaning of features extracted from multiple datasets is aligned with those from our dataset.

(B) Feature distribution for PC and TC under multi-stage discharging conditions, with aligned physical meanings from our dataset.

(C) Feature distribution for PC and TC under multi-stage charging conditions, with aligned physical meaning across datasets.

(D) PC and TC trends over cycles for six datasets, with estimation accuracy at various available data volumes in the target domain, showing close agreement between theoretical and observed data sufficiency (DS) points.

(E) Battery SOH is predicted across six datasets using sufficient data. The transfer scenarios were set with a discharge cutoff voltage shift from 2.7 to 2.5 V in NASA, a temperature shift from 45°C to 25°C in TJU, a discharging rate shift from 1 to 0.5 C in CALCE, a charging rate transition from 6 to 4 C in MIT, a charging rate shift from 3 to 2 C in XJTU, and a discharging rate shift from 4 to 3 C in HUST.

(F) Sensitivity analysis under five different EOL thresholds (75%, 80%, 85%, 90%, and 95%) across the six datasets.

and time without compromising estimation accuracy and transferability.

We must acknowledge that the results are based on the empirical evaluation of existing datasets, which could lead to biased conclusions due to the incomplete inclusion of extreme degradation mechanisms, one of the most determinant factors influencing the PC and TC of physical features. However, the proposed empirical study of DS is validated across 7 datasets with significantly different degradation curves, implying the inclusion of a wide range of degradation mechanisms, as depicted in Figure 2B. Existing studies also indicate that battery degradation mechanisms can be generally divided into three stages: the early formation of the solid electrolyte interphase (SEI) layer, a relatively stable internal state followed by dominant processes of loss of lithium-ion inventory (LLI) and loss of active materials (LAM), and the later stage leading to eventual failure.[48–50] The degradation stage over the whole battery life aligns with data-driven findings that changes in PC and TC occur in distinct phases, reflecting the progression of battery degradation. The trends in PC and TC closely mirror the battery degradation process. Initially, PC is high, which may correspond to the early formation and stabilization of the SEI layer. Then, there is a noticeable decrease in PC, likely due to the completion of activation or SEI growth. During this first stage, TC remains relatively high because there is little differentiation between batteries in the early stages of their lifespan. In the second stage, PC gradually increases, and TC gradually declines, likely reflecting the phase dominated by loss of LLI and LAM. In the final stage, PC experiences a sharp decline, and TC continues to decrease, which may correspond to the battery failure phase. Although degradation mechanisms remain unexplored, the trends observed in our PC and TC metrics can help explain currently recognized degradation processes. The correspondence between the PC and TC values in Figures 3D–3G and the features of the small and large current phases suggests that PC and TC not only have mathematical significance but also potential physical relevance. This can aid in reflecting the internal degradation mechanisms of batteries, particularly those related to thermodynamic and kinetic degradation processes.

As large models continue to evolve, the sufficiency of training data becomes increasingly critical. While larger datasets may require more computational resources, appropriately selected and sufficient data can reduce dependence on high computational power. In battery management, large-scale predictive models rely on high-quality, sufficient data to ensure accurate estimations and effective learning. The complexity of these models and their computational demands highlight the importance of optimizing training data. Ensuring DS, particularly during the early stages of battery degradation, not only reduces computational costs but also improves model accuracy. Striking the right balance between dataset volume and estimation precision is essential as model complexity and computational demands increase. By focusing on early-life data, it is possible to achieve estimation accuracy and transferability while minimizing both data requirements and computational costs, thus making large model training more feasible and cost effective.

Our empirical study opens several pathways for future battery research exploration. One key area is expanding the applicability of PC and TC metrics to a broader range of battery chemistries, operational conditions, and degradation mechanisms. We aim to further uncover deeper insights into the degradation mechanisms of batteries through the analysis of PC and TC when extensive electrochemistry-level data testing is accessible. By exploring their contributions to decoupling thermodynamic and kinetic losses, we hope to validate our findings from a chemical and materials science perspective. Interpretable features not only aid in understanding degradation mechanisms but also provide a foundation for developing more robust estimation models that can handle diverse and unforeseen operational conditions. Based on PC and TC, future work could extend TDS by integrating adaptive sampling strategies and dynamic recalibration mechanisms, enabling robust estimation performance in real-world battery management systems. Additionally, we attempt to apply our method across different battery chemistries to investigate whether similar patterns in PC and TC emerge under more extreme transfer settings. Given that the method is empirical in nature, a larger and more diverse set of data would be required to fully support claims of a universal principle of DS. Nevertheless, our study has been validated on seven datasets with diverse operating conditions, which represents a relatively rich collection compared with previous studies (see Table S6). In future work, we aim to extend validation to even broader operation conditions, such as varying user preferences and historical usage patterns.

METHODS

PC within one domain

The correlation between features and capacity reflects the degree to which a feature can accurately predict capacity, with stronger correlations indicating a higher potential for reliable capacity representation. Thus, we define the PC using the Pearson correlation coefficient. Given that battery degradation is a time-dependent process, it is critical to account for temporal dynamics when assessing PC rather than relying on static evaluations. Therefore, we extend PC to incorporate its evolution over cycles, defining the time-dependent PC_i^j as

$$PC_i^j = \frac{\sum_{j=1}^{m}\left(F_i^j - \overline{F}_i\right)\left(Cap^j - \overline{Cap}\right)}{\sqrt{\sum_{j=1}^{m}\left(F_i^j - \overline{F}_i\right)^2 \sum_{j=1}^{m}\left(Cap^j - \overline{Cap}\right)^2}}, \quad \text{(Equation 1)}$$

$$\overline{F}_i = \frac{1}{m}\sum_{j=1}^{m} F_i^j, \quad \text{(Equation 2)}$$

$$\overline{Cap} = \frac{1}{m}\sum_{j=1}^{m} Cap^j, \quad \text{(Equation 3)}$$

where PC_i^j represents the PC of the i-th feature up to cycle j, F_i^j is the value of the i-th feature at the j-th cycle, Cap^j is the corresponding capacity at the j-th cycle, m is the number of available cycles, $\overline{F}_i$ is the average value of the i-th feature over m cycles, and $\overline{Cap}$ is the average capacity over m cycles.

TC across domains

Battery capacity degradation exhibits inconsistent degradation patterns under varying operational conditions. Similarly, feature distributions can vary across domains due to differences in these conditions. To quantify the distributional divergence of features across domains, we employ the Wasserstein distance (*WD*), which measures the discrepancy between feature distributions.[47] A smaller *WD* indicates better alignment of feature distributions across domains, suggesting that features are less influenced by dynamic conditions and thus possess higher transferability. Accordingly, we define TC using $1 - WD$. As we did with PC, we extend TC to incorporate its evolution over cycles, defining the time-dependent TC as TC_i^j:

$$TC_i^j = 1 - WD_i^j, \quad \text{(Equation 4)}$$

$$WD_i^j = \left(\inf_{\gamma \in \Gamma\left(P_i^j, Q_i^j\right)} \int_{\Omega \times \Omega} \left| X_{S,i}^{\ j} - X_{T,i}^{\ j} \right|^q d\gamma\left(X_{S,i}^{\ j}, X_{T,i}^{\ j}\right) \right)^{1/q}, \quad \text{(Equation 5)}$$

where TC_i^j represents the TC of the i-th feature up to cycle j, W_i quantifies the feature distribution difference between source and target domains, and WD_i^j is the *WD* of the i-th feature up to cycle j, quantifying the distributional discrepancy between the source and target domains for this feature over the first j cycles. P_i^j and Q_i^j are probability distributions of the i-th feature in the source and target domains, respectively, calculated based on the data from the first j cycles. Ω is the domain of feature values. γ is a joint distribution over $\Omega \times \Omega$ with marginal distributions P_i^j and Q_i^j. $X_{S,i}^{\ j}$ and $X_{T,i}^{\ j}$ are the values of the i-th feature in the source and target domains, respectively, over the first j cycles. q is the order of the *WD*, commonly $q = 1$ or $q = 2$. Here, $q = 1$ is chosen to avoid the effect of measurement noises.

LSTM with fine-tuning

To model the sequential degradation behavior of battery capacity, we utilize an LSTM network, which captures the temporal dependencies in the battery data. Specifically, the current capacity of a battery is influenced by its preceding feature states. Therefore, the features from the preceding three cycles are used as input to predict the current capacity. The input feature sequence for the source domain battery is defined as

$$X_{input}^{source} = \left[x_{t-3}, x_{t-2}, x_{t-1}\right], \quad \text{(Equation 6)}$$

where $x_{t-3}, x_{t-2},$ and x_{t-1} represent the features of the battery at three consecutive previous time steps.

The time step is set to be three for the elbow rule (Figure S9). Using the LSTM model trained on the source domain, the predicted capacity of the battery is obtained as $\hat{y}_t^{source}$.

The LSTM model architecture consists of sequentially stacked layers, with the first LSTM layer comprising 96 units. This layer accepts input data with a shape of (T, F), where T denotes the number of time steps and F represents the number of features. The second LSTM layer contains 64 units and outputs only the final time step of the sequence. The outputs from the LSTM layers are subsequently processed through two fully connected (Dense) layers. The first Dense layer consists of 32 units with weights initialized using the Glorot uniform initializer. The final Dense layer contains a single unit, providing the regression output of the model, also initialized with Glorot uniform. Weight initialization is adopted to improve convergence stability in recurrent networks. The model is trained using the Adam optimizer, and the mean squared error is employed as the loss function.

To adapt the pre-trained model to the target domain, fine-tuning is performed using data from the target battery. The input feature sequence for the target domain battery is defined as

$$X_{input}^{target} = \left[x'_{t-3}, x'_{t-2}, x'_{t-1}\right], \quad \text{(Equation 7)}$$

where $x'_{t-3}, x'_{t-2},$ and x'_{t-1} denote the corresponding features in the target domain.

In the fine-tuning stage, the model weights within the first two LSTM layers are frozen, ensuring that the learned temporal representations remain fixed. Assuming the first d cycles of a new battery are available for testing, the model weights within the two Dense layers are fine-tuned using features and actual discharge capacities from the previous d cycles of new batteries. The remaining actual capacities are then used to validate the performance of the fine-tuned model. The adapted model predicts the capacity for the target domain battery as $\hat{y}_t^{target}$. The optimization is performed using the Adam optimizer with a batch size of 64, over 20 epochs. This fine-tuning process ensures that the model leverages the knowledge learned from the source domain while adapting to the specific characteristics of the target domain, improving the transferability of estimations under varying operational scenarios.

Evaluation metric

The MAPE is defined as

$$\text{MAPE} = \frac{\sum_{j=d}^{C} \left| y_j - \hat{y}_j \right|}{\sum_{j=d}^{C} y_j} \times 100\%, \quad \text{(Equation 8)}$$

where y_j and $\hat{y}_j$ are the true capacity and predicted capacity in the j-th cycle, respectively. C is the length of battery cycles, and d is the starting estimation point, which means the available cycles.

ODS

DS refers to the minimum amount of data required to achieve a high estimation accuracy, where further data curation does not lead to a significant improvement in accuracy or reduction in estimation error. To quantify this, we propose the concept of ODS, defined based on the MAPE. The accuracy is defined as

$$Acc = 1 - \text{MAPE}. \quad \text{(Equation 9)}$$

To identify the contribution of each additional data interval, the change in accuracy, $\Delta Acc\ (j)$, between successive intervals is calculated as

$$\Delta Acc\ (j) = Acc\ (j) - Acc\ (j-1). \quad \text{(Equation 10)}$$

The ODS is determined as the cycle index j, where the accuracy continues to increase but subsequently stops improving, defined as

$$\mathrm{ODS} = arg\ \max_j [\Delta Acc\ (j) > 0, \Delta Acc\ (j+1) < 0], \quad \text{(Equation 11)}$$

where ODS is the observed DS, representing the cycle at which the estimation accuracy reaches its peak without further improvement. It identifies the point at which additional data no longer significantly enhance the model's predictive performance, ensuring efficient use of available data while minimizing redundancy.

TDS

PC and TC quantify a feature's ability to predict capacity accurately and maintain performance under varying operating conditions in transfer learning scenarios. The combination of PC and TC forms the foundation for defining TDS. We first calculate the average TC_i^j and average PC_i^j across all features:

$$\mathrm{TC}(j) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{TC}_i^j \quad \text{(Equation 12)}$$

$$\mathrm{PC}(j) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{PC}_i^j. \quad \text{(Equation 13)}$$

TC_i^j is the TC of the i-th feature before cycle j. PC_i^j is the PC of the i-th feature before cycle j. n is the total number of features.

Both $\mathrm{TC}(j)$ and $\mathrm{PC}(j)$ are normalized to the range $[0,1]$ to ensure comparability:

$$\mathrm{TC}_{norm}(j) = \frac{\mathrm{TC}(j)}{\mathrm{TC}(j)_{max}} \quad \text{(Equation 14)}$$

$$\mathrm{PC}_{norm}(j) = \frac{\mathrm{PC}(j)}{\mathrm{PC}(j)_{max}}. \quad \text{(Equation 15)}$$

To account for the decline in $\mathrm{TC}_{norm}(j)$ and $\mathrm{PC}_{norm}(j)$ across consecutive cycles, the following differences are calculated:

$$\Delta\mathrm{TC}_{norm}(j) = \mathrm{TC}_{norm}(j) - \mathrm{TC}_{norm}(j+1) \quad \text{(Equation 16)}$$

$$\Delta\mathrm{PC}_{norm}(j) = \mathrm{PC}_{norm}(j) - \mathrm{PC}_{norm}(j+1). \quad \text{(Equation 17)}$$

The combined metric integrates normalized TC and PC values while emphasizing their decline to highlight the current cycle's comprehensive value:

$$f\ (\mathrm{PC}, \mathrm{TC}) = \mathrm{TC}_{norm}(j) + \mathrm{PC}_{norm}(j) + \Delta\mathrm{TC}_{norm}(j) + \Delta\mathrm{PC}_{norm}(j). \quad \text{(Equation 18)}$$

TDS is defined as the cycle j where $f\ (\mathrm{PC}, \mathrm{TC})$ reaches its maximum:

$$\mathrm{TDS} = arg\ \max_j [f\ (\mathrm{PC}, \mathrm{TC})]. \quad \text{(Equation 19)}$$

In this way, TDS indicates the cycle at which the TC and PC are both high, marking the theoretical point of sufficient data.

SVR

An SVR model is implemented to predict the SOH of the target battery using the same source domain features as input. Prior to model training, all input features and capacity values are normalized to the range $[0,1]$ using a MinMaxScaler, ensuring consistency between the source and target domains. The SVR model employs a radial basis function (RBF) kernel with a regularization parameter of $C = 10$, an epsilon-tube margin of $\varepsilon = 0.1$, and a kernel coefficient of $\gamma = 0.1$. These hyperparameters are selected empirically to balance the model's bias and variance and to ensure stable convergence across cross-domain scenarios. The model is trained on the source battery data and tested on the target battery. After estimation, the normalized capacity outputs are inverse transformed back to the original scale to facilitate direct comparison with the actual capacity values. The model implementation is conducted using the SVR class from the scikit-learn library with standard settings unless otherwise specified.

GPR

A GPR model is developed to predict the SOH of the target domain battery based on the training data from the source domain battery. Prior to model training, all input features and capacity values are normalized to the range $[0,1]$ using a MinMaxScaler, ensuring scale consistency between the source and target domains. The GPR model employs a composite kernel function, defined as the product of a constant kernel $C(1.0)$, with bounds set to $(1 \times 10^{-3}, 1 \times 10^{3})$, and an RBF kernel with an initial length scale of 1.0 and bounds of $(1 \times 10^{-2}, 1 \times 10^{2})$. To mitigate the risk of convergence to local minima during hyperparameter optimization, the model uses ten restarts of the optimizer. An additional noise term ($\alpha = 1 \times 10^{-4}$) is incorporated to account for measurement noise and inherent variability in the data. After estimation, the normalized capacity outputs are inverse transformed to their original scale for evaluation. All model implementation and training procedures are performed using the GaussianProcessRegressor class from the scikit-learn library with default settings unless otherwise specified.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Shengyu Tao (sytao@berkeley.edu).

Materials availability

No materials were used in this work.

Data and code availability

The THU, NASA, TJU, CALCE, MIT, XJTU, and HUST datasets are open source and available in Tao et al.,[39] Xu et al.,[40] Zhu et al.,[41] Xing et al.,[42] Severson et al.,[43] Wang et al.,[44] and Ma et al.[45] The supporting data are provided in the supplemental information. The modeling code and processed data have been deposited at GitHub at https://github.com/terencetaothucb/data-sufficiency-examination-of-transferable-lithium-ion-battery-state-of-health-estimation. The associated DOI is https://doi.org/10.5281/zenodo.17051963 at https://zenodo.org/records/17051963.

ACKNOWLEDGMENTS

This research work was supported by the Key Scientific Research Support Project of the Shanxi Energy Internet Research Institute (no. SXEI2023A002) (X.Z.), the Meituan Scholar Program-International Collaboration Project (no. 202209A) (X.Z.), the Tsinghua Shenzhen International Graduate School-Shenzhen Pengrui Young Faculty Program of Shenzhen Pengrui Foundation (no. SZPR2023007) (G.Z.), and the Guangdong Basic and Applied Basic Research Foundation (no. 2023B1515120099) (G.Z.).

AUTHOR CONTRIBUTIONS

L.S. and S.T. conceptualized, designed, reviewed, revised, and prepared the manuscript draft. W.L. and D.U.S. supervised, discussed, reviewed, revised, and provided practical suggestions and an industry-focused view for this work. G.Z. and X.Z. supervised, reviewed, discussed, and acquired funding for this work. All authors revised the manuscript during revision.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrp.2025.102901.

Received: August 6, 2025
Revised: August 24, 2025
Accepted: September 16, 2025
Published: October 7, 2025

REFERENCES

1. Xu, J., Cai, X., Cai, S., Shao, Y., Hu, C., Lu, S., and Ding, S. (2023). High-Energy Lithium-Ion Batteries: Recent Progress and a Promising Future in Applications. Energy &. Environ. Materials *6*, e12450. https://doi.org/10.1002/eem2.12450.
2. Khan, F.M.N.U., Rasul, M.G., Sayem, A.S.M., and Mandal, N.K. (2023). Design and optimization of lithium-ion battery as an efficient energy storage device for electric vehicles: A comprehensive review. J. Energy Storage *71*, 108033. https://doi.org/10.1016/j.est.2023.108033.
3. Jiang, B., Gent, W.E., Mohr, F., Das, S., Berliner, M.D., Forsuelo, M., Zhao, H., Attia, P.M., Grover, A., Herring, P.K., et al. (2021). Bayesian learning for rapid prediction of lithium-ion battery-cycling protocols. Joule *5*, 3187–3203. https://doi.org/10.1016/j.joule.2021.10.010.
4. Zhao, J., Qu, X., Li, Y., Nan, J., and Burke, A.F. (2025). Real-time prediction of battery remaining useful life using hybrid-fusion deep neural networks. Energy *328*, 136618. https://doi.org/10.1016/j.energy.2025.136618.
5. Huang, X., Tao, S., Liang, C., Chen, J., Shi, J., Li, Y., Xia, B., Zhou, G., and Zhang, X. (2025). Physics-informed mixture of experts network for interpretable battery degradation trajectory computation amid second-life complexities. Preprint at arXiv. https://doi.org/10.48550/arXiv.2506.17755.
6. Zhao, J., Ling, H., Liu, J., Wang, J., Burke, A.F., and Lian, Y. (2023). Machine learning for predicting battery capacity for electric vehicles. eTransportation *15*, 100214. https://doi.org/10.1016/j.etran.2022.100214.
7. Liu, K., Shang, Y., Ouyang, Q., and Widanage, W.D. (2021). A Data-Driven Approach With Uncertainty Quantification for Predicting Future Capacities and Remaining Useful Life of Lithium-ion Battery. IEEE Trans. Ind. Electron. *68*, 3170–3180. https://doi.org/10.1109/TIE.2020.2973876.
8. Xu, L., Deng, Z., Xie, Y., Lin, X., and Hu, X. (2023). A Novel Hybrid Physics-Based and Data-Driven Approach for Degradation Trajectory Prediction in Li-Ion Batteries. IEEE Trans. Transp. Electrific. *9*, 2628–2644. https://doi.org/10.1109/TTE.2022.3212024.
9. Che, Y., Zheng, Y., Wu, Y., Sui, X., Bharadwaj, P., Stroe, D.-I., Yang, Y., Hu, X., and Teodorescu, R. (2022). Data efficient health prognostic for batteries based on sequential information-driven probabilistic neural network. Appl. Energy *323*, 119663. https://doi.org/10.1016/j.apenergy.2022.119663.
10. Zhang, Y., and Li, Y.-F. (2022). Prognostics and health management of Lithium-ion battery using deep learning methods: A review. Renew. Sustain. Energy Rev. *161*, 112282. https://doi.org/10.1016/j.rser.2022.112282.
11. Zhang, H., Li, Y., Zheng, S., Lu, Z., Gui, X., Xu, W., and Bian, J. (2025). Battery lifetime prediction across diverse ageing conditions with inter-cell deep learning. Nat. Mach. Intell. *7*, 270–277. https://doi.org/10.1038/s42256-024-00972-x.
12. Huang, X., Tao, S., Liang, C., Ma, R., Wang, X., Xia, B., and Zhang, X. (2025). Robust and generalizable lithium-ion battery health estimation using multi-scale field data decomposition and fusion. J. Power Sources *642*, 236939. https://doi.org/10.1016/j.jpowsour.2025.236939.
13. Weng, C., Cui, Y., Sun, J., and Peng, H. (2013). On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. J. Power Sources *235*, 36–44. https://doi.org/10.1016/j.jpowsour.2013.02.012.
14. Zhang, Z., Li, L., Li, X., Hu, Y., Huang, K., Xue, B., Wang, Y., and Yu, Y. (2022). State-of-health estimation for the lithium-ion battery based on gradient boosting decision tree with autonomous selection of excellent features. Int. J. Energy Res. *46*, 1756–1765. https://doi.org/10.1002/er.7292.
15. Hu, C., Jain, G., Schmidt, C., Strief, C., and Sullivan, M. (2015). Online estimation of lithium-ion battery capacity using sparse Bayesian learning. J. Power Sources *289*, 105–113. https://doi.org/10.1016/j.jpowsour.2015.04.166.
16. Li, Y., Zou, C., Berecibar, M., Nanini-Maury, E., Chan, J.C.W., van den Bossche, P., Van Mierlo, J., and Omar, N. (2018). Random forest regression for online capacity estimation of lithium-ion batteries. Appl. Energy *232*, 197–210. https://doi.org/10.1016/j.apenergy.2018.09.182.
17. Liu, K., Hu, X., Wei, Z., Li, Y., and Jiang, Y. (2019). Modified Gaussian Process Regression Models for Cyclic Capacity Prediction of Lithium-Ion Batteries. IEEE Trans. Transp. Electrific. *5*, 1225–1236. https://doi.org/10.1109/TTE.2019.2944802.
18. He, N., Wang, Q., Lu, Z., Chai, Y., and Yang, F. (2024). Early prediction of battery lifetime based on graphical features and convolutional neural networks. Appl. Energy *353*, 122048. https://doi.org/10.1016/j.apenergy.2023.122048.
19. You, G.W., Park, S., and Oh, D. (2017). Diagnosis of Electric Vehicle Batteries Using Recurrent Neural Networks. IEEE Trans. Ind. Electron. *64*, 4885–4893. https://doi.org/10.1109/TIE.2017.2674593.
20. Wang, Y., Zhu, J., Cao, L., Gopaluni, B., and Cao, Y. (2023). Long Short-Term Memory Network with Transfer Learning for Lithium-ion Battery Capacity Fade and Cycle Life Prediction. Appl. Energy *350*, 121660. https://doi.org/10.1016/j.apenergy.2023.121660.
21. Zhao, J., Li, D., Li, Y., Shi, D., Nan, J., and Burke, A.F. (2025). Battery state of health estimation under fast charging via deep transfer learning. iScience *28*, 112235. https://doi.org/10.1016/j.isci.2025.112235.
22. Tao, S., Guo, R., Lee, J., Moura, S., Casals, L.C., Jiang, S., Shi, J., Harris, S., Zhang, T., Chung, C.Y., et al. (2025). Immediate remaining capacity estimation of heterogeneous second-life lithium-ion batteries via deep generative transfer learning. Energy Environ. Sci. *18*, 7413–7426. https://doi.org/10.1039/D5EE02217G.
23. Tao, S., Ma, R., Zhao, Z., Ma, G., Su, L., Chang, H., Chen, Y., Liu, H., Liang, Z., Cao, T., et al. (2024). Generative learning assisted state-of-health estimation for sustainable battery recycling with random retirement conditions. Nat. Commun. *15*, 10154. https://doi.org/10.1038/s41467-024-54454-0.

24. Sun, J., Gu, A., and Kainz, J. (2025). A solution framework for the experimental data shortage problem of lithium-ion batteries: Generative adversarial network-based data augmentation for battery state estimation. J. Energy Chem. *103*, 476–497. https://doi.org/10.1016/j.jechem.2024.12.010.

25. Wang, Y., Feng, X., Guo, D., Hsu, H., Hou, J., Zhang, F., Xu, C., Chen, X., Wang, L., Zhang, Q., and Ouyang, M. (2024). Temperature excavation to boost machine learning battery thermochemical predictions. Joule *8*, 2639–2651. https://doi.org/10.1016/j.joule.2024.07.002.

26. Jin, H., Ru, R., Cai, L., Meng, J., Wang, B., Peng, J., and Yang, S. (2025). A synthetic data generation method and evolutionary transformer model for degradation trajectory prediction in lithium-ion batteries. Appl. Energy *377*, 124629. https://doi.org/10.1016/j.apenergy.2024.124629.

27. Park, S., Lee, H., Scott-Nevros, Z.K., Lim, D., Seo, D.-H., Choi, Y., Lim, H., and Kim, D. (2023). Deep-learning based spatio-temporal generative model on assessing state-of-health for Li-ion batteries with partially-cycled profiles. Mater. Horiz. *10*, 1274–1281. https://doi.org/10.1039/D3MH00013C.

28. Li, Y., Li, K., Liu, X., Wang, Y., and Zhang, L. (2021). Lithium-ion battery capacity estimation — A pruned convolutional neural network approach assisted with transfer learning. Appl. Energy *285*, 116410. https://doi.org/10.1016/j.apenergy.2020.116410.

29. Zhou, Z., Liu, Y., You, M., Xiong, R., and Zhou, X. (2022). Two-stage aging trajectory prediction of LFP lithium-ion battery based on transfer learning with the cycle life prediction. Green Energy and Intelligent Transportation *1*, 100008. https://doi.org/10.1016/j.geits.2022.100008.

30. Deng, Z., Lin, X., Cai, J., and Hu, X. (2022). Battery health estimation with degradation pattern recognition and transfer learning. J. Power Sources *525*, 231027. https://doi.org/10.1016/j.jpowsour.2022.231027.

31. Lin, M., You, Y., Meng, J., Wang, W., Wu, J., and Stroe, D.-I. (2023). Lithium-ion battery degradation trajectory early prediction with synthetic dataset and deep learning. J. Energy Chem. *85*, 534–546. https://doi.org/10.1016/j.jechem.2023.06.036.

32. Fei, Z., Yang, F., Tsui, K.-L., Li, L., and Zhang, Z. (2021). Early prediction of battery lifetime via a machine learning based framework. Energy *225*, 120205. https://doi.org/10.1016/j.energy.2021.120205.

33. Saxena, S., Ward, L., Kubal, J., Lu, W., Babinec, S., and Paulson, N. (2022). A convolutional neural network model for battery capacity fade curve prediction using early life data. J. Power Sources *542*, 231736. https://doi.org/10.1016/j.jpowsour.2022.231736.

34. Zhang, C., Du, J., Li, S., Zhang, L., and Zhang, W. (2024). Cross-domain machine transfer learning for capacity aging trajectory prediction of lithium-ion batteries. J. Power Sources *624*, 235534. https://doi.org/10.1016/j.jpowsour.2024.235534.

35. Fu, S., Tao, S., Fan, H., He, K., Liu, X., Tao, Y., Zuo, J., Zhang, X., Wang, Y., and Sun, Y. (2024). Data-driven capacity estimation for lithium-ion batteries with feature matching based transfer learning method. Appl. Energy *353*, 121991. https://doi.org/10.1016/j.apenergy.2023.121991.

36. Shen, S., Sadoughi, M., Li, M., Wang, Z., and Hu, C. (2020). Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. Appl. Energy *260*, 114296. https://doi.org/10.1016/j.apenergy.2019.114296.

37. Guo, Y., Wang, Y., Ding, P., and Huang, K. (2023). Future degradation trajectory prediction of lithium-ion battery based on a three-step similarity evaluation criterion for battery selection and transfer learning. J. Energy Storage *72*, 108763. https://doi.org/10.1016/j.est.2023.108763.

38. Liang, C., Tao, S., Huang, X., Wang, Y., Xia, B., and Zhang, X. (2025). Stochastic state of health estimation for lithium-ion batteries with automated feature fusion using quantum convolutional neural network. J. Energy Chem. *106*, 205–219. https://doi.org/10.1016/j.jechem.2025.02.030.

39. Tao, S., Zhang, M., Zhao, Z., Li, H., Ma, R., Che, Y., Sun, X., Su, L., Sun, C., Chen, X., et al. (2025). Non-destructive degradation pattern decoupling for early battery trajectory prediction via physics-informed learning. Energy Environ. Sci. *18*, 1544–1559. https://doi.org/10.1039/D4EE03839H.

40. Xu, H., Peng, Y., and Su, L. (2018). Health State Estimation Method of Lithium Ion Battery Based on NASA Experimental Data Set. IOP Conf. Ser. Mater. Sci. Eng. *452*, 032067. https://doi.org/10.1088/1757-899X/452/3/032067.

41. Zhu, J., Wang, Y., Huang, Y., Bhushan Gopaluni, R., Cao, Y., Heere, M., Mühlbauer, M.J., Mereacre, L., Dai, H., Liu, X., et al. (2022). Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. Nat. Commun. *13*, 2261. https://doi.org/10.1038/s41467-022-29837-w.

42. Xing, Y., Ma, E.W.M., Tsui, K.-L., and Pecht, M. (2013). An ensemble model for predicting the remaining useful performance of lithium-ion batteries. Microelectron. Reliab. *53*, 811–820. https://doi.org/10.1016/j.microrel.2012.12.003.

43. Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D., et al. (2019). Data-driven prediction of battery cycle life before capacity degradation. Nat. Energy *4*, 383–391. https://doi.org/10.1038/s41560-019-0356-8.

44. Wang, F., Zhai, Z., Zhao, Z., Di, Y., and Chen, X. (2024). Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. Nat. Commun. *15*, 4332. https://doi.org/10.1038/s41467-024-48779-z.

45. Ma, G., Xu, S., Jiang, B., Cheng, C., Yang, X., Shen, Y., Yang, T., Huang, Y., Ding, H., and Yuan, Y. (2022). Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning. Energy Environ. Sci. *15*, 4083–4094. https://doi.org/10.1039/D2EE01676A.

46. Wood, K.N., Hawley, W.B., and Less, G. (2024). Scalable and Non-Destructive Analysis of Battery Degradation and Performance. Meet. Abstr. *697*, MA2024-01. https://doi.org/10.1149/MA2024-014697mtgabs.

47. Tao, S., Sun, C., Fu, S., Wang, Y., Ma, R., Han, Z., Sun, Y., Li, Y., Wei, G., Zhang, X., et al. (2023). Battery Cross-Operation-Condition Lifetime Prediction via Interpretable Feature Engineering Assisted Adaptive Machine Learning. ACS Energy Lett. *8*, 3269–3279. https://doi.org/10.1021/acsenergylett.3c01012.

48. Han, X., Lu, L., Zheng, Y., Feng, X., Li, Z., Li, J., and Ouyang, M. (2019). A review on the key issues of the lithium ion battery degradation among the whole life cycle. eTransportation *1*, 100005. https://doi.org/10.1016/j.etran.2019.100005.

49. Liao, Y., Zhang, H., Peng, Y., Hu, Y., Liang, J., Gong, Z., Wei, Y., and Yang, Y. (2024). Electrolyte Degradation During Aging Process of Lithium-Ion Batteries: Mechanisms, Characterization, and Quantitative Analysis. Adv. Energy Mater. *14*, 2304295. https://doi.org/10.1002/aenm.202304295.

50. Zhao, M., Zhang, Y., and Wang, H. (2024). Battery degradation stage detection and life prediction without accessing historical operating data. Energy Storage Mater. *69*, 103441. https://doi.org/10.1016/j.ensm.2024.103441.