Review

Uncertainty-aware physics-guided digital twins for lithium-ion batteries: Toward integrated health, safety, and fast-charging management

Hanfeng Wu , Yuanxing Zhang*, and Daobin Mu*

School of Materials Science and Engineering, Beijing Institute of Technology, Beijing 100081, China.

***Correspondence to:** Prof. Yuanxing Zhang and Prof. Daobin Mu, School of Materials Science and Engineering, Beijing Institute of Technology, Beijing 100081,China. E-mail: 7520250041@bit.edu.cn; mudb@bit.edu.cn

Received: 12 May 2026 | Approved: 25 May 2026 | Online: 25 May 2026

Abstract

Lithium-ion battery management is moving from scalar state estimation toward integrated supervision of health, safety, and fast-charging capability. Existing reviews have addressed battery informatics, probabilistic prognostics, physics-guided learning, or digital-twin concepts as largely separate topics, but the connection among observability, mechanistic state representation, uncertainty quantification, and decision-oriented control remains underdeveloped. This Review reorganizes the field around uncertainty-aware physics-guided digital twins and argues that health loss, safety risk, and fast-charge limitations originate from coupled electrochemical, thermal, mechanical, and interfacial processes rather than from independent objectives. The discussion first identifies the shared physicochemical basis of degradation, with fast charging treated as a stringent validation case for a useful twin. The literature is then structured around four coupled layers: observability and data, physics-based and hybrid models, uncertainty-aware inference, and decision-facing control. On this basis, the Review examines health management beyond scalar state of health, plating-aware safety diagnosis, risk-constrained fast charging, and deployment from cell to pack and fleet levels. Finally, benchmarking, standardization, semantic data governance, and battery-passport requirements are discussed together with future opportunities in multimodal sensing, knowledge-enhanced twins, and autonomous closed-loop optimization. The central conclusion is that a deployable battery digital twin must be physics-grounded, uncertainty-calibrated, and explicitly designed to support decisions.

Keywords: Battery digital twin, lithium-ion batteries, uncertainty quantification, physics-guided machine learning, fast charging

INTRODUCTION

Lithium-ion batteries underpin electrified transportation, grid buffering, and portable electronics, but large-scale deployment remains constrained by the coupled requirements of high energy density, short charging time, long lifetime, low cost, and safety. Recent reviews in Advanced Materials and npj Computational Materials show that machine learning has accelerated the transition from empirical trial-and-error workflows toward data-centric battery science, while also exposing the limitations of purely empirical optimization when data coverage, task relevance, and experimental validation are inadequate[1-3]. For electric vehicles and stationary storage, the central challenge is therefore no longer limited to discovering better materials; it is to manage each battery system safely and efficiently as it ages under heterogeneous real-world use.

This systems-level challenge creates a natural role for digital twins. In the battery context, a digital twin should be understood as a continuously updated representation that integrates measurements, mechanistic priors, and inference to estimate internal states, predict degradation, quantify risk, and support operational decisions[4,5]. Such a representation is valuable only if it connects hidden electrochemical and thermal states to actionable outcomes: health estimation, safety warning, fast-charging supervision, maintenance, and lifecycle decisions.

Fast charging illustrates the promise and difficulty of this framework. Closed-loop learning has shown that high-dimensional charging protocols can be searched efficiently[6], whereas plating-aware sensing and operando diagnostics indicate that the practical limits of fast charging are controlled by latent electrochemical, thermal, and mechanical states rather than by voltage-current thresholds alone[7-10]. Fast charging should therefore be treated as a state- and risk-dependent control problem, not as a fixed protocol-design problem.

The main barrier is the gap between laboratory observability and field deployment. Vehicle batteries rarely experience ideal full cycles; instead, they operate through fragmented charging, partial discharge windows, thermal variability, cell-to-cell inconsistency, and sparse labels for true health states. Recent Communications Engineering and Nature Communications studies demonstrate that large EV datasets can support more realistic state-of-health estimation, but they also reveal persistent ambiguity in SOH definition, pack-level heterogeneity, and operational domain shift[11,12].

Existing reviews often cover only one part of this landscape—battery informatics, probabilistic health prognostics, physics-guided machine learning, or cloud battery management—and therefore stop short of a unified account of how sensing, modeling, inference, and decision making should be connected[2-5]. This revised review adopts a different organizing principle. Instead of cataloguing algorithms, it focuses on uncertainty-aware physics-guided digital twins as decision systems for integrated health, safety, and fast-charging management.

What is still missing is a framework that is sufficiently mechanistic to remain meaningful under distribution shift, yet sufficiently adaptive to remain useful under fragmented operating data and evolving battery conditions. This missing middle is precisely where battery digital twins become relevant. In the present context, a digital twin should not be reduced to a high-fidelity simulator, a software dashboard, or a machine-learning predictor. It should instead be understood as a decision-oriented representation of the battery that is continuously updated by measurements, informed by mechanistic priors, and evaluated against uncertainty. Such a definition is especially important in electrochemical systems because multiple internal degradation pathways can generate similar terminal voltage and capacity trajectories. Without an explicit internal-state representation, state estimation can become numerically accurate but physically ambiguous[13,35].

A second unresolved issue concerns scale. Much of the literature on battery diagnostics is still framed around single cells under well-controlled laboratory protocols, whereas practical battery management must operate at the level of packs, fleets, and lifecycle decision chains. Thermal gradients, current maldistribution, module imbalance, charging-station heterogeneity, and incomplete ground truth all alter the meaning of state estimation once deployment moves from the cell cycler to the vehicle or storage system[11-13]. Consequently, a review focused only on cell-level inference or only on algorithm families risks missing the systems problem that motivates digital twins in the first place.

A third unresolved issue concerns interpretation. A growing body of recent work argues that the central variables for battery management are not single scalar labels, but latent representations that connect what can be observed to what must be controlled. For health, this means degradation modes rather than capacity alone. For safety, it means a latent state of safety rather than post hoc threshold alarms. For charging, it means a risk-conditioned action set rather than a universally optimal protocol. These developments suggest that the most useful future battery-management frameworks will be structured around hidden states, uncertainty, and control relevance instead of around isolated prediction tasks[5,13,19,27,35].

This review therefore aims to provide a more closed analytical narrative than is typical in existing battery-ML surveys. Rather than treating health estimation, safety diagnosis, fast charging, sensing, uncertainty quantification, and deployment as parallel themes, the article argues that they are different expressions of the same multi-layered problem. The central thesis is that a deployable battery digital twin must be simultaneously physics-grounded, uncertainty-aware, and decision-oriented. The discussion is organized accordingly: Section 2 identifies the shared physicochemical basis of health loss, safety escalation, and fast-charging limitation; Section 3 defines the architecture of uncertainty-aware physics-guided digital twins; Section 4 examines digital-twin-enabled health management beyond scalar SOH; Section 5 turns to safety diagnosis and risk-constrained fast charging; Section 6 addresses deployment, benchmarking, and data governance; and Sections 7-8 synthesize future directions and the main conclusion.

Table 1 positions representative prior work according to the role it plays in this Review: data-centric battery science, probabilistic prognostics, physics-guided modeling, operando observability, field-data deployment, and lifecycle data governance.

Table 1. Representative prior studies grouped by conceptual role in the present review.

Conceptual strand	Representative article	Role in the present review
Data-centric battery research	Lv *et al.*, "Machine Learning for Materials Development and State Prediction in Lithium-Ion Batteries" (Adv. Mater. 2022) [1]	Establishes the broader AI landscape and motivates the shift from empirical workflows to data-centric battery science.
Battery informatics	Ling, "A review of the recent progress in battery informatics" (npj Comput. Mater. 2022) [2]	Clarifies the main data sources, bottlenecks, and machine-learning task families in battery research.
Application-oriented ML	Wang, "Application-oriented design of machine learning paradigms for battery science" (npj Comput. Mater. 2025) [3]	Reframes battery ML around task relevance, data adequacy, and experimental validation.

Conceptual strand	Representative article	Role in the present review
Physics + ML for management	Borah *et al.*, "Synergizing physics and machine learning for advanced battery management" (Commun. Eng. 2024) [4]	Provides the conceptual backbone for internal/external integration and battery digital twins.
Probabilistic battery prognostics	Thelen *et al.*, "Probabilistic machine learning for battery health diagnostics and prognostics" (npj Mater. Sustain. 2024) [5]	Defines uncertainty types, task taxonomy, and calibration-focused evaluation.
Fast-charging optimization	Attia *et al.*, "Closed-loop optimization of fast-charging protocols" (Nature 2020) [6]	Demonstrates learning-in-the-loop charging-policy optimization.
Plating-aware operando sensing	Huang *et al.* (pressure) [7], Zeng *et al.* (thermal-wave) [8], Wasylowski *et al.* (ultrasound) [9], Ishigaki *et al.* (MHz electromagnetics) [10]	Shows that plating and related side reactions are becoming partially observable during operation.
Automotive field data	von Bülow *et al.* (Commun. Eng. 2024) [11]; Liu *et al.* (Nat. Commun. 2025) [12]	Shows both the promise and the limitations of real-world data for pack-level health estimation.

Conceptual strand	Representative article	Role in the present review
Degradation-mode interpretation	Li *et al.*, "The importance of degradation mode analysis ..." (Nat. Commun. 2025) [13]	Explains why capacity-only fitting is insufficient and why degradation modes should be explicit state targets.
Hybrid and physics-informed learning	Wang *et al.* (PINN, Nat. Commun. 2024) [18]; Che *et al.* (mechanistic residual learning, Nat. Commun. 2026) [20]	Representative routes for embedding mechanism into learnable battery twins.
Digital-twin-enabled diagnosis	Guo *et al.*, "Digital Twin-Assisted Degradation Diagnosis During Fast Charging" (Adv. Energy Mater. 2024) [19]	Links fast-charging actions directly to aging modes and mechanistic interpretation.
Data semantics and governance	Clark *et al.*, "Toward a Unified Description of Battery Data" (Adv. Energy Mater. 2022) [16]; EU Battery Regulation 2023/1542 [32]	Shows why ontology, standardization, and battery passports are central to deployable battery twins.

SHARED PHYSICOCHEMICAL ORIGINS OF HEALTH, SAFETY, AND FAST CHARGING

Health estimation, safety protection, and fast charging are often handled as separate engineering problems, yet in practice they are different manifestations of one coupled degradation landscape. Mechanism-aware analyses show that capacity fade, power fade, plating propensity, and safety escalation are linked through shared electrochemical and mechanical pathways, including SEI growth, electrolyte depletion, lithium plating, active-material loss, and particle cracking[5,13]. This observation matters because very different internal states can produce similar external capacity-loss trajectories, especially when only voltage or resistance are used as labels[13]. A useful digital twin must therefore explain not only how much performance has been lost, but why it has been lost.

This is why degradation modes are more informative than scalar health labels. The now common decomposition into lithium inventory loss (LLI), positive-electrode loss of active material (LAM_PE), and negative-electrode loss of active material (LAM_NE) offers a physically meaningful middle layer between latent mechanisms and measurable performance[5,13]. Once health is expressed in this intermediate form, state estimation, remaining-life prediction, and charging control can be connected more naturally because the twin can reason over the mechanism space instead of over a single capacity number.

Fast charging is the harshest validation case for such reasoning. Under low temperature, high current, or heterogeneous internal transport conditions, local overpotentials rise and lithium plating becomes more likely, while thermal and structural nonuniformity amplify the damage[6,7]. Operando pressure, thermal-wave, ultrasound, and MHz electromagnetic studies further show that plating and related side reactions are spatially distributed, time-varying phenomena rather than discrete threshold events[7-10]. As a result, the limiting quantity for fast charging is not current alone but the joint evolution of latent electrochemical, thermal, and risk states.

The same logic extends from cell to pack scale. In practical EV systems, the relevant state space includes not only cell-level SOC and degradation modes, but also pack-level thermal gradients, current maldistribution, worst-cell margin, and evolving operational context[11,12]. The lab-to-field gap therefore changes the task definition of a digital twin: the objective is no longer the accurate fitting of curated cycling data, but robust state reconstruction and decision support under partial observability, domain shift, and sparse ground truth.

Two analytical consequences follow from this observation. The first is that identifiability becomes a central issue. If different combinations of LLI, LAM_NE, interfacial impedance growth, electrolyte depletion, and structural heterogeneity can reproduce similar external degradation curves, then capacity-only model fitting is intrinsically underdetermined. A management framework built only on terminal variables may therefore remain blind to whether the limiting factor is lithium inventory, active-material accessibility, transport resistance, or plating risk. This is not a semantic nuance. It directly affects whether a battery should be charged more conservatively, thermally conditioned, derated, repurposed, or retired[10,13].

The second consequence is that health, safety, and fast charging cannot be separated by time scale as cleanly as they often are in engineering practice. Slow interphase growth, transition-metal dissolution, and active-material isolation reshape local transport pathways; those transport limitations alter overpotential distributions during fast charging; and those overpotentials change the likelihood of plating, gas generation, or localized heat release. In that sense, fast charging is not an independent objective layered on top of a healthy battery. It is an operational regime that exposes the evolving internal state of the battery more aggressively than mild cycling does[6,7,9,36].

This coupling is particularly visible when degradation is viewed through the lens of propagation. At the cell level, the relevant progression may be from interfacial growth to transport limitation, then to plating or self-heating. At the pack level, the same progression is amplified by nonuniform cooling, electrical interconnection, and balancing constraints. A cell that is only moderately aged in isolation can become the worst cell in a series-connected string, limiting both pack power and safe charging current. The cell-to-pack transition is therefore not a mere scaling problem; it changes which latent variables become safety-critical and which observables retain diagnostic value[11,12].

Recent safety-focused reviews have begun to articulate this bridge more explicitly by treating degradation, fault evolution, and thermal runaway as a continuum rather than as isolated failure classes[13,35,50,58]. That viewpoint is highly relevant here because battery digital twins must eventually operate across the same continuum. A useful twin should therefore not be designed only to estimate gradual ageing or only to detect rare events. It should instead be able to represent how long-term ageing modifies near-term hazard thresholds and how abnormal operation reshapes future degradation trajectories.

Figure 1 summarizes this coupled landscape from application need to latent mechanism. The application map in Figure 1a sets the management requirements. Figure 1b and Figure 1c then show that capacity fade and power fade arise from interacting degradation submodels and degradation modes rather than from a single ageing coordinate. Figure 1d and Figure 1e add spatially resolved plating and electrothermal nonuniformity. Figure 1f, Figure 1g, and Figure 1h show how these mechanisms become difficult to infer in EV field data. Figure 1i therefore frames the digital twin as a systems-level response to a many-to-one inverse problem: multiple latent degradation pathways can project onto similar terminal signals, so management requires an internal-state representation rather than a scalar label alone.

A useful way to make this coupling more concrete is to separate the latent failure landscape from its observable manifestations. At the latent level, the primary actors are interphase growth, active-material isolation, electrolyte consumption, plating, crack nucleation, oxygen release, and heat generation. At the observable level, the quantities seen by the BMS are far more compressed: voltage trajectory, apparent capacity, impedance rise, relaxation behavior, pressure, temperature, and, occasionally, gas emission. The mapping from the first set to the second is many-to-one. The same terminal voltage depression can arise from lithium loss, transport limitation, thermal imbalance, or contact degradation. Conversely, a small change in one observable, such as dP/dQ, may carry disproportionate information about a specific latent mechanism such as plating[7,12,13,35,36,43]. The digital-twin problem is therefore fundamentally an inverse problem under structural non-uniqueness.

Recent degradation-and-safety studies further suggest that ageing modifies not only the mean performance trajectory but also the shape of the hazard landscape. Aged cells may exhibit lower tolerance to mechanical abuse, altered separator resilience, earlier self-heating onset, modified gas-evolution chemistry, and different thermal runaway heat release. This means that degradation cannot be treated as a neutral background trend that simply lowers capacity. It actively changes the triggering conditions, escalation pathways, and observability of extreme events. Put differently, ageing does not merely precede safety failure; it conditions the entire fault-to-hazard transition[10,13,41,42,46,47]. For a digital twin, this implies that safety states should be conditioned on health states, and health states should be updated with awareness of safety-relevant operating history.

A related point is that fast charging should be understood as an information-rich perturbation. Under mild cycling, many harmful mechanisms evolve slowly and remain partly latent. Under fast charging, local overpotentials, thermal gradients, transport limitations, and interfacial instabilities are amplified, making otherwise subtle differences among cells more observable. This is one reason why fast charging is such a stringent validation scenario for battery twins: it magnifies both the opportunity for state discrimination and the penalty for model misspecification. A twin that remains well calibrated under fast charging is likely to be informative under gentler conditions; the converse is not necessarily true[6-10,19,36,39].

The coupling among degradation, fault development, and catastrophic failure can also be viewed through a timescale hierarchy. Slow processes such as SEI growth, lithium inventory loss, and contact degradation reshape the baseline electrochemical state over weeks to months. Intermediate processes, such as localized lithium deposition, gas generation, or thermal imbalance under repeated high-rate operation, evolve over cycles or hours. Very fast processes, such as internal short circuits, violent gas release, or thermal runaway, can unfold in minutes or seconds. A battery digital twin that aims to be decision-useful must therefore preserve continuity across these timescales. It is not enough to estimate long-term health without relating it to short-term hazard, nor to detect short-term anomalies without situating them within the ageing state that conditioned them[10,13,35,41,42].

This continuity matters because interventions also operate on different timescales. Design changes, electrolyte choices, and formation protocols act before the battery is deployed. Thermal conditioning, charging control, and balancing operate during use. Maintenance, warranty, repurposing, and recycling decisions act later in the lifecycle. The same hidden state may influence all of them, but in different ways. For example, a pack with modest average capacity fade but a strongly degraded worst cell may still be acceptable for some use scenarios while being inappropriate for extreme fast charging or second-life deployment. An integrated digital twin is attractive precisely because it can preserve this causal continuity across lifecycle stages rather than allow the battery to be reinterpreted independently at each stage[16,17,31,32].

The degradation-to-safety bridge can also be interpreted in terms of progressive versus abrupt failure regimes. In a progressive regime, latent damage accumulates through repeated cycling, thermal gradients, local plating, gas accumulation, or contact degradation until the battery enters a state in which even routine operation becomes hazardous. In an abrupt regime, an acute trigger such as severe short circuit or mechanical damage drives the battery into failure on a short timescale, yet the severity of that event is still conditioned by the pre-existing degradation state. This distinction is useful because it shows that battery safety is neither wholly slow nor wholly sudden. It is conditioned by ageing but revealed by events. A practical digital twin must therefore represent both the gradual movement of the battery toward risk and the acute transitions by which that risk becomes visible[13,35,41,42,46,50].

A remaining scientific gap is to translate degradation modes into risk thresholds that are both physically identifiable and operationally actionable. Recent studies on localized high-temperature instability, side-reaction amplification, thermal-runaway onset, and aged-cell risk asymmetry show that the same nominal capacity loss can correspond to very different hazard envelopes depending on interfacial chemistry, heat release characteristics, and abuse history. In other words, the relevant state is not merely how much the battery has degraded, but how that degradation reshapes the latent pathways by which faults escalate. This is precisely the level of closure a useful digital twin must achieve: it must connect LLI/LAM-type degradation coordinates to transport limitation, self-heating propensity, short-circuit susceptibility, and thermal-runaway severity rather than treating these as independent labels[68,70,71,85-96].

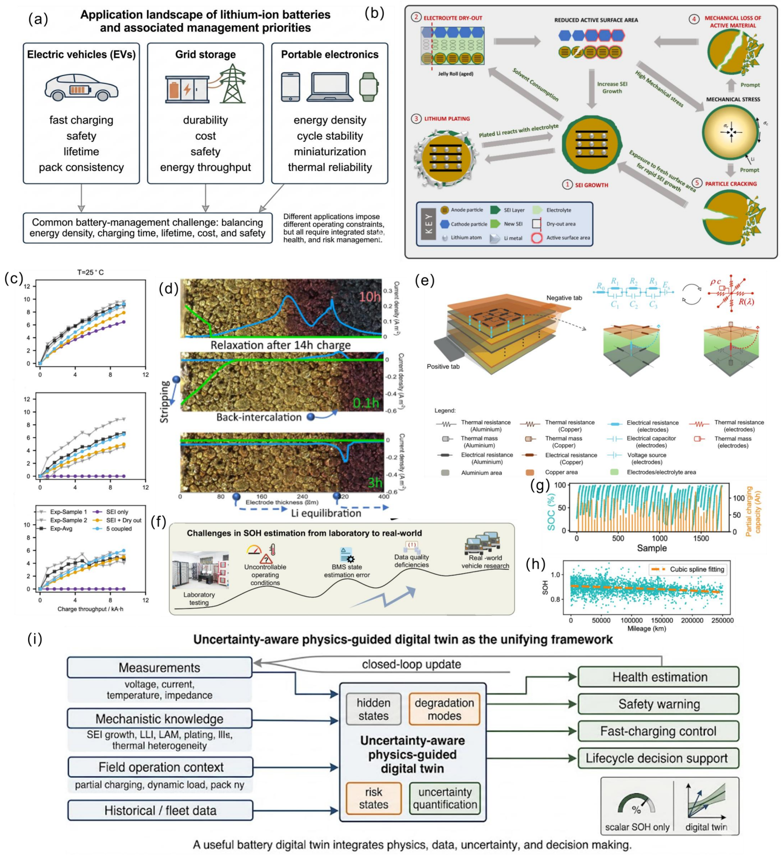

Figure 1. Shared physicochemical origins of health, safety, and fast charging. (a) Application landscape and management priorities for lithium-ion batteries; (b) Coupled degradation submodels, including SEI growth, electrolyte dry-out, lithium plating, active-material loss, and particle cracking[13]; (c) Degradation-mode trajectories involving LLI, LAMNE, and LAMPE[13]; (d) Phase-field prediction of lithiation/delithiation and plating/stripping current distributions during charge and relaxation[35]; (e) Electrothermal network representation of a pouch cell under spatially resolved thermal gradients[36]; (f) Real-world EV SOH-estimation challenges; (g) Fragmented EV charging behavior; (h) Fleet-level SOH variability[12]; (i) Uncertainty-aware physics-guided digital twin as the unifying framework.

ARCHITECTURE OF UNCERTAINTY-AWARE PHYSICS-GUIDED DIGITAL TWINS

Definition and scope

In this review, an uncertainty-aware physics-guided digital twin is defined as a continuously updated digital representation of a battery system that couples measurement streams, mechanistic priors, and data-driven adaptation to estimate hidden states, predict degradation, quantify risk, and support control[4,5]. This definition deliberately distinguishes a digital twin from both a standalone simulator and a standalone machine-learning model. A simulator may provide mechanistic consistency without sufficient adaptability, while a black-box predictor may interpolate well without clarifying what is physically happening inside the cell.

The qualifier uncertainty-aware is essential. Battery behavior varies across cells, across usage histories, and across operational domains, so any practically useful twin must represent what is known and what remains ambiguous. Following recent probabilistic battery reviews, uncertainty should be carried through the full stack: from noisy or missing observations, to imperfect model form and drifting parameters, to predictive risk under unseen conditions[5]. This is especially important when twin outputs are used to determine charging aggressiveness or safety margins, where overconfidence can be more harmful than modest point-prediction error.

In practice, this definition implies that the twin must sit between measurement and action, rather than at either end of the pipeline. If it is reduced to an offline model, it cannot adapt. If it is reduced to a direct control policy, it cannot explain or calibrate. If it is reduced to a static dashboard of measured signals, it cannot infer latent states. The distinctive value of a digital twin therefore lies in preserving a mechanistically meaningful internal state while remaining open to data-driven correction and uncertainty-aware updating[4,5,13,28].

A second implication is that the twin should be evaluated by the quality of the decisions it enables, not only by the error of the intermediate variables it predicts. Accurate SOH estimation is important, but the engineering question is whether that estimate leads to more appropriate charging limits, thermal interventions, maintenance recommendations, or second-life decisions. Likewise, a physically elegant model is not automatically a useful twin if it cannot be synchronized with measurements or if its uncertainty is not interpretable to downstream control layers. This decision-facing perspective is what differentiates a battery digital twin from an advanced battery model[4,5,38].

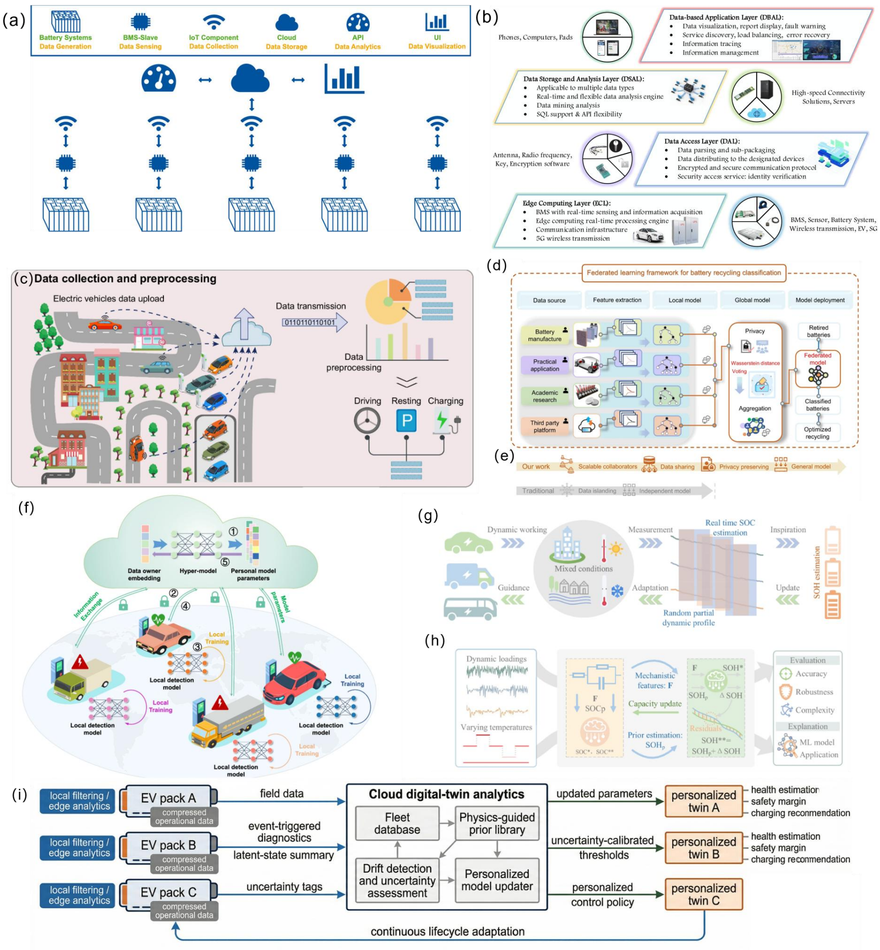

Figure 2. Closed-loop architecture of uncertainty-aware physics-guided digital twins. (a) Closed-loop battery digital twin linking observability, physics-based representation, uncertainty-aware inference, and decision feedback[37]; (b) Cloud BMS architecture for data generation, sensing, collection, cloud storage, analytics, and visualization[37]; (c) Multimodal field-data feature construction for EV SOH estimation, including voltage maps, charge-capacity and temperature sequences, and point features[12]; (d) Multi-contributor federated learning framework for retired-battery sorting without raw-data exchange[38]; (e) Privacy-preserving collaboration compared with the conventional data-islanding paradigm[38]; (f) Personalized federated fault-warning architecture using a central hyper-model and local owner-specific models for massive EV fleets[28]; (g) Closed-loop residual-learning demonstration for battery monitoring[20]; (h) Mechanistically inspired SOC/SOH monitoring pipeline for lifelong state correction[20]; (i) Cloud-edge-fleet updating and personalization loop for battery digital twins.

DOI: 10.20517/scierxiv202605.0340.v1

Figure 2 should be read as the architectural bridge between the conceptual definition of a battery digital twin and the observability layer discussed below. Figure 2a defines the minimum closed loop in which observations are translated into a mechanistic state representation, uncertainty is carried through inference, and the estimated state constrains action. Figure 2b extends this loop into a cloud BMS architecture for data sensing, storage, analytics, and visualization. Figure 2c shows how field data are converted into multimodal health features rather than treated as unstructured telemetry. Figure 2d and Figure 2e introduce privacy-preserving retired-battery sorting, where data contributors train local models and share parameters instead of raw data. Figure 2f moves the same logic to fleet-scale fault warning through a central hyper-model and owner-specific local models. Figure 2g and Figure 2h then show how mechanistic priors and residual learning can keep SOC/SOH estimates synchronized as cells age. Figure 2i abstracts these elements into a cloud-edge-fleet personalization loop. The figure therefore does not merely collect management platforms; it defines the information pathway by which a twin remains synchronized, individualized, and decision-relevant under field conditions.

Observability, data sources, and the lab-to-field gap

The first architectural pillar is observability. Conventional voltage, current, and surface temperature signals remain indispensable because they are inexpensive, continuous, and already available in most BMS hardware[11]. Yet recent work makes clear that these channels alone are not sufficient for many of the most consequential hidden variables. Online impedance concepts[15], data ontologies for battery semantics[16], and multiscale characterization frameworks spanning electrochemical, thermal, and microstructural measurements[17] together suggest that digital twins should be built from a hierarchy of data modalities rather than from one monolithic timeseries.

This hierarchy should combine at least four sources of information: operational data from field use, laboratory cycling and diagnostic data, high-fidelity simulation or virtual experiments, and structured knowledge extracted from the literature or manufacturing records[2,11,16,17]. The practical implication is that data engineering becomes part of model design. What the vehicle stores, what the cloud aggregates, and what the twin treats as a label or latent state all shape the eventual notion of battery health and risk.

From the standpoint of observability, the key challenge is not only that the most consequential variables are hidden, but that the available observables have unequal diagnostic value across conditions. Voltage is highly informative near equilibrium but loses specificity during dynamic load transients. Temperature is useful for thermal imbalance but often lags internal reactions. Pressure, ultrasound, high-frequency electromagnetics, and impedance each reveal different slices of the internal state, but none offers a complete picture on its own. A deployable twin therefore requires modality selection rather than modality accumulation: each signal should be judged by the hidden states it makes more identifiable, the conditions under which it remains informative, and the cost of acquiring it at scale[7-10,15,58-74].

This is one reason why the lab-to-field gap should be treated as an observability problem rather than merely as a data-volume problem. In the laboratory, one can often impose full cycles, control temperature, and measure reference quantities with high fidelity. In the field, one receives fragmented charging windows, irregular driving loads, varying ambient conditions, missing metadata, and pack-level interactions. The question is therefore not simply whether a laboratory-trained model generalizes, but whether the field data contain enough information to reconstruct the variables that the model was built to estimate[11,12,20,25,41].

The answer may differ by task. For capacity-oriented health tracking, partial-charge or fragmented-charge features may suffice[7,14]. For degradation-mode diagnosis, richer signatures such as impedance, relaxation, or controlled perturbations may be required[20,21]. For safety and fast charging, multimodal sensing can become decisive because the relevant precursors are spatial, local, and strongly condition-dependent[7-10,26,39]. The design of a digital twin therefore begins not with architecture diagrams but with a rigorous statement of which hidden variables are needed for which decisions, and which measurements make those variables recoverable.

This point also implies that observability is partly an experimental-design problem. If the relevant latent states are weakly expressed in passive field trajectories, then the twin must either exploit naturally informative events—such as charging segments, relaxation intervals, or pack imbalance transients—or call on lightweight active probing, for example through impedance excitation, controlled balancing perturbations, or fault-residual monitoring. Recent work on hybrid fault diagnosis, degradation diagnostics, and safety-oriented estimation makes clear that the quality of a battery twin depends not only on the model class employed but also on whether the available evidence is sufficiently informative for the variables being estimated[48,97-110].

An equally important principle is identifiability under limited stimulation. In laboratory diagnostics, one can deliberately impose informative experiments such as full CC-CV cycles, relaxation sequences, impedance sweeps, or controlled thermal perturbations. In real operation, the battery often experiences whatever the user or the application demands. This means that the twin must infer as much as possible from passively observed trajectories while also recognizing when passive observation is insufficient. The most useful architectures may therefore combine passive monitoring with occasional active probing, for example through controlled charging segments, balancing-resistor perturbations, or opportunistic impedance measurements[15,21,25,29,65-74].

Table 2 maps these observability trade-offs by relating each modality to the hidden states it makes more identifiable, its strengths, its deployment limits, and its present maturity.

Table 2. Observability modalities for deployable battery digital twins.

Modality	Hidden states informed	Advantages	Deployment limits	Maturity

Modality	Hidden states informed	Advantages	Deployment limits	Maturity
Voltage, current, temperature	SOC; ohmic trend; pack imbalance; coarse thermal state	Low cost; continuous; already in BMS	Limited specificity for local interfacial events	High
Impedance / EIS	Charge-transfer resistance; diffusion limit; interfacial evolution	Mechanism-rich; supports probabilistic forecasting	Excitation/synchronization burden; limited field deployment	Medium
Pressure / strain	Plating; gas generation; swelling; mechanical integrity	Early sensitivity to local side reactions	Packaging dependence; difficult pack integration	Medium
Thermal-wave / internal sensing	Effective thermal conductivity; degradation pathway; self-heating precursor	Pathway-sensitive thermal signatures	Additional hardware and contact design required	Medium

Modality	Hidden states informed	Advantages	Deployment limits	Maturity
Ultrasound / magnetic diagnostics	Internal heterogeneity; plating morphology; current maldistribution	Non-destructive access to spatially distributed states	High instrumentation complexity; limited automotive maturity	Low-Medium
Field / fleet metadata	Usage context; transfer regime; owner-specific risk; lifecycle traceability	Critical for deployment realism and personalization	Incomplete, heterogeneous, and weakly standardized	High collection / low semantic integration

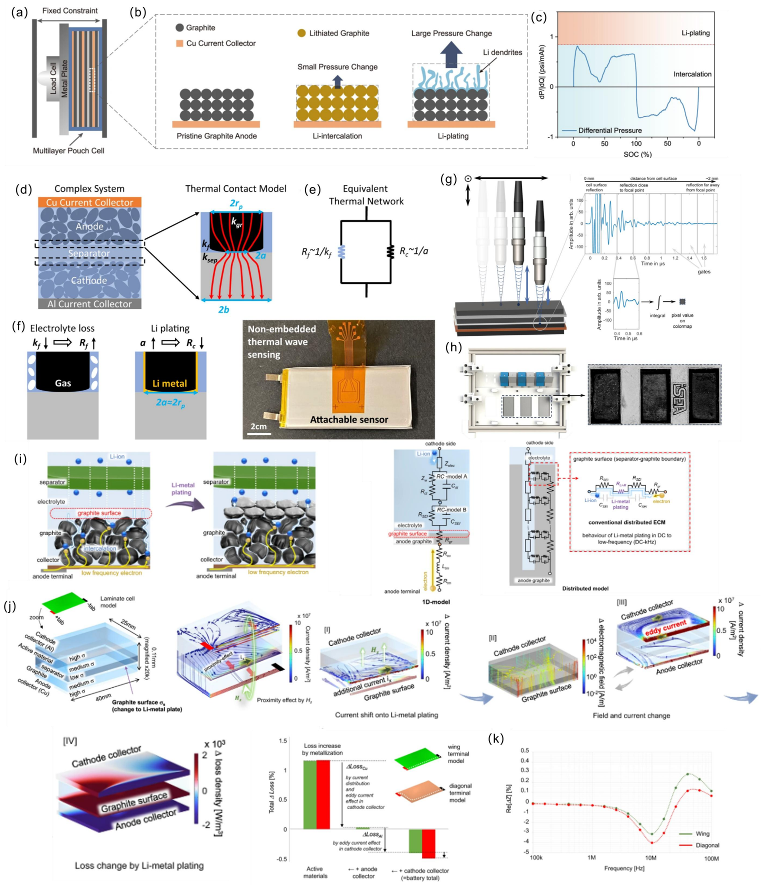

Figure 3. Multimodal observability for hidden electrochemical, thermal, and mechanical states. (a) Operando differential-pressure sensing configuration for pouch cells[7]; (b) Pressure-change mechanism distinguishing lithium plating from intercalation[7]; (c) Pressure-derived lithium-plating threshold from the dP/dQ profile[7]; (d) Thermal-wave unit-cell and contact model for porous electrodes[8]; (e) Equivalent thermal network for separator-electrode contact[8]; (f) Attachable thermal-wave sensing for pathway-resolved degradation interpretation[8]; (g) Ultrasound scanning and gated signal acquisition for operando plating visualization[9]; (h) Signal-to-image reconstruction for ultrasound imaging[9]; (i) Low-frequency electromagnetic and equivalent-circuit interpretation of Li-metal plating[10]; (j) High-frequency FEM analysis of current-density and electromagnetic-field response to plating[10]; (k) Terminal-impedance response over a wide frequency range[10].

Together, Figure 3a, Figure 3b, and Figure 3c show why mechanical pressure can expose lithium plating earlier than voltage-only monitoring. Figure 3d, Figure 3e, and Figure 3f extend observability to thermal-contact and degradation-pathway information. Figure 3g and Figure 3h add spatially resolved ultrasound imaging, whereas Figure 3i,

Figure 3j, and Figure 3k show that high-frequency electromagnetic response can encode metallic-lithium signatures. These examples support the central point that deployable twins require state-targeted observability rather than indiscriminate sensor accumulation.

Physics-based models and hybrid learning

The second architectural pillar is a task-matched model hierarchy. For high-frequency onboard tasks, such as SOC updates or power limitation, low-order equivalent circuit models remain attractive because they are lightweight and interpretable. For root-cause analysis, plating prediction, or lifetime extrapolation, higher-fidelity electrochemical-thermal-degradation models are more appropriate because they explicitly represent concentration gradients, overpotentials, transport bottlenecks, and mechanism-specific aging[4,13]. The key point is not that one model class is universally superior, but that different time scales and decisions require different levels of mechanistic detail.

Physics-guided machine learning becomes valuable precisely at this interface. Internal integration approaches embed physical relations into the architecture or loss function, as illustrated by physics-informed SOH modeling[18]. External integration approaches use mechanistic models as priors, feature generators, simulators, or residual backbones, as seen in digital-twin-assisted degradation diagnosis under fast charging[19] and mechanistically guided residual learning for lifelong state monitoring[20]. In both cases, the goal is not to decorate a neural network with physical terminology but to improve sample efficiency, extrapolation, and state interpretability by constraining learning within a physically meaningful representation.

A useful way to interpret this hierarchy is through task allocation. Equivalent-circuit models remain attractive when computational parsimony and real-time execution dominate, for example during onboard SOC estimation, online filtering, or control-oriented state tracking. Reduced-order electrochemical models become valuable when concentration gradients, overpotentials, or plating thresholds must be represented explicitly but full P2D fidelity is still unaffordable. High-fidelity electrochemical-thermal-degradation models become necessary when the management task itself depends on mechanism attribution, such as understanding how charging protocols redistribute ageing modes or why a local hotspot appears under a given cooling topology[4,13,19,38].

This also explains why model simplicity and model usefulness are not the same thing. A model can be numerically stable but semantically poor if it suppresses the very states that matter for decision making. Conversely, a high-fidelity model can be scientifically valuable yet operationally irrelevant if it cannot be synchronized with the available measurements or cannot be updated online. The key design question is therefore not whether a model is simple or complex, but whether its state variables are aligned with the management objective. In battery digital twins, state-space design is therefore a central scientific act rather than an implementation detail[4,13,25].

Hybrid learning becomes meaningful in this context because it allows the twin to inherit the state semantics of physical models while using data-driven components to absorb unresolved heterogeneity, missing physics, or field-induced drift. Physics-informed neural networks, residual learners, reduced-order surrogates, and domain-adaptive mappings all belong to this broader strategy. Their value is greatest when the physical prior remains informative but incomplete, which is precisely the regime encountered in real battery systems[8,18-21,25,38].

Figure 4 links the physical model hierarchy to hybrid and probabilistic learning in sequential order. Figure 4a introduces a cross-scale electrochemical-mechanical model that connects particle-scale strain, electrode-scale deformation, and cell-level voltage-strain response, providing a mechanistic backbone distinct from the electrothermal heterogeneity shown in Figure 1e[39]. Figure 4b identifies the chemistry and usage heterogeneity that any transferable model must accommodate, Figure 4c shows the short-window feature extraction used for physics-informed SOH estimation, and Figure 4d gives the PINN architecture that constrains degradation learning[18]. Figure 4e, Figure 4f, and Figure 4g then show training, DNN-swarm structure, and estimation in a domain-adaptive workflow[23]. Figure 4h, Figure 4i, and Figure 4j move to impedance-based forecasting, where battery state and future action jointly determine performance[21,25]. The resulting hierarchy should be read as a mapping between physical fidelity, state semantics, and decision relevance, not as a simple ranking of model complexity.

This suggests an additional design principle: battery digital twins should be built around state variables that are operationally sufficient rather than merely physically exhaustive. An exhaustive state description of a lithium-ion cell would be intractable for most deployment contexts. An operationally sufficient description, by contrast, preserves the latent variables needed to support the decisions of interest. For health management, this may be a vector of capacity, resistance, LLI, LAM, and uncertainty. For safety management, it may be a latent hazard state coupled to internal temperature, pressure, and escalation likelihood. For fast charging, it may be the admissible action set conditioned on plating risk and thermal margin. The architecture should therefore be judged by sufficiency for intervention rather than by formal fidelity alone[4,5,13,28,38].

A second design principle is synchronization. A digital twin that cannot remain synchronized with real measurements is not a twin in any useful sense. Synchronization is difficult in batteries because the information content of incoming data is highly nonuniform across time. A full CC-CV charge may be highly informative; a brief partial charge may not be. A thermal transient may be diagnostic under one condition and uninformative under another. Real systems therefore require event-aware updating rules that weight incoming measurements by their state informativeness, rather than simply by their recency. This is one reason why cloud-assisted and edge-assisted battery twins are attractive: they allow asynchronous data aggregation, event-triggered diagnostics, and model updates at timescales that match the underlying physics[12,20,25-28,37,38,47].

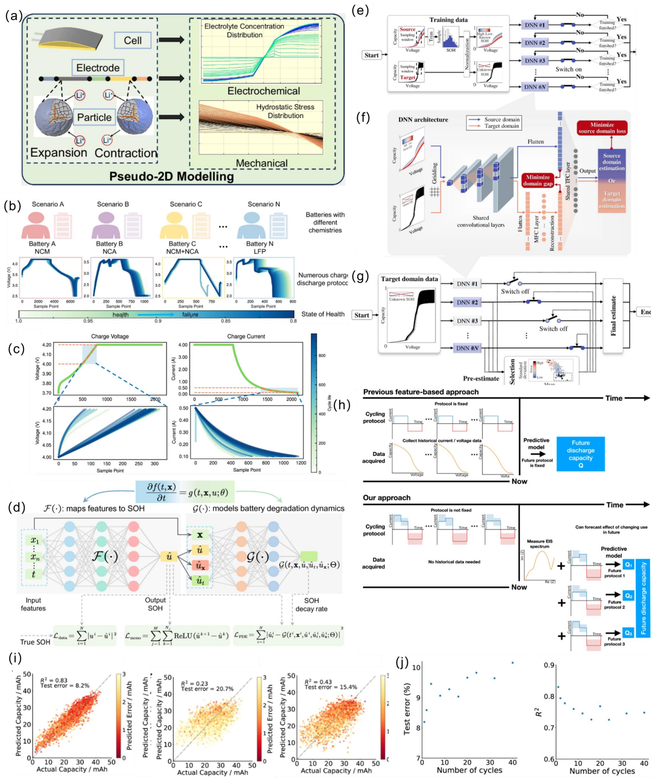

Figure 4. Physics-based models, hybrid learning, and uncertainty-aware inference. (a) Cross-scale electrochemical-mechanical modeling framework linking particle-scale strain, electrode-scale displacement, and cell-level voltage-strain response for high-fidelity battery twins[39]; (b) Chemistry and user-dependent degradation heterogeneity motivating physics-informed learning[18]; (c) Short charging-window feature extraction for SOH estimation[18]; (d) Physics-informed neural-network architecture for stable degradation modeling and SOH prognosis[18]; (e) Training procedure for domain-adaptive SOH estimation[23]; (f) DNN-swarm architecture for source-target transfer[23]; (g) Estimation procedure without additional degradation experiments[23]; (h) Impedance-based forecasting framework under variable future protocols[21]; (i) State-action forecasting evidence showing the joint need for EIS-derived state and protocol action[21]; (j) Multi-step forecasting performance under uneven usage[21,25].

Uncertainty-aware inference and decision layers

The third architectural pillar is uncertainty-aware inference. Rather than returning a single SOH value, the twin should estimate a probability distribution over current states, degradation modes, and future risk, ideally with calibrated measures of epistemic and aleatory uncertainty[5]. Impedance-based probabilistic forecasting illustrates the value of this view: battery condition is better described as a multidimensional state vector than a single scalar health index, and predictive uncertainty becomes part of the state itself when future usage is variable or partially unknown[21].

The fourth pillar is decision and deployment. A useful twin must map estimated states and uncertainties onto actions such as charging-rate selection, thermal intervention, power derating, maintenance recommendation, or fleet-level model update[4,5]. This naturally motivates a cloud-edge-fleet architecture: the vehicle handles real-time filtering and safety execution, while cloud resources perform heavier health diagnosis, cross-vehicle learning, drift detection, and personalized model adaptation. Seen this way, the digital twin is less a single model than a layered decision system built around observability, mechanism, inference, and control.

In battery systems, uncertainty has at least four operationally distinct origins. Measurement uncertainty arises from sensor noise, drift, limited resolution, and missing data. Parametric uncertainty arises because latent electrochemical and degradation parameters are not directly measured and may drift with age or environment. Structural uncertainty arises because all models suppress part of the underlying physics. Finally, deployment uncertainty arises because the operating domain itself shifts across users, climates, battery types, charging infrastructures, and lifecycle stages. A digital twin that collapses these sources into a single deterministic estimate may still interpolate well on familiar data, but it cannot support trustworthy management when conditions change[5,11,12,25,38].

The practical consequence is that uncertainty must be propagated to the decision layer rather than reported only at the estimator output. A charging controller should not only know the inferred plating threshold, but also how uncertain that threshold currently is. A maintenance decision should not only use the estimated health state, but also the uncertainty around whether the limiting mechanism is LLI, LAM, impedance rise, or a pack-level imbalance effect. In this sense, calibration is not an optional reporting metric; it is part of the physical usefulness of the twin[5,21,27].

DOI: 10.20517/scierxiv202605.0340.v1

The relation between uncertainty and deployment is particularly strong in battery applications because many operational choices are inherently asymmetric. A slightly over-conservative charging limit may cost time; an overconfident charging limit may trigger plating or self-heating. A slightly conservative fault alarm may cause service interruption; an overconfident miss may allow hazard escalation. This asymmetry is why battery digital twins should be regarded as risk-mediating systems. Their purpose is not to eliminate uncertainty, which is impossible, but to structure it in a form that can inform action[5,18,23,27].

A third design principle is semantic continuity. Battery models are often built for one lifecycle phase at a time: formation, first use, ageing, second life, or recycling. However, deployment increasingly requires continuity across these phases. Manufacturing genealogy affects initial heterogeneity; operating history affects second-life value; safety incidents affect recycling logistics; regulatory descriptors affect what can be audited. A battery twin that fragments across lifecycle stages loses much of its practical value. This is why ontology, battery passport, and digital-thread concepts should be understood not as external data-management layers but as part of the architecture of deployable twin systems[12,16,17,31,32,37].

The same task-matched logic applies to model selection: the appropriate model depends on which hidden states must be recovered and which decisions must ultimately be supported. In practice, model hierarchy should therefore be interpreted as a mapping between physical fidelity, state semantics, and decision relevance rather than as a simple ladder from low to high complexity.

This view also clarifies the role of machine learning in battery twins. Machine learning is not most useful when it replaces the mechanistic model entirely, but when it learns which parts of the state are identifiable under current data conditions, how uncertainty should be recalibrated after domain shift, and where mechanistic priors systematically fail. In other words, the strongest digital-twin use cases are not those in which AI suppresses physics, but those in which AI helps determine when physics, measurement, and decision constraints are jointly informative enough to act[4,5,26-28,38,44-46].

DOI: 10.20517/scierxiv202605.0340.v1

For this reason, calibration should be treated as a design objective from the outset rather than as a post hoc evaluation metric. A twin that is slightly less accurate in mean estimation but reliably calibrated may be more useful than a sharper but overconfident alternative, especially in charging and safety applications. This is because battery-management decisions are almost always asymmetric under uncertainty. The cost of an unnecessary warning or moderate derating is often much smaller than the cost of missing a true escalation. The architecture should therefore be built to preserve not only state semantics but also calibration semantics, so that downstream control layers can interpret what level of confidence is being reported and act accordingly[5,21,27,38].

A related requirement is fault tolerance in the architecture itself. Practical twins must continue to function under partial sensor loss, missing metadata, asynchronous uploads, or occasional drift in the underlying models. This argues for modular architectures in which sensing, state estimation, hazard assessment, and decision support can degrade gracefully rather than fail monolithically. In that respect, the most valuable future twins may be those that are not only accurate when fully informed, but also robustly useful when partially informed—a condition that better reflects deployment reality[11,12,28,37,38].

Table 3 organizes the model layer by state variables, strengths, limitations, and best-fit tasks, emphasizing that model choice in a digital twin is task-matched rather than universal.

Table 3. Model classes and their deployment roles in battery digital twins.

Model class	State variables	Advantages	Limits	Best-fit tasks

Model class	**State variables**	**Advantages**	**Limits**	**Best-fit tasks**
Equivalent-circuit models	SOC; polarization; lumped resistance/capacitance	Fast; interpretable; control-oriented	Weak mechanism resolution; limited extrapolation for plating/safety	Onboard filtering; SOC; power limitation
Reduced-order electrochemical models	Concentration gradients; overpotentials; effective transport states	Better mechanism fidelity at manageable cost	Still requires calibration and simplification	Fast charging; diagnostic supervision; adaptive control
Electrochemical-thermal-degradation models	Electrode potentials; heat generation; plating tendency; degradation modes	Strong physical interpretability; causal structure	Heavy parameterization; computational burden	Root-cause analysis; safety; cloud diagnosis
Physics-informed learning	Learned health trajectories constrained by physics	Improved sample efficiency; state consistency	Depends on adequacy of embedded physics	SOH/SOP/SOS under partial data

Model class	State variables	Advantages	Limits	Best-fit tasks
Residual / hybrid learning	Mechanistic priors + data-driven corrections	Good deployment compromise; adapts to drift	Requires stable prior model and calibration loop	Lifelong monitoring; personalization; cloud-edge updates
Probabilistic models	Distributions over states, parameters, and risk	Explicit uncertainty; supports risk-aware decisions	Calibration and dataset realism are critical	Safety; charging; maintenance; decision support

DIGITAL-TWIN-ENABLED HEALTH MANAGEMENT

One of the clearest consequences of this framework is that health management can no longer be organized around a single scalar SOH definition. Vehicle-oriented discussions of SOH emphasize that capacity-based, energy-based, and resistance-based definitions each capture only part of the future usefulness of a battery[11]. A twin that is meant to support charging, safety, second-life grading, or warranty decisions therefore needs a vector-like health representation that can simultaneously encode capacity loss, power limitation, degradation-mode composition, and risk margins[21].

Practical health management begins with deployable health indicators. Early machine-learning pipelines demonstrated that carefully engineered features extracted from partial charge curves can estimate SOH with usable confidence intervals[22]. More recent domain-knowledge-guided work moves this idea closer to the field by constructing indicators that remain meaningful under fragmented charge windows and real driving discharges[14]. This is an important conceptual shift: useful health indicators are not simply statistically predictive; they must also remain computable under realistic operating constraints.

A second shift is from capacity estimation to degradation diagnostics. If multiple internal pathways can produce similar capacity-loss curves, then maintenance and charging decisions require more than a scalar estimate. Mechanism-aware studies now argue that LLI, LAM_NE, and LAM_PE should be treated as explicit diagnostic targets, not merely as post hoc explanations[5,13]. In this context, information-rich measurements such as electrochemical impedance can act as compressed fingerprints of hidden electrochemical processes and support probabilistic forecasting even when detailed usage history is unavailable[21].

The third shift concerns generalization. Health models trained on single chemistries or narrow laboratory protocols often fail when transferred to new manufacturers, broader operating conditions, or field data. Domain-adaptive deep learning has shown that usable SOH estimation can be obtained without repeating full degradation experiments for every target battery[23], while inter-cell deep learning has demonstrated improved lifetime prediction across diverse ageing conditions and even across chemistries[24]. These results suggest that the right question is no longer whether transfer is possible, but how much mechanistic and statistical structure is needed to make transfer trustworthy.

Field data are now sufficiently rich to make this transition concrete. Open-source EV datasets have enabled multimodal SOH estimation from large numbers of vehicles, exposing both the opportunity and the complexity of pack-level health assessment[12]. At the same time, the historical lesson from early-life prediction is still important: deliberately designed datasets and informative early-cycle features remain powerful when paired with physically meaningful targets[25]. For digital twins, the implication is that cell-level diagnosis, pack-level heterogeneity management, and fleet-level learning should be treated as a single service continuum rather than as isolated modeling problems.

This is also where hybrid digital twins show practical strength. Fast-charging-specific twins can translate observed aging into mechanism-level interpretations[19], while mechanistically guided residual learning offers a route to continuous monitoring throughout life without repeated offline recalibration[20]. Taken together, these developments point to a health-management architecture in which lightweight onboard monitoring is periodically corrected by richer cloud-level inference, and both are anchored to mechanism-aware latent states rather than to a single empirical health score.

The notion of vector health is useful precisely because it resolves several ambiguities that have historically been hidden inside scalar SOH. Capacity loss does not distinguish between loss of lithium inventory and loss of active material. Resistance increase does not by itself identify whether the limiting process lies in charge transfer, ionic transport, contact loss, or heterogeneous utilization. Even the same numerical SOH can imply very different remaining safe charging windows, thermal sensitivities, and second-life values depending on how that SOH was reached. A digital twin can only support lifecycle decisions if its health representation preserves some of this mechanism-level structure[10,13,21,44].

This is why early-cycle information has regained importance. The value of early-cycle features is not merely that they allow earlier forecasting, but that they can encode intrinsic tendencies of a cell before heavy ageing obscures them. Recent work on generative and foundation-model approaches suggests that early signals can be embedded into richer trajectory models, potentially allowing the twin to update both its current health estimate and its expected future ageing path as new data arrive[20,25,44-46]. In practice, this would move health management from static diagnosis toward continual prognosis.

Generalization remains the decisive bottleneck. A model that performs well only on a single chemistry, protocol, or fleet is not a digital-twin solution but a local regression model. Real deployment requires transfer across manufacturers, climates, pack topologies, and charging habits. This is why domain adaptation, inter-cell learning, cross-condition benchmarking, and data-efficient uncertainty-aware learning are central rather than auxiliary topics[14,21,23-25,28,29,40-43]. A mature health twin should not merely survive domain shift; it should indicate when a new operating regime lies outside the reliable envelope of its prior experience.

The health-management literature is therefore converging toward a representation in which observables, latent states, and future trajectories are learned jointly rather than sequentially. Fragmented-charge capacity, multimodel fusion, degradation-mode decoupling, early-trajectory prediction, and cloud-assisted health updating all point in the same direction: battery health should be treated as a structured latent process whose meaning depends on transferability, uncertainty, and intervention relevance. This perspective also clarifies why health estimation, prognostics, and second-life grading should not be developed as isolated subfields, but as connected inference tasks sharing a common state representation[111-125].

Figure 5 translates this argument into health-management workflows in the order from observable feature construction to lifelong prognosis. Figure 5a shows multimodal feature engineering from EV field data, and Figure 5b embeds these features into a deep-learning SOH-estimation framework[12]. Figure 5c and Figure 5d connect a deployable power-autocorrelation indicator to capacity loss[14]. Figure 5e addresses available-capacity estimation under fragmented charging, where complete laboratory cycles are unavailable[29]. Figure 5f and Figure 5g emphasize that lifetime prediction must be evaluated across diverse ageing factors and capacity-fade trajectories[24]. Figure 5h extends health inference toward mechanistically guided residual learning for lifelong SOC/SOH correction[20], whereas Figure 5i illustrates early full-lifecycle prediction through BatteryGPT[43]. The figure therefore supports a vector-health view in which observables, latent mechanisms, transfer regime, and uncertainty are jointly relevant to decisions.

The shift from scalar SOH to vector health also changes how one should think about interpretability. In much of the battery literature, interpretability is treated as an attribute of the algorithm. In practice, the more consequential notion is interpretability of the state representation. A highly transparent model is of limited value if it estimates a quantity that is not aligned with the maintenance or charging decision. Conversely, a moderately complex model may still be useful if it yields a health representation that can be decomposed into physically meaningful limiting factors. For battery twins, interpretability therefore begins with what is estimated, not only with how it is estimated[5,13,21,44].

This has immediate implications for pack management. A battery pack does not fail gracefully according to the average cell. It is constrained by the least healthy cell, the hottest region, the most resistive connection, or the most uncertain safety margin. A twin that estimates only fleet-average or pack-average health may therefore underestimate operational risk. In contrast, a vector-health formulation can preserve pack dispersion, worst-cell margin, and uncertainty around both. This is essential when health outputs feed into balancing, charging derating, warranty assessment, or second-life grading[11,12,20,25,28].

Another underappreciated issue is that health estimation and prognosis operate on different evidence regimes. Current-state estimation can rely on short windows and local indicators; long-horizon prognosis requires assumptions about future operating conditions. This means that the most credible health twins will usually combine mechanistic current-state estimation with scenario-conditioned prognosis rather than extrapolate a single trend line indefinitely. The increasing use of generative sequence models, foundation models, and multimodal latent representations may improve this transition, but only if such models remain anchored to physically interpretable targets and calibrated uncertainty estimates[20,25,44-46,55-57].

Second-life management makes these issues even sharper. A battery entering repurposing does not arrive with a clean, uniformly sampled health trajectory. It arrives with missing context, uneven stress history, unknown micro-fault burden, and heterogeneous cell dispersion. A practical twin for second-life decisions therefore cannot rely only on nominal SOH labels. It must infer residual capability, uncertainty, and risk under incomplete information, ideally using representations that remain meaningful across the transition from first-life to second-life applications[10,28,37,54].

This is also why cloud assistance is likely to remain central to high-quality health management. Fleet-scale data provide the diversity needed to learn which features remain stable across vehicles, climates, and charging infrastructures; onboard systems provide the immediacy required for real-time adaptation. Neither alone is sufficient. The most useful health-management architectures are therefore likely to remain hybrid not only in the model sense, but also in the deployment sense: lightweight local monitoring, event-triggered uplink, and richer cloud-side inference that can return updated priors, thresholds, and control suggestions to the edge[11,12,20,28,37,38].

There is also a methodological lesson here for benchmark construction. Health models should not be compared exclusively at equal levels of target observability. Some methods assume complete charge curves; some assume only partial windows; some rely on impedance or pulse data; some infer trajectory from early cycles. A review that compares only endpoint RMSE without declaring these observability assumptions can mislead readers about what is actually deployable. In digital twins, the operational question is whether the assumed evidence stream exists in the intended application and whether it can be acquired repeatedly without unacceptable cost or disruption[7,12,14,21,25,28].

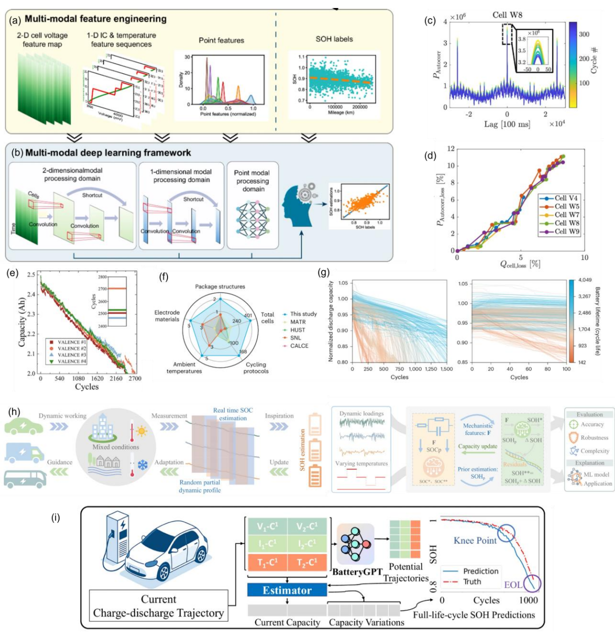

Figure 5. Digital-twin-enabled health management beyond scalar SOH. (a) Multimodal feature engineering from EV field data[12]; (b) Deep-learning SOH-estimation framework using 2D voltage, 1D sequence, and point-feature domains[12]; (c)

Power-autocorrelation profiles calculated from discharge data[14]; (d) Relation between power-autocorrelation loss and capacity loss[14]; (e) Available-capacity degradation under fragmented charging windows[29]; (f) Diverse ageing-factor coverage for inter-cell lifetime prediction[24]; (g) Long- and short-term degradation behaviors across ageing conditions[24]; (h) Mechanistically guided residual-learning pipeline for lifelong SOC/SOH monitoring[20]; (i) BatteryGPT pipeline linking early-cycle data to full-lifecycle SOH, knee point, and EOL prediction[43].

SAFETY DIAGNOSIS AND RISK-CONSTRAINED FAST CHARGING

Safety diagnosis and fast-charging control are the most stringent tests of whether a digital twin is truly decision-ready. Conventional BMS strategies rely heavily on threshold-based protection using voltage, current, and surface temperature. This is necessary but increasingly insufficient because the most dangerous transitions—lithium plating onset, local heat accumulation, gas evolution, or the approach to thermal runaway—develop internally and can remain weakly visible at the surface until late in the process[7,26].

Recent sensing studies show how a twin can move beyond this limitation. Differential pressure monitoring detects plating during fast charging[7], thermal-wave sensing differentiates degradation pathways in commercial cells[8], ultrasound visualizes plating evolution within multilayer pouch cells[9], and MHz-band electromagnetics reveals cycle-by-cycle changes associated with metallic lithium deposition[10]. In parallel, lab-on-fiber measurements demonstrate that internal temperature and pressure signals can provide early-warning signatures of thermal runaway before traditional external indicators become decisive[26]. These results do not eliminate the need for modeling; instead, they create the observability needed for more informative latent safety states.

This motivates a transition from threshold alarms to state-of-safety estimation. A safety-oriented twin should infer a time-varying hazard state that integrates thermal, electrochemical, and mechanical evidence rather than issuing binary warnings only after thresholds are crossed. Recent state-of-safety formulations move in this direction by combining multiple observables into a unified early-warning indicator and quantifying warning lead time under abuse conditions[27]. The practical value of such a formulation is that uncertainty can be translated into conservative control before irreversible failure develops.

In fast charging, this latent-state perspective becomes unavoidable. The fundamental control problem is not to force the battery through a fixed current profile as quickly as possible, but to exploit whatever fast-charging freedom remains after the current health state, temperature field, and risk margin are taken into account. Put differently, the safe charging envelope is not static. It contracts or expands with age, heterogeneity, environment, and model confidence. The role of a digital twin is to infer this envelope online and update it continuously rather than treat it as a fixed engineering constant[6,9,10,13,19,26,39].

Fast charging should be interpreted within the same framework. Closed-loop protocol optimization proved that data-driven exploration can identify better charging policies more efficiently than brute-force experimentation[6]. However, a digital-twin perspective pushes the objective further: the output should not be a universally optimal protocol, but a safe and context-specific action set conditioned on present health, temperature, heterogeneity, and predictive confidence. Digital-twin-assisted fast-charging studies show that this is feasible because charging actions can be linked directly to changes in dominant aging modes rather than only to terminal capacity retention[19].

A useful way to express this distinction is through the difference between protocol optimization and action-space regulation. Protocol optimization seeks a globally good charging schedule under assumed conditions. Digital-twin charging seeks the set of locally admissible actions under current inferred conditions. This shift matters because a protocol that is safe for a fresh cell at 25 °C may not be safe for an aged cell, a pack with a strong thermal gradient, or a vehicle after repeated high-rate operation. Risk-constrained fast charging is therefore inseparable from state estimation, uncertainty quantification, and pack-level supervision[6-10,19,26,39,82-89].

Fast-charging safety is also inseparable from chemistry. The charging window is not only determined by current, temperature, and state of charge, but by the interphase chemistry that governs charge transfer, lithium transport, gas evolution, and plating reversibility. This is why recent fast-charging advances increasingly couple control with electrolyte design, interphase characterization, and electrode-level potential sensing[7-10,39,68,86-89]. A battery digital twin that ignores this chemistry layer may still regulate current successfully in the short term, but it will remain blind to why the safe envelope changes over repeated cycling.

For safety diagnosis more broadly, the most important conceptual shift is from event detection to state-of-safety estimation. Batteries do not move from 'safe' to 'unsafe' in a single physical step. They transition through a sequence of latent precursor states characterized by subtle changes in internal temperature, pressure, stress, gas evolution, impedance, and local current distribution. A warning framework that only reacts after voltage collapse or sharp temperature rise has already surrendered much of its intervention window. A state-of-safety framework instead asks how evidence should be fused before catastrophic failure is visible at the pack surface[13,18,26,27,35,50-60].

This transition is especially important at pack and fleet scales. The same local cell event may remain weakly visible in aggregate pack telemetry until it is amplified by thermal propagation or balancing constraints. Conversely, fleet-level monitoring can reveal rare-event signatures and owner-specific risk patterns that are invisible in single-vehicle analysis. A complete safety twin therefore spans cell-level precursors, pack-level escalation, and fleet-level collaborative learning. Its output is not merely a warning flag, but a graded hazard state that can support derating, charging reconfiguration, thermal conditioning, or shutdown[12,18,23,27].

At larger scales, safety intelligence also becomes a distributed learning problem. Privacy-preserving collaborative fault warning demonstrates that fleet-level safety models can be trained from heterogeneous charging-station data without centralizing raw records[28]. Meanwhile, fragmented charge-capacity estimation shows that useful capacity-related information can still be recovered even when full charge traces are unavailable[29]. Together, these studies imply that safety and fast-charging twins must be designed for limited onboard observability, fragmented data access, and privacy constraints from the outset, not retrofitted to them after laboratory success.

A second unresolved issue in this section is how safety-relevant information should be prioritized when multiple sensing modalities are available. The emerging battery-safety literature increasingly shows that current, voltage, temperature, pressure, gas, strain, and embedded thermal signals are not interchangeable; rather, they resolve different precursors at different stages of hazard escalation. This is especially important for pack applications, where aggregate telemetry can mask local precursor states. Recent work on safety detection technology, multiphysics warning, smart BMS architectures, and AI-assisted thermal management suggests that the next generation of safety twins will be judged not only by early-warning lead time, but also by their ability to map specific evidence streams onto intervention classes such as derating, reconfiguration, isolation, or shutdown[126-145].

Figure 6 and Figure 7 distinguish the charging and safety faces of the same latent-state problem. Figure 6a and Figure 6b illustrate how negative-electrode potential and Ti-O covalency can expand the materials window for fast charging[44]. Figure 6c, Figure 6d, Figure 6e, and Figure 6f show pressure-triggered dynamic current regulation and graphite-anode evidence for suppressing lithium plating[7]. Figure 6g and Figure 6h show post-plating manipulation through electric-field relaxation[45], while Figure 6i and Figure 6j connect chemistry-enabled fast charging to CC/CV-stage behavior and cycling retention[46]. Figure 7 then shifts from charging control to hazard-state estimation. Figure 7a and Figure 7b define the lab-on-fiber measurement platform, Figure 7c and Figure 7d show internal temperature/pressure traces and thermal-runaway event sequence, and Figure 7e links these measurements to mechanism decoding and derivative-based warning before venting[26]. Figure 7f, Figure 7g, and Figure 7h introduce strain-temperature-voltage evidence, electrode-level damage, and state-of-safety trajectories[27], while Figure 7i adds heat and mass-output variability under thermal runaway[47]. Together, Figure 6 and Figure 7 show why fast charging and safety should be treated as coupled control problems rather than as separate threshold rules.

From a sensing perspective, the central safety challenge is that many precursors are local while most deployed signals are aggregate. Internal pressure can rise before surface temperature. A local hotspot can form before pack-level temperature appears abnormal. Lithium plating can begin in a limited region before mean voltage or current reveals anything unusual. This mismatch is one reason why a purely threshold-based BMS frequently detects late. It also explains the growing interest in richer detection modalities such as optical fibers, strain sensing, gas detection, ultrasound, lock-in thermography, magnetic-field imaging, and operando spectroscopy[26,27,39,50,58-81]. A digital twin does not replace these modalities; it provides the latent-state framework in which their signals become collectively interpretable.

Battery safety also unfolds across multiple timescales. Some hazards develop over months through progressive degradation and fault accumulation. Others, such as sudden short-circuit-triggered escalation, evolve within seconds or minutes. A mature digital twin should therefore separate long-horizon hazard conditioning from short-horizon hazard monitoring. The former depends on degradation state, installed environment, and operating history; the latter depends on fast-evolving thermal, mechanical, and electrochemical evidence. Conflating these layers can cause either false confidence or false alarms. This multi-timescale distinction is especially important for large battery packs in which ancillary equipment, pack topology, and thermal management subsystems can influence the final hazard outcome as strongly as the cell itself[13,35,42,45,50,58].

For fast charging, the analogous distinction is between feasibility and aggressiveness. A charge action can be electrochemically feasible yet operationally unwise if its uncertainty margin is too small, if thermal nonuniformity is increasing, or if the pack contains an outlier cell approaching a plating boundary. Conversely, a conservative charge action can still be suboptimal if the twin underestimates the available safety margin because it lacks sufficient observability. The practical task is therefore not simply to minimize charging time, but to maximize useful charging rate under explicit uncertainty-aware safety constraints. This is the setting in which digital twins become more valuable than fixed protocols or purely reactive controls[6-10,19,26,39,82-89].

The broader safety-detection literature reinforces this conclusion. Recent high-level reviews emphasize that large battery packs introduce a distinct safety regime relative to

DOI: 10.20517/scierxiv202605.0340.v1

single cells because failure propagation, auxiliary-equipment coupling, and thermal management complexity become first-order variables. The practical difficulty is that the signals available to the BMS are often aggregate measurements, whereas the relevant precursors may arise at the level of a single cell, tab, separator region, or cooling defect. This mismatch explains why safety detection in packs cannot simply be extrapolated from cell-level fault detection, and why digital twins are attractive as integrative layers for reconciling multi-source evidence across scales[13,35,50,58-74].

Recent work also underscores that safety monitoring is expanding beyond voltage-current-temperature triplets. Mechanical stress and swelling, gas composition, acoustic signatures, ultrasonic echoes, X-ray or tomography-derived structural information, optical spectroscopy, and EIS-based thermal or interfacial signatures all offer orthogonal safety information under different constraints. The scientific challenge is no longer only sensor invention; it is how to convert heterogeneous signals into a coherent latent hazard state. This is a quintessential digital-twin problem because it requires model-based fusion, uncertainty-aware interpretation, and state-dependent action selection rather than one-to-one threshold mapping[26,27,39,50,63-81].

From this perspective, safety and fast charging become two coupled supervision problems. Fast charging asks how aggressively one may act under current hidden-state uncertainty. Safety asks how early a hazard state can be inferred from imperfect evidence. In both cases, the most relevant outputs are not raw measurements but decision variables: maximum admissible current, preheating request, fault-severity class, safety margin, or shutdown trigger. Battery twins matter because they allow these outputs to be conditioned simultaneously on mechanism, measurement, and uncertainty rather than on fixed thresholds alone.

Pack propagation deserves particular attention. A large fraction of the safety literature concerns how single-cell abnormalities trigger local runaway, but practical electrification increasingly depends on whether such events remain local, propagate to neighboring cells, or are amplified by module enclosure, gas pathways, and pack topology. The digital twin therefore needs a state representation that can move across scales: from cell-level precursors, to module-level coupling, to pack-level escalation

DOI: 10.20517/scierxiv202605.0340.v1

probability. This cross-scale hazard representation is still underdeveloped, and its construction remains one of the clearest opportunities for future work[13,18,23,24,35,50-59].

There is a complementary issue on the control side. A warning that a hazard state is rising does not specify what intervention should be taken. Depending on time scale, the correct response may be current reduction, thermal conditioning, balancing, charging interruption, derating, cell isolation, or shutdown. An effective safety twin should therefore map inferred hazard states to intervention classes, and not only to alarms. This is another reason why state-of-safety formulations are more useful than threshold alarms alone: they are inherently closer to the control problem that the BMS must solve[27,39,50,58].

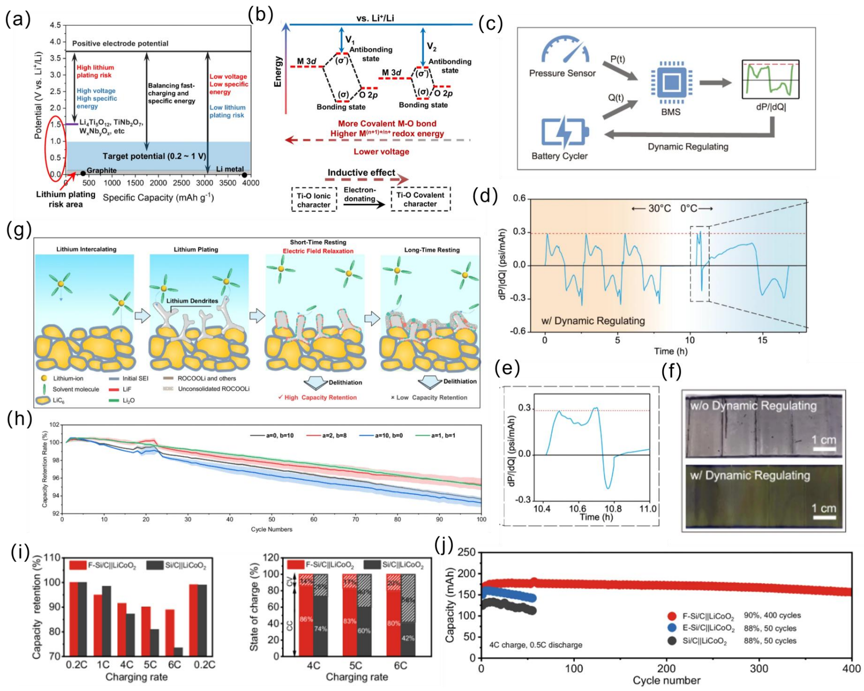

Figure 6. Fast-charging supervision as risk-constrained control. (a) Materials-level potential window for balancing fast charging and specific energy[44]; (b) Ti-O covalency strategy for regulating lithium-insertion potential[44]; (c) Pressure-triggered self-regulated charging scheme[7]; (d) Dynamic current regulation at 30 °C[7]. (e) Dynamic current regulation at 0 °C[7]; (f) Graphite-anode optical evidence showing suppression of lithium plating[7]; (g) Electric-field-relaxation mechanism for lithium-dendrite manipulation[45]; (h) Degradation response of commercial Gr||LFP cells under 3 C plating-prone fast charging and different manipulation protocols[45]; (i)

DOI: 10.20517/scierxiv202605.0340.v1

Capacity retention and CC/CV-stage segmentation under chemistry-enabled fast charging[46]; (j) Pouch-cell cycling stability under 4 C charging[46].

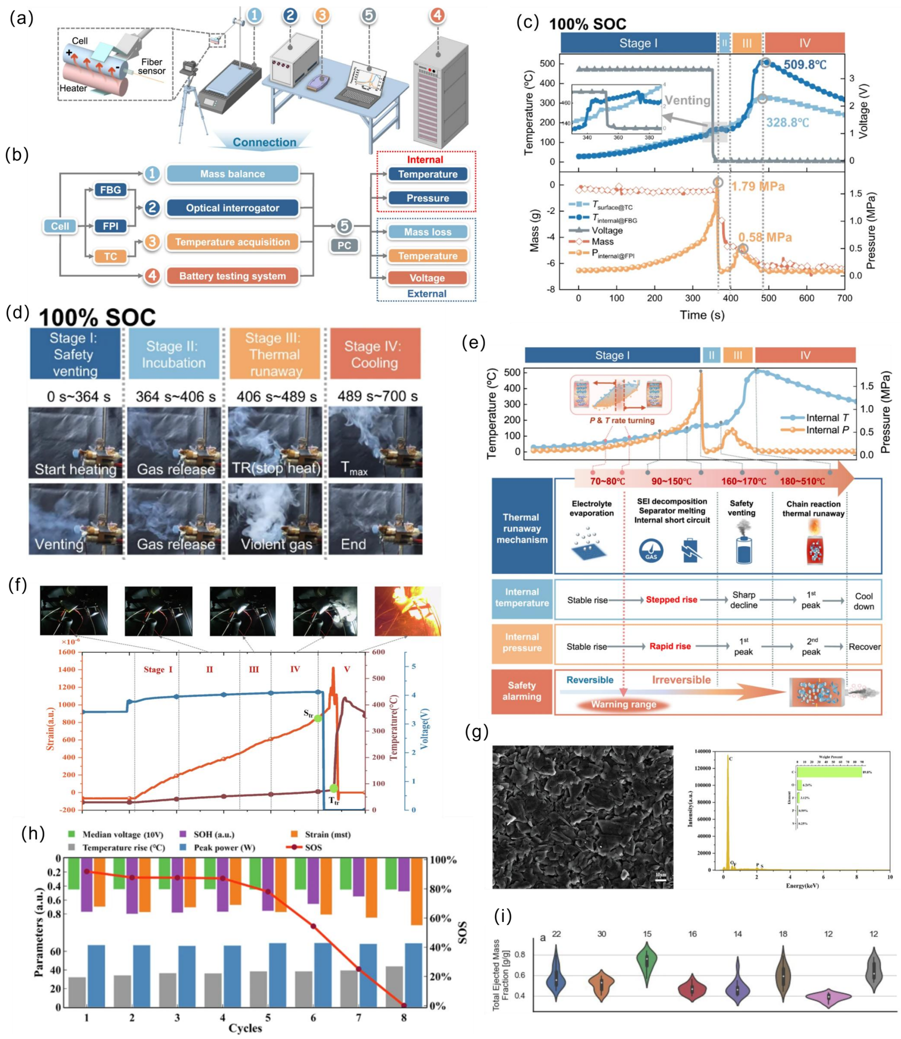

Figure 7. Safety diagnosis, early warning, and state-of-safety estimation. (a) Lab-on-fiber experimental platform for internal temperature and pressure monitoring[26]. (b) Measurement logic and functions of the experimental setup[26]. (c) Internal temperature/pressure traces, surface temperature, voltage, and mass evolution during thermal runaway[26]. (d) Thermal-runaway event sequence in 18,650 cells[26]. (e) Mechanistic decoding of thermal runaway and derivative-based early warning before venting[26]. (f) Strain, temperature, and voltage evolution under abuse[27]. (g) Negative-electrode morphology and composition after runaway[27]. (h) State-of-safety quantitative assessment and trajectory[27]. (i) Distribution of heat and mass-output variability under thermal runaway[47].

Table 4 aligns health estimation, prognosis, fast-charging supervision, safety diagnosis, and lifecycle governance by their outputs, dominant uncertainty sources, decision relevance, and evaluation metrics.

Table 4. Task classes, outputs, uncertainty sources, and decision metrics.

Task	Primary outputs	Dominant uncertainty sources	Decision relevance	Recommended metrics
Health estimation	SOH, vector health state, degradation modes	Partial cycles, temperature variation, pack heterogeneity	Maintenance, warranty, second life, charging limits	RMSE/MAE + calibration + robustness
Health prognosis	Future SOH, knee point, EOL, RUL	Future usage uncertainty, model-form uncertainty	Maintenance planning, lifecycle value	Trajectory error, NLL, prediction-interval quality
Fast-charging supervision	Admissible charging current, plating margin, thermal margin	Latent plating threshold, thermal nonuniformity, state-estimation error	Charging control, preheating, derating	Time-to-target + safety-margin violation rate + calibration

Task	Primary outputs	Dominant uncertainty sources	Decision relevance	Recommended metrics
Safety diagnosis	State of safety, warning lead time, escalation class	Rare events, incomplete observability, pack propagation	Alarm, shutdown, isolation, emergency response	Lead time, false-alarm rate, missed-event rate, robustness
Deployment and governance	Personalized twin update, battery passport fields, lifecycle traceability	Semantic inconsistency, missing metadata, domain shift	Auditability, compliance, repurposing, recycling	Latency, interoperability, reproducibility, auditability

Deployment Realism, Benchmarking, and Data Governance

The deployment challenge is often underestimated. In practice, a battery digital twin is not deployed once; it is sustained across sensing, logging, synchronization, update, and decision layers. Vehicle batteries differ in initial grading, cooling conditions, aging history, and operational patterns, so a deployable twin must explicitly manage cell inconsistency, worst-cell limitation, and pack-level thermal heterogeneity[11,12]. This favors a layered architecture in which the pack is not collapsed into a single lumped state unless the information loss is acceptable for the intended decision.

Deployment changes the criteria by which a battery digital twin should be judged. In laboratory studies, the implicit objective is often predictive fidelity under controlled conditions. In field operation, the objective is usually narrower and more consequential: maintain safe operation, preserve useful life, and keep the model synchronized with incomplete, noisy, and partially labeled data. This means that deployability depends not only on model accuracy, but also on data compression, synchronization cadence, event-triggered updating, edge-side computation, and the ability to distinguish which information must remain local and which can be learned at fleet scale[11,12,28,37,38].

This in turn reframes the value of cloud-edge-fleet architectures. Their purpose is not simply to centralize data, but to distribute intelligence by time scale and by decision type. Local filtering and immediate protection remain naturally onboard. Heavy health diagnosis, drift detection, personalization, and fleet-level anomaly learning are more naturally cloud-mediated. The edge layer becomes important whenever synchronization latency, privacy, or bandwidth constraints make raw data transfer undesirable. The most useful architecture is therefore not the most centralized one, but the one that minimizes information loss while respecting the physical urgency of the decision[12,25-28,37,38,47].

Benchmarking must follow the same realism. Many current datasets and evaluations still emphasize narrow laboratory cycling, whereas digital twins are meant to operate across protocol changes, temperature shifts, pack configurations, and field data with scarce labels[5,12,24]. Benchmark design should therefore separate interpolation from extrapolation, laboratory from field transfer, cell from pack configuration shift, and health prediction from safety warning. Without this separation, apparently strong headline accuracy can hide weak deployment robustness.

Evaluation metrics should also move beyond point-error reporting. Probabilistic battery reviews argue that calibration-oriented metrics such as negative log-likelihood, expected calibration error, and sparsification-based measures are necessary whenever predictions are used to guide control[5]. For safety tasks, warning lead time, false-alarm rate, and missed-event rate are just as important as classification accuracy[27]. Open-access resources such as the Battery Failure Databank can play an important role here because they offer a common basis for benchmarking rare but consequential thermal-runaway behaviors[30].

A second prerequisite for reproducible evaluation is better semantic and regulatory infrastructure. Vehicle-level health measurement itself is not yet standardized, and recent work has argued for repeatable onboard procedures for energy- and capacity-based SOH at the pack level[31]. At the same time, the earlier data-centric lesson from battery research still holds: broad and diverse datasets become much more

DOI: 10.20517/scierxiv202605.0340.v1

useful when they are organized around meaningful early descriptors and clearly defined targets[25]. A robust benchmark therefore needs both shared data and shared measurement definitions.

These concerns now extend into lifecycle governance. Unified battery data descriptions[16] and multiscale digital modeling across the full lifecycle[17] provide the semantic backbone needed to connect manufacturing genealogy, testing history, operational data, and end-of-life decisions. This is increasingly important under the EU Battery Regulation, which requires battery passports for EV batteries, light means of transport batteries, and industrial batteries above 2 kWh from 18 February 2027[32]. In that context, the digital twin becomes more than an internal engineering asset; it becomes part of the auditable digital thread that links performance, safety, value retention, repurposing, and recycling.

Benchmark design deserves separate emphasis because the field is still too tolerant of accuracy numbers generated under unrealistically narrow conditions. A battery digital twin should be benchmarked along at least four dimensions: deployment realism, system scale, task type, and uncertainty quality. Deployment realism asks whether the data originate from controlled cycling, semi-realistic duty cycles, or genuine field operation. System scale asks whether the inference problem is at the material/interface, cell, pack, or fleet level. Task type distinguishes health estimation, degradation diagnosis, safety warning, and action recommendation. Uncertainty quality asks whether the model is calibrated, robust, and explicit about domain shift. Without this four-dimensional benchmark frame, strong headline accuracy can hide weak operational usefulness[5,12,13,18,21,23-25,27,31,35].

Data governance is not an administrative afterthought to this problem; it is part of the scientific architecture. Once the digital twin is deployed beyond a single cell test stand, battery identity, provenance, manufacturing history, beginning-of-life characterization, field usage, maintenance events, second-life grading, and end-of-life processing all become causally relevant. Without a semantically unified description of these records, the twin loses continuity across lifecycle stages. Battery ontology, battery passport, and lifecycle digital-thread concepts are therefore not simply compliance tools; they are the information backbone required to convert a sequence of disconnected models into a deployable and auditable battery-intelligence system[16,17,31,32,35,74,82,83].

The governance problem also has a scientific dimension. Standardization, pack-level measurement procedures, and semantic continuity determine which deployment claims are actually testable across institutions, fleets, and lifecycle stages. A benchmark that ignores privacy constraints, owner heterogeneity, measurement protocol, or regulatory fields may still be publishable, but it is not deployable. Recent work on embedded sensing, gas and strain detection, standardized SOH measurement, smart BMS design, machine-learning-assisted thermal management, and federated estimation under privacy constraints reinforces the need to treat deployment realism, auditability, and semantic interoperability as first-class scientific requirements rather than as downstream engineering details[146-160].

Figure 8 and Figure 9 close the deployment and benchmarking argument in sequential order. Figure 8a and Figure 8b show privacy-preserving collaboration for retired-battery sorting without raw-data exchange[38]. Figure 8c and Figure 8d expose battery-type heterogeneity across charging-station owners, and Figure 8e shows how a personalized federated fault-warning architecture handles that heterogeneity[28]. Figure 8f illustrates second-life SOH estimation under random retirement conditions and deployment without extra SOC conditioning[48]. Figure 8g adds the lifecycle digital-thread layer that connects battery identity, provenance, health, compliance, repurposing, and recycling. Figure 9 then shifts to benchmark realism: Figure 9a and Figure 9b show real-world DC fast-charging capacity distribution and temperature-dependent charging-profile heterogeneity[49]; Figure 9c and Figure 9d address standardized pack-level SOH measurement and differential-voltage transfer from cell to vehicle level[31,50]; Figure 9e and Figure 9f show that evaluation should include realistic data flow, data-frame heterogeneity, ROC curves, and reconstruction error rather than point accuracy alone[51]; Figure 9g adds pack-voltage variability under fixed voltage windows[31]; and Figure 9h summarizes the benchmark matrix and roadmap for deployable twins. Together, Figure 8 and Figure 9 show that deployment and benchmarking are not external to the digital-twin problem; they define its engineering meaning.

Dataset design is central to this transition. A dataset that is suitable for academic model comparison is not automatically suitable for battery-twin deployment. Deployment-oriented datasets must preserve context: battery identity, chemistry, pack topology, environment, sampling cadence, control history, maintenance events, and the provenance of labels. Otherwise, two apparently identical voltage traces may correspond to different lifecycle phases or different risk regimes. This is why field datasets, battery-failure databases, and real-world charging archives are beginning to matter as much as classical laboratory cycling datasets in this literature[12,13,18,21,23,25,28,30,61].

The notion of benchmark realism should therefore be broadened beyond data source. Realism also concerns the perturbations that the benchmark allows. Does it test adaptation to temperature shift? To pack scaling? To partial charging? To rare-event imbalance? To missing labels? To changing decision cost? In battery digital twins, a benchmark that does not expose the model to these distortions may be scientifically informative yet operationally misleading. We therefore argue that benchmark design should distinguish at least four regimes: interpolation within a known laboratory domain, extrapolation across known but unseen conditions, out-of-domain transfer across chemistries or configurations, and in-the-wild deployment under fragmented data and uncertain labels. These regimes should not be collapsed into a single accuracy number[5,13,23-25,31,35].

Regulation further sharpens these requirements. Battery passports, traceability mandates, and lifecycle reporting obligations introduce an additional meaning of correctness: the twin must not only predict well, but also support standardized, explainable, and auditable decisions. This is particularly relevant for health-based warranty, second-life valuation, and recycling readiness, where the decision output of the twin may have commercial and regulatory consequences. A battery-management framework that is opaque, weakly calibrated, or semantically inconsistent across lifecycle phases will struggle in such environments even if its cell-level accuracy is high[12,16,17,31,32,35,37].

Another practical issue is that large-scale deployment requires explicit coexistence of model diversity. The same fleet may contain different chemistries, form factors, pack topologies, software versions, and sensor layouts. A digital-twin architecture that assumes one model per fleet or one metric per chemistry will quickly become brittle. More realistic architectures will likely require shared semantic layers, chemistry-aware priors, modular models, and benchmark suites that declare clearly which forms of transfer are being attempted. In that sense, interoperability is not only a software-engineering concern; it is a scientific requirement for any twin that claims lifecycle usefulness[12,16,17,25,31,32,37].

A final deployment consideration concerns auditability. If the twin is used only for internal analysis, modest opacity may be tolerable. If it influences warranty decisions, insurance risk, second-life valuation, charging limits, or safety interventions, then the rationale for its outputs must be inspectable. This does not mean that every subsystem must be fully transparent, but it does mean that inputs, assumptions, uncertainty, and action logic must remain traceable. For battery passports and lifecycle governance, this traceability is likely to be as important as the numerical output itself. This is one reason why the future battery twin is better understood as an auditable intelligence layer than as a single predictive model[12,31,32,35,37,82,83].

Data resources are likely to determine the pace of progress as much as any individual algorithm. The field now has access to open EV datasets, retired-battery datasets, battery-failure databases, and increasingly rich lab datasets with impedance, imaging, and operando sensing. Yet these resources are rarely aligned semantically. Different datasets define health differently, record metadata unevenly, and expose different slices of the lifecycle. A major opportunity for the next phase of digital-twin research is therefore not only more data, but better connected data: field records that can be traced to manufacturing descriptors, lab diagnostics that can be linked to pack operation, and failure data that can be interpreted in light of prior degradation history[12,13,25,30,31,35,37,61].

This is also why future benchmark initiatives should not separate data quality from governance. A benchmark that is impossible to reproduce, impossible to interpret semantically, or incompatible with lifecycle traceability will not support industrial translation even if it produces informative machine-learning leaderboards. Strong benchmark design in this field therefore requires both scientific and infrastructural discipline: explicit metadata, declared label generation procedures, unambiguous task definitions, standardized evaluation scripts, and clarity about whether the benchmark is laboratory, semi-realistic, or field-based. In our view, these requirements are no longer optional for high-quality battery-digital-twin research[21,23,25,31,32].

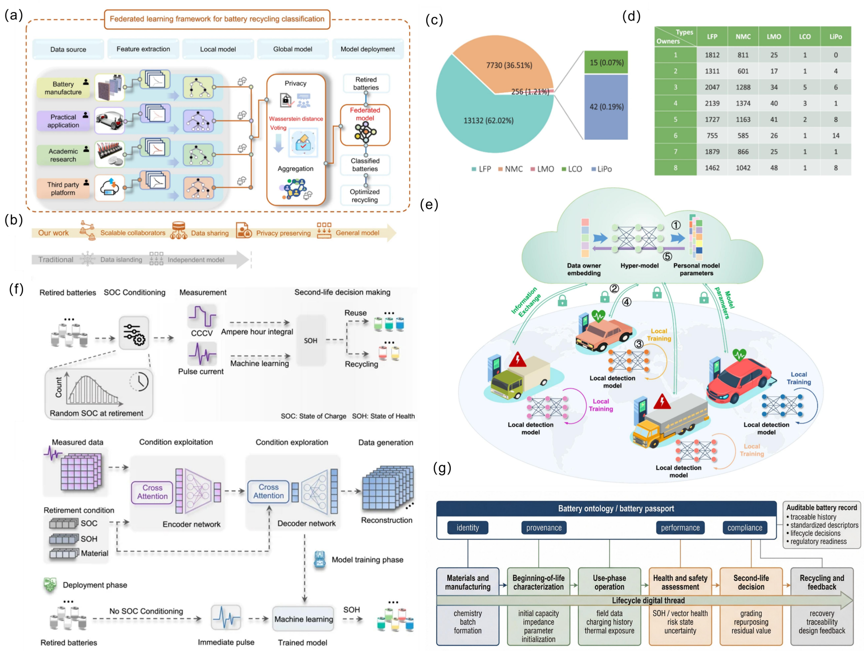

Figure 8. Deployment, privacy-preserving collaboration, and lifecycle digital thread. (a) Federated retired-battery sorting using distributed data contributors without raw-data exchange[38]; (b) Privacy-preserving collaboration compared with the conventional data-island paradigm[38]; (c) Overall battery-type distribution in heterogeneous charging-station data[28]; (d) Owner-specific battery-type distribution[28]; (e) Personalized federated fault-warning architecture across charging-station owners[28]; (f) Generative-learning-assisted retired-battery SOH estimation under random retirement conditions, including pretreatment, pulse-response generation, and deployment without additional SOC conditioning[48]; (g) Battery-passport, ontology, and lifecycle digital-thread layer connecting identity, provenance, performance, compliance, second-life use, and recycling.

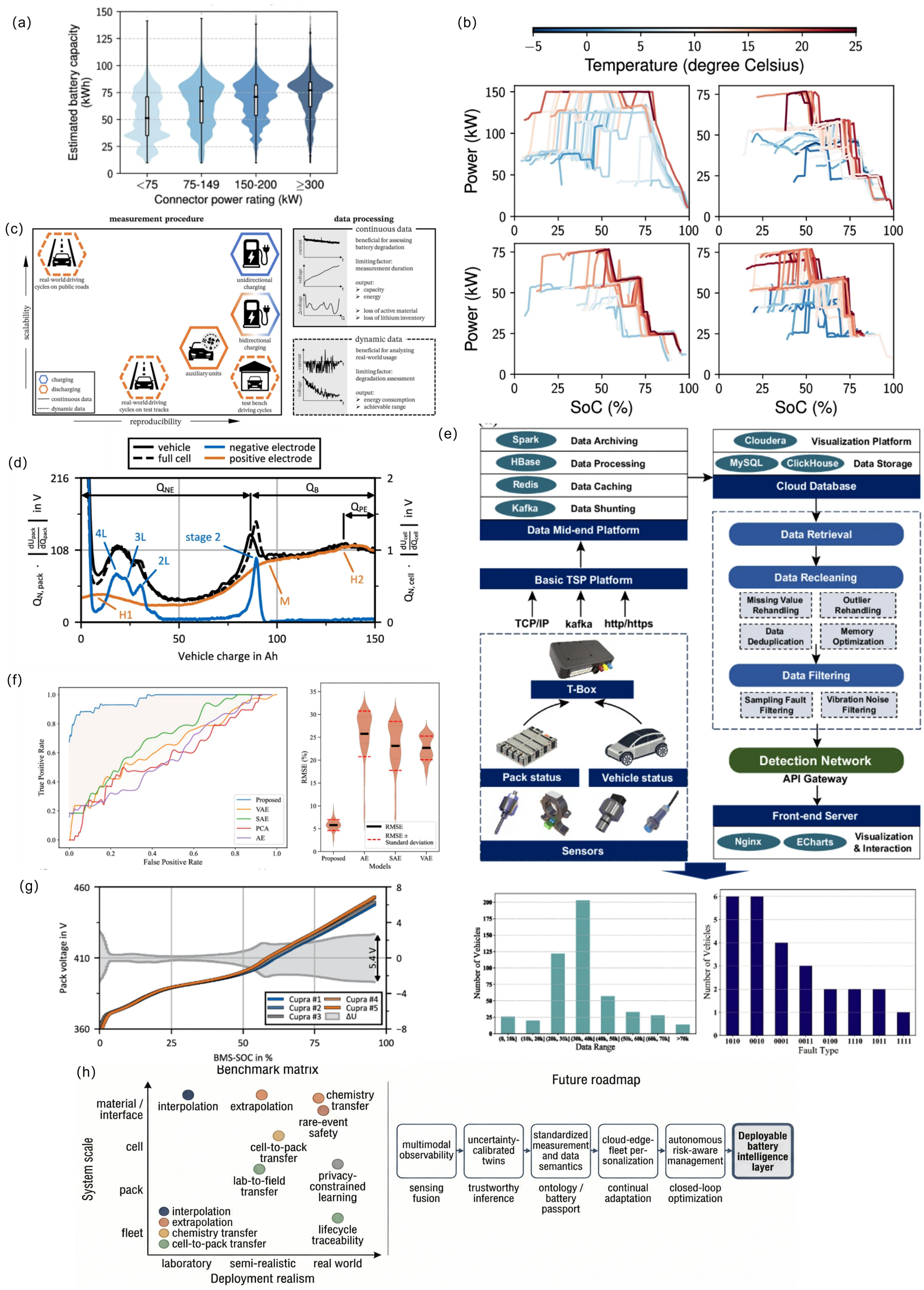

Figure 9. Benchmarking, standardization, and future directions. (a) Real-world DC fast-charging capacity distribution across connector power ratings[49]; (b) Temperature-dependent real-world charging-profile heterogeneity[49]; (c) Vehicle-level SOH standardization based on scalable and reproducible measurement procedures[31]; (d) Differential-voltage transfer from cell-level diagnostics to pack-level EV measurement[31,50]; (e) Online fault-diagnosis data flow and field dataset structure under stochastic operating conditions[51]; (f) ROC and RMSE-based evaluation of fault detection and prediction accuracy[51]; (g) Pack-voltage variability under fixed voltage windows, illustrating limits of nonstandardized SOH comparison[31]; (h) Benchmark

DOI: 10.20517/scierxiv202605.0340.v1

matrix and roadmap for deployable battery twins, spanning deployment realism, system scale, calibration, robustness, latency, and auditability.

Finally, the maturation of battery digital twins will depend on whether the community can align scientific rigor with industrial traceability. A high-performing health model is not sufficient if its labels are not portable across platforms, if its uncertainty is uncalibrated, or if its outputs cannot be audited at the pack, fleet, or regulatory level. Conversely, a robust digital thread without mechanism-aware inference will still fail to deliver actionable intelligence. The next phase of progress therefore lies at the intersection of electrochemical understanding, reliable data semantics, lifecycle governance, and computational decision theory. Only by treating these as coupled requirements, rather than as separate research tracks, can battery digital twins become a stable technical foundation for fast charging, predictive safety, warranty management, repurposing, and circular battery value chains.

An equally important open issue concerns the coupling between management decisions and information acquisition. Most battery studies still assume that sensing is passive and fixed, but practical digital twins should also decide when to probe the system more aggressively. A short rest period, a diagnostic pulse, a low-amplitude impedance measurement, or a temporary charging-rate modulation may all be interpreted as information-seeking actions, because they improve the observability of hidden states such as plating tendency, resistance growth, or pack imbalance. In this sense, battery management is not only an optimization problem over energy and power, but also an optimization problem over uncertainty reduction. Future twins will likely combine passive monitoring with selective active interrogation, thereby coupling estimation, control, and diagnostic experiment design within the same closed-loop framework. This shift from passive prediction to active state clarification is particularly relevant for safety-critical operation, where uncertainty itself is often the variable that must be managed.

A further conceptual challenge is that the twin must remain identifiable under changing evidence regimes. In early life, electrochemical signatures are information-rich but the damage state is small; in mid-life, degradation modes become separable but operating histories diverge; late in life, pack constraints, thermal asymmetry, and cell dispersion dominate the practical risk picture. A deployable twin should therefore not be calibrated once and then only updated numerically. It should adapt its own reliance on model classes, latent states, and observables as the battery ages. For example, a newly commissioned pack may benefit from impedance-informed initialization and formation-aware priors, whereas an aged field pack may rely more on uncertainty-aware residual learning, abnormality scoring, and conservative control thresholds. This age-dependent reweighting of evidence is rarely made explicit in current literature, yet it is essential if digital twins are to remain scientifically interpretable and operationally useful across the full battery lifecycle.

OUTLOOK: TOWARD TRUSTWORTHY AND AUTONOMOUS BATTERY TWINS

The preceding sections suggest that progress in battery digital twins will be limited less by isolated algorithmic accuracy than by whether the field can construct state representations that are observable, identifiable, calibrated, and useful for intervention. Three questions should therefore be addressed together: what the twin can see, what physical state it represents, and how uncertainty is converted into action. Treating sensing, modeling, control, and governance as separate topics is no longer sufficient for deployable battery intelligence.

The first priority is active and economical observability. Future twins will not infer plating, thermal-runaway precursors, degradation-mode evolution, or pack imbalance reliably from voltage, current, and surface temperature alone. Pressure, strain, thermal-wave, ultrasound, magnetic, optical-fiber, gas, and impedance measurements provide complementary information, but each adds cost, packaging constraints, calibration burden, and reliability risk. The central issue is therefore which sensing combination maximizes hidden-state identifiability for a specific decision and deployment context[7-10,15,58-74].

The second priority is state-sufficient multiphysics modeling. A useful twin does not need to reproduce every microscopic detail of a cell; it must retain the internal variables needed for the intended decision. Cross-scale electrochemical-mechanical models, electrothermal-aging models, mechanistically guided residual learners, and physics-informed neural networks all contribute to this goal when their state variables are aligned with management tasks[4,17-21,39]. The open problem is to define, for each task, which latent coordinates are sufficient: degradation modes for health management, plating and thermal margins for charging, state of safety for hazard mitigation, and traceable descriptors for lifecycle governance.

The third priority is calibrated uncertainty under distribution shift. Field operation changes the evidence regime continuously. Early-life data are information-rich but weakly degraded, mid-life trajectories are mechanism-separating but user-dependent, and late-life operation is dominated by pack dispersion, thermal asymmetry, and risk concentration. A deployable twin should therefore adapt not only its parameters but also its reliance on model classes, observables, and priors as the battery ages. This age-dependent reweighting is rarely explicit in current literature, yet it is essential for maintaining interpretability across the full lifecycle[5,11,12,20,24,43].

The fourth priority is benchmark realism. Battery twins should be tested across distinct regimes: interpolation within known laboratory distributions, extrapolation across controlled but unseen operating conditions, transfer across chemistry or configuration, and in-the-wild operation under fragmented data and uncertain labels. Reporting a single point-error metric across these regimes is not informative. Evaluation should include calibration, warning lead time, robustness to missing data, action-safety violations, latency, and auditability whenever the twin is used for charging, safety, maintenance, or lifecycle decisions[5,24,27,30,31].

The fifth priority is semantic and regulatory continuity. Battery identity, manufacturing genealogy, beginning-of-life characterization, field operation, maintenance events, safety incidents, second-life grading, and recycling decisions are not peripheral metadata once a twin is deployed beyond the laboratory. They determine whether health labels are portable, whether uncertainty is interpretable, and whether decisions can be audited. Battery ontology, battery passports, and lifecycle digital threads therefore form part of the scientific infrastructure of deployable twins rather than an external compliance layer[16,17,31,32].

DOI: 10.20517/scierxiv202605.0340.v1

A final direction is cautious autonomy. Closed-loop optimization and reinforcement learning can move battery management from monitoring to supervisory action, but autonomy is credible only when the action policy remains constrained by physical state, calibrated uncertainty, and traceable decision logic. The objective is not to replace electrochemical understanding with end-to-end automation. It is to use learning to determine when measurements, mechanisms, and uncertainty are sufficient for a safe action, when additional probing is required, and when conservative intervention is necessary[6,33,34,168].

From this perspective, the long-term value of digital twins lies in reorganizing battery management around state sufficiency, uncertainty calibration, and lifecycle accountability. If this reorganization succeeds, apparently separate functions—state estimation, degradation diagnosis, fast-charging supervision, fault warning, second-life grading, and regulatory traceability—become different queries to the same evolving representation of the battery system.

CONCLUSION

This Review argues that lithium-ion battery health management, safety diagnosis, and fast-charging control should not be treated as independent software functions. They are different decision faces of the same coupled electrochemical, thermal, mechanical, and interfacial degradation system. A digital twin that only predicts scalar SOH or reproduces terminal voltage is therefore insufficient for modern battery management because it does not preserve the mechanism-to-action link required for safe operation.

A deployable battery digital twin must be physics-grounded, uncertainty-calibrated, and decision-oriented. It should fuse multimodal observability with task-matched mechanistic models, infer hidden degradation and hazard states with calibrated uncertainty, and translate those states into control-relevant quantities such as charging margins, safety warnings, maintenance priorities, and lifecycle decisions. This is a stricter requirement than high prediction accuracy under laboratory cycling; it requires robustness under fragmented field data, pack heterogeneity, privacy constraints, and evolving evidence regimes.

The most consequential advances will therefore come from explicit coupling of observability, mechanism, uncertainty, and action across the battery lifecycle. When benchmarking, data semantics, and governance mature alongside model development, the digital twin can move from an attractive concept to an auditable battery-intelligence layer for safe fast charging, predictive safety, durable operation, repurposing, and circular battery value management.

DECLARATIONS

Authors' contributions

drafted and organized the manuscript: HF.W;

contributed to conceptualization, supervision, critical revision, and correspondence: YX.Z, DB.M;

All authors reviewed and approved the final manuscript.

Availability of data and materials

Not applicable.

Financial support and sponsorship

This work was financially supported by the National Key Research and Development Program of China (2024YFF0505900), the China Postdoctoral Science Foundation General Fund (No. 77, 2025M774198), and the Postdoctoral Fellowship Program of CPSF under Grant Number GZC20252688.

Conflicts of interest

The authors declare no conflict of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

REFERENCES

[1] C. Lv, X. Zhou, L. Zhong, et al., Machine Learning: An Advanced Platform for Materials Development and State Prediction in Lithium-Ion Batteries, *Adv. Mater.* 2022, 34, 2101474.[DOI:10.1002/adma.202101474]

[2] C. Ling, A review of the recent progress in battery informatics, *npj Comput. Mater.* 2022, 8, 33.[DOI: 10.1038/s41524-022-00713-x]

[3] Y. Wang, Application-oriented design of machine learning paradigms for battery science, *npj Comput. Mater*. 2025, 11, 89.[DOI: 10.1038/s41524-025-01575-9]

[4] M. Borah, Q. Wang, S. Moura, et al., Synergizing physics and machine learning for advanced battery management, *Commun. Eng.* 2024, 3, 134.[PMID: 39300192 PMCID: PMC11413374 DOI: 10.1038/s44172-024-00273-6]

[5] A. Thelen, X. Huan, N. Paulson, et al., Probabilistic machine learning for battery health diagnostics and prognostics, *npj Mater. Sustain.* 2024, 2, 14.[DOI: 10.1038/s44296-024-00011-1]

[6] P. M. Attia, A. Grover, N. Jin, et al., Closed-loop optimization of fast-charging protocols for batteries with machine learning, *Nature* 2020, 578, 397.[PMID: 32076218 DOI: 10.1038/s41586-020-1994-5]

[7] W. Huang, Y. Ye, H. Chen, et al., Onboard early detection and mitigation of lithium plating in fast-charging batteries, *Nat. Commun.* 2022, 13, 7091.[PMID: 36402759 PMCID: PMC9675798 DOI: 10.1038/s41467-022-33486-4]

[8] Y. Zeng, F. Shen, B. Zhang, et al., Nonintrusive thermal-wave sensor for operando quantification of degradation in commercial batteries, *Nat. Commun.* 2023, 14, 8203.[PMID: 38081869 PMCID: PMC10713567 DOI: 10.1038/s41467-023-43808-9]

[9] D. Wasylowski, H. Ditler, M. Sonnet, et al., Operando visualisation of lithium plating by ultrasound imaging of battery cells, *Nat. Commun.* 2024, 15, 10237.[PMID: 39592572 PMCID: PMC11599900 DOI: 10.1038/s41467-024-54319-6]

[10] M. Ishigaki, K. Ishikawa, T. Usuki, H. Kondo, S. Komagata, T. Sasaki, Operando Li metal plating diagnostics via MHz band electromagnetics, *Nat. Commun.* 2023, 14, 7275.[PMID: 37949884 PMCID: PMC10638420 DOI: 10.1038/s41467-023-43138-w]
[11] F. von Bülow, F. Heinrich, W. A. Paxton, The future of battery data and the state of health of lithium-ion batteries in automotive applications, *Commun. Eng.* 2024, 3, 173.[PMID: 39562828 PMCID: PMC11577075 DOI: 10.1038/s44172-024-00299-w]
[12] H. Liu, C. Li, X. Hu, et al., Multi-modal framework for battery state of health evaluation using open-source electric vehicle data, *Nat. Commun.* 2025, 16, 1137.[PMID: 39880811 PMCID: PMC11779878 DOI: 10.1038/s41467-025-56485-7]
[13] R. Li, N. D. Kirkaldy, F. F. Oehler, et al., The importance of degradation mode analysis in parameterising lifetime prediction models of lithium-ion battery degradation, *Nat. Commun.* 2025, 16, 2776.[PMID: 40113781 PMCID: PMC11926349 DOI: 10.1038/s41467-025-57968-3]
[14] A. Lanubile, P. Bosoni, G. Pozzato, et al., Domain knowledge-guided machine learning framework for state of health estimation in lithium-ion batteries, *Commun. Eng.* 2024, 3, 168.[PMID: 39533114 PMCID: PMC11557864 DOI: 10.1038/s44172-024-00304-2]
[15] A. Blömeke, H. Zappen, F. Ringbeck, et al., Balancing resistor-based online electrochemical impedance spectroscopy in battery systems: opportunities and limitations, *Commun. Eng.* 2024, 3, 62.[DOI: 10.1038/s44172-024-00203-6]
[16] S. Clark, F. L. Bleken, S. Stier, et al., Toward a Unified Description of Battery Data, *Adv. Energy Mater.* 2022, 12, 2102702.[DOI: 10.1002/aenm.202102702]
[17] X. Liu, L. Zhang, H. Yu, et al., Bridging Multiscale Characterization Technologies and Digital Modeling to Evaluate Lithium Battery Full Lifecycle, *Adv. Energy Mater.* 2022, 12, 2200889.[DOI: 10.1002/aenm.202200889]
[18] F. Wang, Z. Zhai, Z. Zhao, Y. Di, X. Chen, Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis, *Nat. Commun.* 2024, 15, 4332.[DOI: 10.1038/s41467-024-48779-z]
[19] W. Guo, Z. Sun, J. Guo, et al., Digital Twin-Assisted Degradation Diagnosis and Quantification of NMC Battery Aging Effects During Fast Charging, *Adv. Energy Mater.* 2024, 14, 2401644.[DOI: 10.1002/aenm.202401644]

[20] Y. Che, Y. Zheng, J. Rhyu, et al., Mechanistically guided residual learning for battery state monitoring throughout life, *Nat. Commun.* 2026, 17, 855.[PMID: 41540022 PMCID: PMC12827459 DOI: 10.1038/s41467-025-67565-z]
[21] P. K. Jones, U. Stimming, A. A. Lee, Impedance-based forecasting of lithium-ion battery performance amid uneven usage, *Nat. Commun.* 2022, 13, 4806.[PMID: 35974010 PMCID: PMC9381522 DOI: 10.1038/s41467-022-32422-w]
[22] D. Roman, S. Saxena, V. Robu, M. Pecht, D. Flynn, Machine learning pipeline for battery state-of-health estimation, *Nat. Mach. Intell.* 2021, 3, 447.[PMID: 41851236 PMCID: PMC13048978 DOI: 10.1038/s41598-026-43692-5]
[23] J. Lu, R. Xiong, J. Tian, C. Wang, F. Sun, Deep learning to estimate lithium-ion battery state of health without additional degradation experiments, *Nat. Commun.* 2023, 14, 2760.[PMID: 37179411 PMCID: PMC10183024 DOI: 10.1038/s41467-023-38458-w]
[24] H. Zhang, Y. Li, S. Zheng, et al., Battery lifetime prediction across diverse ageing conditions with inter-cell deep learning, *Nat. Mach. Intell.* 2025, 7, 270.[DOI: 10.1038/s42256-024-00972-x]
[25] K. A. Severson, P. M. Attia, N. Jin, et al., Data-driven prediction of battery cycle life before capacity degradation, *Nat. Energy* 2019, 4, 383.[DOI: 10.1038/s41560-019-0356-8]
[26] W. Mei, Z. Liu, C. Wang, et al., Operando monitoring of thermal runaway in commercial lithium-ion cells via advanced lab-on-fiber technologies, *Nat. Commun.* 2023, 14, 5251.[PMID: 37640698 PMCID: PMC10462619 DOI: 10.1038/s41467-023-40995-3]
[27] X. Gu, Y. Shang, J. Li, et al., Early warning of thermal runaway based on state of safety for lithium-ion batteries, *Commun. Eng.* 2025, 4, 106.[PMID: 40494989 PMCID: PMC12152129 DOI: 10.1038/s44172-025-00442-1]
[28] H. Yang, J. Tian, W. Mai, et al., Privacy-preserving collaborative battery fault warning for massive electric vehicles by heterogeneous data from charging stations, *Nat. Commun.* 2026, 17, 974.[PMID: 41419740 PMCID: PMC12847864 DOI: 10.1038/s41467-025-67703-7]
[29] Z. Zhang, X. Gu, Y. Zhu, et al., Data-driven available capacity estimation of lithium-ion batteries based on fragmented charge capacity, *Commun. Eng.* 2025, 4, 32.[PMID: 39994361 PMCID: PMC11850593 DOI: 10.1038/s44172-025-00372-y]

[30] D. P. Finegan, J. Billman, J. Darst, et al., The battery failure databank: insights from an open-access database of thermal runaway behaviors of Li-ion cells and a resource for benchmarking risks, *J. Power Sources* 2024, 597, 234106.[DOI: 10.1016/j.jpowsour.2024.234106]
[31] P. Bilfinger, M. Schreiber, P. Rosner, et al., Why we need a standardized state of health measurement procedure for electric vehicle battery packs—a proposal for energy- and capacity-based metrics, *npj Clean Energy* 2025, 1, 10.[DOI: 10.1038/s44406-025-00010-8]
[32] Regulation (EU) 2023/1542 of the European Parliament and of the Council of 12 July 2023 concerning batteries and waste batteries, *Off. J. Eur. Union* 2023, L 191, 1.
[33] A. B. Faheem, Z. Han, D. Wu, H. Li, AI-Driven Big Data Frameworks for Electrode-Electrolyte Interphases in Batteries, *Adv. Mater.* 2026, 38, e21975.[DOI: 10.1002/adma.202521975]
[34] Z. He, Z. Wang, Y. Dong, et al., Coupling Data-Driven and Reinforcement Learning for Material Development and Device Management in Batteries, *Adv. Mater.* 2026, 38, e72332.[PMID: 41645654 DOI: 10.1002/adma.72332]
[35] X. Lu, M. Lagnoni, A. Bertei, S. Das, R.E. Owen, Q. Li, K. O'Regan, A. Wade, D.P. Finegan, E. Kendrick, M.Z. Bazant, D.J.L. Brett, P.R. Shearing, Multiscale dynamics of charging and plating in graphite electrodes coupling operando microscopy and phase-field modelling, *Nat. Commun.* 14 (2023) 1-14.[PMID: 37620348 PMCID: PMC10449918 DOI: 10.1038/s41467-023-40574-6]
[36] S. Li, C. Zhang, Y. Zhao, G.J. Offer, M. Marinescu, Effect of thermal gradients on inhomogeneous degradation in lithium-ion batteries, *Commun. Eng.* 2 (2023).[DOI: 10.1038/s44172-023-00124-w]
[37] W. Li, M. Rentemeister, J. Badeda, D. Jöst, D. Schulte, D. Sauer, Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation, *Journal of Energy Storage* 30 (2020) 101557.[DOI: 10.1016/j.est.2020.101557]
[38] S. Tao, H. Liu, C. Sun, et al., Collaborative and privacy-preserving retired battery sorting for profitable direct recycling via federated machine learning, *Nat. Commun.* 14 (2023) 8032.[PMID: 38052823 PMCID: PMC10697957 DOI: 10.1038/s41467-023-43883-y]

[39] X. Li, S. Zhao, C. Gwan, Z. Wang, S. Jiao, Y. Zhu, High-fidelity hierarchical modeling of lithium-ion batteries: a cross-scale electrochemical-mechanical framework, *Commun. Eng*. 2026, 5, 10.[PMID: 41402441 PMCID: PMC12808785 DOI: 10.1038/s44172-025-00567-3]

[40] G. Ma, S. Xu, B. Jiang,et al., Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning, *Energy Environ. Sci.* 15 (2022) 4083-94.[DOI:10.1039/D2EE01676A]

[41] Y. Preger, H.M. Barkholtz, A. Fresquez, et al., Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions, *J. Electrochem. Soc.* 167 (2020) 120532.[DOI: 10.1149/1945-7111/abae37]

[42] W. He, N. Williard, M. Osterman, M. Pecht, Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method, *J. Power Sources* 196 (2011) 10314-21.[DOI: 10.1016/j.jpowsour.2011.08.040]

[43] J. Hu, P. Fu, Z. Wei, Y. Huang, J. Early, A. Fly, Y. Zhang, Early prediction of lithium-ion battery degradation with a generative pre-trained transformer, *Nat. Commun.* 2026, 17, 126.[DOI: 10.1038/s41467-025-66819-0]

[44] J. Huang, Q. Yang, A. Hu, et al., Enhanced specific energy in fast-charging lithium-ion batteries negative electrodes via Ti-O covalency-mediated low potential, *Nat. Commun.* 16 (2025) 1-13.[PMID: 40624036 PMCID: PMC12234993 DOI: 10.1038/s41467-025-61461-2]

[45] X. Han, S. Mao, Y. Wang, et al., Manipulation of lithium dendrites based on electric field relaxation enabling safe and long-life lithium-ion batteries, *Nat. Commun.* 16 (2025) 1-11.[PMID: 40251155 PMCID: PMC12008231 DOI: 10.1038/s41467-025-58818-y]

[46] Y. Ou, B. Zhang, R. Zhan, et al., A salt-free medium facilitating electrode prelithiation towards fast-charging and high-energy lithium-ion batteries, *Nat. Commun.* 16 (2025).[PMID: 40858605 PMCID: PMC12381243 DOI: 10.1038/s41467-025-63257-w]

[47] K. Masalkovaitė, P. Gasper, D.P. Finegan, Predicting the heat release variability of Li-ion cells under thermal runaway with few or no calorimetry data, *Nat. Commun.* 15 (2024) 1-11.[PMID: 39333523 PMCID: PMC11437010 DOI: 10.1038/s41467-024-52653-3]

[48] S. Tao, R. Ma, Z. Zhao, G. Ma, L. Su, H. Chang, et al., Generative learning assisted state-of-health estimation for sustainable battery recycling with random retirement conditions, *Nat. Commun.* 2024, 15, 10154.[PMID: 39578484 PMCID: PMC11584641 DOI: 10.1038/s41467-024-54454-0]
[49] S. Li, M. Zhang, R. Doel, B. Ross, M. D. Piggott, Deep learning predicts real-world electric vehicle direct current charging profiles and durations, *Nat. Commun.* 2025, 16, 10921.[PMID: 41350260 PMCID: PMC12680626 DOI: 10.1038/s41467-025-65970-y]
[50] P. Bilfinger, P. Rosner, M. Schreiber, et al., Battery pack diagnostics for electric vehicles: Transfer of differential voltage and incremental capacity analysis from cell to vehicle level, *ETransportation* 22 (2024) 100356.[DOI: 10.1016/j.etran.2024.100356]
[51] R. Cao, Z. Zhang, R. Shi, J. Lu, Y. Zheng, Y. Sun, X. Liu, S. Yang, Model-constrained deep learning for online fault diagnosis in Li-ion batteries over stochastic conditions, *Nat. Commun.* 16 (2025) 20-25.[PMID: 39952987 PMCID: PMC11829048 DOI: 10.1038/s41467-025-56832-8]
[52] Y. Zhao, Z. Wang, L. Yan, Z. Sun, P. Liu, L. Zhang, W. Li, Bridging battery degradation and safety: challenges and opportunities, *eTransportation* 2025, 26, 100497.[DOI: 10.1016/j.etran.2025.100497]
[53] Y. Wei, M. Wang, M. Zhang, T. Cai, Y. Huang, M. Xu, Advancements, challenges, and future trajectories in advanced battery safety detection, *Electrochem. Energy Rev.* 2025, 8, 10.[DOI: 10.1007/s41918-025-00245-0]
[54] Dubarry M, Howey D, Wu B. Enabling battery digital twins at the industrial scale. *Joule* 2023;7:1134-44.[DOI: 10.1016/j.joule.2023.05.005]
[55] Zhao J, Feng X, Tran M-K, Fowler M, Ouyang M, Burke AF. Battery safety: fault diagnosis from laboratory to real world. *J Power Sources* 2024;598:234111.[DOI: 10.1016/j.jpowsour.2024.234111]
[56] Li W, Chen J, Quade K, Luder D, Gong J, Sauer DU. Battery degradation diagnosis with field data, impedance-based modeling and artificial intelligence. *Energy Storage Mater* 2022;53:391-403.[DOI: 10.1016/j.ensm.2022.08.021]
[57] Xu Y, Ge X, Guo R, Shen W. Recent advances in model-based fault diagnosis for lithium-ion batteries: a comprehensive review. *Renew Sustain Energy Rev* 2025; 207:114922.[DOI: 10.1016/j.rser.2024.114922]

[58] Hu X, Zhang K, Liu K, Lin X, Dey S, Onori S. Advanced fault diagnosis for lithium-ion battery systems: a review of fault mechanisms, fault features, and diagnosis procedures. *IEEE Ind Electron Mag* 2020;14:65-91.[DOI:10.36227/techrxiv.11777448]
[59] Shang Y, Wang S, Tang N, Fu Y, Wang K. Research progress in fault detection of battery systems: a review. *J Energy Storage* 2024;98:113079.[DOI: 10.1016/j.est.2024.113079]
[60] Yu Q, Wang C, Li J, Xiong R, Pecht M. Challenges and outlook for lithium-ion battery fault diagnosis methods from the laboratory to real world applications. *Etransportation* 2023;17:100254.[DOI: 10.1016/j.etran.2023.100254]
[61] Hu X, Gao F, Xiao Y, et al. Advancements in the safety of lithium-ion battery: the trigger, consequence and mitigation method of thermal runaway. *Chem Eng J* 2024:148450.[DOI: 10.1016/j.cej.2023.148450]
[62] Ding Y, Lu L, Zhang H. Multi-physics simulation and risk analysis of internal thermal runaway propagation in lithium-ion batteries. *eTransportation* 2025;24: 100427.[DOI: 10.1016/j.etran.2025.100427]
[63] Wang G, Gao W, He X, et al. Numerical investigation on thermal runaway propagation and prevention in cell-to-chassis lithium-ion battery system. *Appl Therm Eng* 2024;236:121528.[DOI: 10.1016/j.applthermaleng.2023.121528]
[64] Feng X, Ouyang M, Liu X, Lu L, Xia Y, He X. Thermal runaway mechanism of lithium ion battery for electric vehicles: a review. *Energy Storage Mater* 2018;10: 246-67.[DOI: 10.1016/j.ensm.2017.05.013]
[65] S. Ding, L. Wang, H. Dai, X. He, Prognosticating nonlinear degradation in lithium-ion batteries: operando pressure as an early indicator preceding other signals of capacity fade and safety risks, *Energy Storage Mater.* 2025, 75, 103998.[DOI: 10.1016/j.ensm.2024.103998]
[66] Che Y, Hu X, Teodorescu R. Opportunities for battery ageing mode diagnosis of renewable energy storage. *Joule* 2023;7:1405-7.[DOI: 10.1016/j.joule.2023.06.014]
[67] Song S, Tang X, Sun Y, et al. Fault evolution mechanism for lithium-ion battery energy storage system under multi-levels and multi-factors. *J Energy Storage* 2024;80:110226.[DOI: 10.1016/j.est.2023.110226]
[68] Zhou Y, Zhu X, Wang Z, Shan T, Zhang J, Sun Z. Safety assessment of thermal runaway behavior of lithium-ion cells with actual installed state. *Appl Therm Eng* 2023;229:120617.[DOI: 10.1016/j.applthermaleng.2023.120617]

[69] Z. Wang, J. Yuan, X. Zhu, et al., Overcharge-to-thermal-runaway behavior and safety assessment of commercial lithium-ion cells with different cathode materials: a comparison study, *J. Energy Chem.* 2021, 55, 484-498.[DOI: 10.1016/j.jechem.2020.07.028]

[70] Lyu P, Liu X, Qu J, et al. Recent advances of thermal safety of lithium ion battery for energy storage. *Energy Storage Mater* 2020;31: 195-220.[DOI: 10.1016/j.ensm.2020.06.042]

[71] Hou J, Lu L, Wang L, et al. Thermal runaway of Lithium-ion batteries employing LiN(SO2F)2-based concentrated electrolytes. *Nat Commun* 2020;11:5100.[DOI: 10.1038/s41467-020-18868-w]

[72] Zhou P, Liang J, Liu Y, Wu J, Song Q, Li X. Capacity estimation of retired lithium-ion batteries using random charging segments from massive real-world data. *CR- PHYS-SC* 2025;6:102444.[DOI: 10.1016/j.xcrp.2025.102444]

[73] Sudarshan M, Serov A, Jones C, Ayalasomayajula SM, García RE, Tomar V. Data- driven autoencoder neural network for onboard BMS Lithium-ion battery degradation prediction. *J Energy Storage* 2024;82:110575.[DOI: 10.1016/j.est.2024.110575]

[74] X. Yang, H. Xie, L. Zhang, et al., Early-stage degradation trajectory prediction for lithium-ion batteries: a generalized method across diverse operational conditions, *J. Power Sources* 2024, 612, 234808.[DOI: 10.1016/j.jpowsour.2024.234808]

[75] Ye J, Xie Q, Lin M, Wu J. A method for estimating the state of health of lithium-ion batteries based on physics-informed neural network. *Energy* 2024;294: 130828.[DOI: 10.1016/j.energy.2024.130828]

[76] Fan Y, Xiao F, Li C, Yang G, Tang X. A novel deep learning framework for state of health estimation of lithium-ion battery. *J Energy Storage* 2020;32:101741.[DOI: 10.1016/j.est.2020.101741]

[77] Ospina Agudelo B, Zamboni W, Postiglione F, Monmasson E. Battery state-of-health estimation based on multiple charge and discharge features. *Energy* 2023; 263:125637.[DOI: 10.1016/j.energy.2022.125637]

[78] Wang J, Xiang Y. Fast modeling of the capacity degradation of lithium-ion batteries via a conditional temporal convolutional encoder-decoder. *IEEE Trans Transp Electrific* 2022;8:1695-709.[DOI: 10.1109/TTE.2021.3128018]

[79] Hu J, Fu P, Hu X, et al., Toward advanced estimation of state of health for integral lithium-ion battery pack. *CR-PHYS-SC* 2025;6.[DOI: 10.1016/j.xcrp.2024.102363]
[80] Ning Y, Yang F, Zhang Y, et al., Bridging multimodal data and battery science with machine learning. *Matter* 2024;7:2011-32.[DOI: 10.1016/j.matt.2024.04.030]
[81] Zuo W, Zheng H, He T, Vishwanath V, Chan MKY, Stevens RL, et al. Large language models for batteries. *Joule* 2025;9.[DOI: 10.1016/j.joule.2025.102037]
[82] Sun W, Wu C, Xie C, et al., Fine-tuning enables state of health estimation for lithium-ion batteries via a time series foundation model. *Energy* 2025;318:134177.[DOI: 10.1016/j.energy.2024.134177]
[83] M. Talal Qasem, J. Stubblefield, M. Qandil, Y. Yassin, M. Haddadin, M. Krishnamurthy, *eTransportation* 2026, 27, 100528.
[84] S. Kim, H. Lee, J. Lim, J. Park, Y. M. Lee, Digital Twin Battery Modeling and Simulations: A New Analysis and Design Tool for Rechargeable Batteries, *ACS Energy Lett.* 2024, 9, 5225-39.[DOI: 10.1021/acsenergylett.4c01931]
[85] Zhang X, Jia H, Xu Y, Zou L, Engelhard MH, Matthews BE, et al. Unravelling high- temperature stability of lithium-ion battery with lithium-rich oxide cathode in localized high-concentration electrolyte. *J Power Sources Adv* 2020;5:100024.[DOI: 10.1016/j.powera.2020.100024]
[86] Du H, Wang Y, Kang Y, et al., Side reactions/changes in lithium-ion batteries: mechanisms and strategies for creating safer and better batteries. *Adv Mater* 2024;36:2401482.[PMID: 38695389 DOI: 10.1002/adma.202401482]
[87] Bj¨orklund E, Xu C, Dose WM, et al., Cycle-induced interfacial degradation and transition-metal cross-over in $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$-Graphite cells. *Chem Mater* 2022;34:2034-48.[PMID: 35557994 PMCID: PMC9082506 DOI: 10.1021/acs.chemmater.1c02722]
[88] Edge JS, O'Kane S, Prosser R, et al., Lithium ion battery degradation: what you need to know. *Phys Chem Chem Phys* 2021;23: 8200-21.[PMID: 33875989 DOI: 10.1039/d1cp00359c]
[89] Hu X, Xu L, Lin X, Pecht M. Battery lifetime prognostics. *Joule* 2020;4:310-46.[DOI: 10.1016/j.joule.2019.11.018]
[90] Zhang Y, Cheng S, Mei W, et al., Understanding of thermal runaway mechanism of LiFePO4 battery in-depth by three-level analysis. *Appl Energy* 2023;336:120695.[DOI: 10.1016/j.apenergy.2023.120695]

[91] Tang S, Liang Y, Zhong C, et al., Revisiting the overdischarge process as a novel accelerated ageing method for LiFePO4/Graphite batteries through the unveiling of SEI evolution mechanism. *Energy Storage Mater* 2025; 74:103916.[DOI: 10.1016/j.ensm.2024.103916]
[92] Kasnatscheew J, B¨orner M, Streipert B, et al., Lithium ion battery cells under abusive discharge conditions: electrode potential development and interactions between positive and negative electrode. *J Power Sources* 2017;362:278-82.[DOI: 10.1016/j.jpowsour.2017.07.044]
[93] Gao T, Bai J, Ouyang D, et al., Effect of ageing temperature on thermal stability of lithium-ion batteries: part A - high-temperature ageing. *Renew Energy* 2023;203:592-600.[DOI: 10.1016/j.renene.2022.12.092]
[94] Yang C, Singh A, Pu X, et al., Addressing the safety of next-generation batteries. *Nature* 2025;645:603-13.[PMID: 40962979 DOI: 10.1038/s41586-025-09358-4]
[95] Zhang L, Liu L, Terekhov A, Warnberg D, Zhao P. Thermal runaway of Li-ion battery with different ageing histories. *Process Saf Environ Prot* 2024;185:910-7.[DOI: 10.1016/j.psep.2024.03.077]
[96] G. Zhang, X. Wei, X. Wang, et al., Lithium-ion battery sudden death: safety degradation and failure mechanism, *eTransportation* 2024, 20, 100328.[DOI: 10.1016/j.etran.2024.100333]
[97] Maradesa A, Py B, Huang J, et al., Advancing electrochemical impedance analysis through innovations in the distribution of relaxation times method. *Joule* 2024;8:1958-81.[DOI: 10.1016/j.joule.2024.05.008]
[98] Messing M, Shoa T, Habibi S. Estimating battery state of health using electrochemical impedance spectroscopy and the relaxation effect. *J Energy Storage* 2021;43:103210.[DOI: 10.1016/j.est.2021.103210]
[99] Fan K, Wan Y, Wang Z, Jiang K. Time-efficient identification of lithium-ion battery temperature-dependent OCV-SOC curve using multi-output Gaussian process. *Energy* 2023;268:126724.[DOI: 10.1016/j.energy.2023.126724]
[100] Gou B, Xu Y, Feng X. An ensemble learning-based data-driven method for online state-of-health estimation of lithium-ion batteries. *IEEE Trans Transp Electrific* 2021;7:422-36.[DOI: 10.1109/TTE.2020.3029295]

[101] Wang Q, Wang Z, Liu P, Zhang L, Sauer DU, Li W. Large-scale field data-based battery ageing prediction driven by statistical features and machine learning. *CR-PHYS-SC* 2023;4:101720.[DOI: 10.1016/j.xcrp.2023.101720]

[102] Deng Z, Xu L, Liu H, Hu X, Wang B, Zhou J. Rapid health estimation of in-service battery packs based on limited labels and domain adaptation. *J Energy Chem* 2024;89:345-54.[DOI: 10.1016/j.jechem.2023.10.056]

[103] Tao S, Zhang M, Zhao Z, et al., Non-destructive degradation pattern decoupling for early battery trajectory prediction via physics-informed learning. *Energy Environ Sci* 2025;18:1544-59.[DOI: 10.1039/D4EE03839H]

[104] Yan L, Peng J, Gao D, et al., A hybrid method with cascaded structure for early-stage remaining useful life prediction of lithium-ion battery. *Energy* 2022;243:123038.[DOI: 10.1016/j.energy.2021.123038]

[105] Uy OM, Maurer RH. Fault tree safety analysis of a large Li/SOCl2 spacecraft battery. *J Power Sources* 1987;21:207-25.[DOI: 10.1016/0378-7753(87)80055-1]

[106] Wu C, Zhu C, Ge Y, Zhao Y. A review on fault mechanism and diagnosis approach for Li-Ion batteries. *J Nanomater* 2015;2015:631263.[DOI: 10.1155/2015/631263]

[107] Zhu X, Wang Z, Wang Y, et al., Overcharge investigation of large format lithium-ion pouch cells with $Li(Ni_{0.6}Co_{0.2}Mn_{0.2})O_2$ cathode for electric vehicles: thermal runaway features and safety management method. *Energy* 2019;169:868-80.[DOI: 10.1016/j.energy.2018.12.041]

[108] Chen W, Chen W-T, Saif M, Li M-F, Wu H. Simultaneous fault isolation and estimation of lithium-ion batteries via synthesized design of luenberger and learning observers. *IEEE Trans Automat Control* 2014;22:290-8.[DOI: 10.1109/TCST.2013.2239296]

[109] Chen Z, Xiong R, Tian J, Shang X, Lu J. Model-based fault diagnosis approach on external short circuit of lithium-ion battery used in electric vehicles. *Appl Energy* 2016;184:365-74.[DOI: 10.1016/j.apenergy.2016.10.026]

[110] Gao W, Zheng Y, Ouyang M, Li J, Lai X, Hu X. Micro-short-circuit diagnosis for series-connected lithium-ion battery packs using mean-difference model. *IEEE Trans Ind Electron* 2019;66:2132-42.[DOI: 10.1109/TIE.2018.2838109]

[111] Cao R, Zhang Z, Lin J, et al., Reliable online internal short circuit diagnosis on lithium-ion battery packs via voltage anomaly detection based on the mean-difference

model and the adaptive prediction algorithm. *Batteries* 2022;8:224.[DOI: 10.3390/batteries8110224]
[112] Wu X, Wei Z, Wen T, Du J, Sun J, Shtang AA. Research on short-circuit fault-diagnosis strategy of lithium-ion battery in an energy-storage system based on voltage cosine similarity. *J Energy Storage* 2023;71:108012.[DOI: 10.1016/j.est.2023.108012
[113] Li X, Dai K, Wang Z, Han W. Lithium-ion batteries fault diagnostic for electric vehicles using sample entropy analysis method. *J Energy Storage* 2020;27:.[DOI: 10.1016/j.est.2019.101121]
[114] Xia B, Shang Y, Nguyen T, Mi C. A correlation based fault detection method for short circuits in battery packs. *J Power Sources* 2017;337:1-10.[DOI: 10.1016/j.jpowsour.2016.11.007]
[115] Sun Z, Han Y, Wang Z, et al., Detection of voltage fault in the battery system of electric vehicles using statistical analysis. *Appl Energy* 2022; 307:118172.[DOI: 10.1016/j.apenergy.2021.118172]
[116] Zhao J, Ling H, Wang J, Burke AF, Lian Y. Data-driven prediction of battery failure for electric vehicles. *iScience* 2022;25:104172.[PMID: 35434566 PMCID: PMC9010759 DOI: 10.1016/j.isci.2022.104172]
[117] Zhang X, Yang W, Yan L, Kaleem MB, Liu W. Adaptive internal short-circuit fault detection for lithium-ion batteries of electric vehicles. *J Energy Storage* 2024;84:.[DOI: 10.1016/j.est.2024.110874]
[118] Zhang J, Wang Y, Jiang B, et al., Realistic fault detection of li-ion battery via dynamical deep learning. *Nat Commun* 2023;14:5940.[DOI: 10.1038/s41467-023-41226-5]
[119] Qiao D, Wei X, Jiang B, et al., Quantitative diagnosis of internal short circuit for lithium-ion batteries using relaxation voltage. *IEEE Trans Ind Electron* 2024:1-10.[DOI: 10.1109/TIE.2023.3342289]
[120] Zhao Y, Deng J, Liu P, et al., Enhancing battery durable operation: Multi-Fault diagnosis and safety evaluation in series-connected lithium-ion battery systems. *Appl Energy* 2025;377:124632.[DOI: 10.1016/j.apenergy.2024.124632]
[121] Cai L, Wang H, Dong Z, He Z, Gao M, Song Y. A multi-fault diagnostic method based on category-reinforced domain adaptation network for series-connected battery packs. *J Energy Storage* 2023;60:106690.[DOI: 10.1016/j.est.2023.106690]

[122] Tian, E.G., Chen, H., et al., Security-ensured state of charge estimation of lithium-ion batteries subject to malicious attacks. *IEEE Trans. Smart Grid* 14, 2250-2261 (2023).[DOI: 10.1109/TSG.2022.3202811]

[123] Liu, K.L., Wei, Z.B., et al., Towards long lifetime battery: AI-based manufacturing and management. *IEEE-CAA J. Automatica Sin.* 9, 1139-1165 (2022).[DOI: 10.1109/JAS.2022.105599]

[124] Zhang, Y.Z., Wik, T., Bergström, J., et al., State of health estimation for lithium-ion batteries under arbitrary usage using data-driven multimodel fusion. *IEEE Trans. Transp. Electrif.* 10, 1494-1507 (2024).[DOI: 10.1109/TTE.2023.3267124]

[125] Zhao, J.Y., Feng, X.N., Pang, Q.Q., et al., Battery safety: machine learning-based prognostics. *Prog. Energy Combust. Sci.* 102, 101142 (2024).[DOI: 10.1016/j.pecs.2023.101142]

[126] Feng, Y., Zhou, L.M., Ma, H., et al., Challenges and advances in wide-temperature rechargeable lithium batteries. *Energy Environ. Sci.* 15, 1711-1759 (2022).[DOI: 10.1039/D1EE03292E]

[127] Du, Y.T., Fujita, K., Shironita, S., et al.: Capacity fade characteristics of nickel-based lithium-ion secondary battery after calendar deterioration at 80 °C. *J. Power Sources* 501, 230005 (2021).[DOI: 10.1016/j.jpowsour.2021.230005]

[128] Hu, A.J., Li, F., Chen, W., et al., Ion transport kinetics in low-temperature lithium metal batteries. *Adv. Energy Mater.* 12, 2202432 (2022).[DOI: 10.1002/aenm.202202432]

[129] Byun, S., Park, J., Appiah, W.A., et al., The effects of humidity on the self-discharge properties of Li(Ni1/3Co1/3Mn1/3)O2/graphite and -LiCoO2/graphite lithium-ion batteries during storage. *RSC Adv.* 7, 10915-10921 (2017).[DOI: 10.1039/C6RA28516C]

[130] Duan, X.D., Wang, H.C., Jia, Y.K., et al., A multiphysics understanding of internal short circuit mechanisms in lithium-ion batteries upon mechanical stress abuse. *Energy Storage Mater.* 45, 667-679 (2022).[DOI: 10.1016/j.ensm.2021.12.018]

[131] Yang, S., Wang, W.W., Lin, C., et al., Investigation of internal short circuits of lithium-ion batteries under mechanical abusive conditions. *Energies* 12, 1885 (2019).[DOI: 10.3390/en12101885]

[132] Sun, Z.Y., Zhao, J.W., Zhu, M., et al., Critical problems and modification strategies of realizing high-voltage -LiCoO2 cathode from electrolyte engineering. *Adv. Energy Mater.* 14, 2303498 (2024).[DOI: 10.1002/aenm.202303498]

[133] Vinodkumar, E., Rotem, M., Ran, E., et al.: Challenges in the development of advanced Li-ion batteries: a review. *Energy Environ. Sci.* 4, 9 (2011).[DOI: 10.1039/C1EE01598B]

[134] Klein, S., van Wickeren, S., Röser, S., et al., Understanding the outstanding high-voltage performance of NCM523||Graphite lithium ion cells after elimination of ethylene carbonate solvent from conventional electrolyte. *Adv. Energy Mater.* 11, 2003738 (2021).[DOI: 10.1002/aenm.202003738]

[135] Deshpande, R., Verbrugge, M., Cheng, Y.T., et al., Battery cycle life prediction with coupled chemical degradation and fatigue mechanics. *J. Electrochem. Soc.* 159, A1730-A1738 (2012).[DOI: 10.1149/2.049210je]

[136] Birkl, C.R., Roberts, M.R., McTurk, E., et al., Degradation diagnostics for lithium ion cells. *J. Power Sources* 341, 373-386 (2017).[DOI: 10.1016/j.jpowsour.2016.12.011]

[137] Liu, Y.Y., Liao, C.L., Zhang, W.J., et al., Internal short circuit diagnosis of lithium-ion battery based on mechanism model and deep learning. *J. Electrochem. Soc.* 169, 100514 (2022).[DOI: 10.1149/1945-7111/ac91ab]

[138] Xu, J.Y., Ma, J., Zhao, X., et al., Detection technology for battery safety in electric vehicles: a review. *Energies* 13, 4636 (2020).[
DOI: 10.3390/en13184636]

[139] Ren, D.S., Feng, X.N., Liu, L.S., et al., Investigating the relationship between internal short circuit and thermal runaway of lithiumion batteries under thermal abuse condition. *Energy Storage Mater.* 34, 563-573 (2021).[DOI: 10.1016/j.ensm.2020.10.020]

[140] Hendricks, C., Williard, N., Mathew, S., et al., A failure modes, mechanisms, and effects analysis (FMMEA) of lithium-ion batteries. *J. Power Sources* 297, 113-120 (2015).[DOI: 10.1016/j.jpowsour.2015.07.100]

[141] Zhang, S.Z., Guo, X., Dou, X.X., et al., A rapid online calculation method for state of health of lithium-ion battery based on coulomb counting method and differential voltage analysis. *J. Power Sources* 479, 228740 (2020).[DOI: 10.1016/j.jpowsour.2020.228740]

[142] Lee, S., Hong, S., Park, W., et al., High accuracy open-type current sensor with a differential planar hall resistive sensor. *Sensors* 18, 2231 (2018).[PMID: 30002315 PMCID: PMC6068747 DOI: 10.3390/s18072231]
[143] Baorda, R.P., Rosa, T., Angelini, P., et al., Integrated metal shunt with matched sensing resistor for high-side current sensing. In: ESSCIRC 2023- IEEE 49th European Solid State Circuits Conference (ESSCIRC). *Lisbon, Portugal. IEEE,* pp. 241-4(2023).[DOI:10.1109/ESSCIRC59616.2023.10268715]
[144] Li, J.J., Ren, W., Luo, Y.S., et al., Design of fluxgate current sensor based on magnetization residence times and neural networks. *Sensors* 24, 3752 (2024).[PMID: 38931534 PMCID: PMC11207243 DOI: 10.3390/s24123752]
[145] Garcha, P., Schaffer, V., Haroun, B., et al., A duty-cycled integratedfluxgate magnetometer for current sensing. *IEEE J. Solid-State Circuit* 57, 2741-2751 (2022).[DOI: 10.1109/JSSC.2022.3156572]
[146] Kim, C.H., Kim, M.Y., Moon, G.W.: A modularized charge equalizer using a battery monitoring IC for series-connected Li-ion battery strings in electric vehicles. IEEE Trans. *Power Electron.* 28, 8 (2012).[DOI:10.1109/ICPE.2011.5944609]
[147] Nugroho, B., Yahya, A., Rahim, A., et al.: Magnetic field relationship between distance and induced voltage generated by electromagnetic pulse (EMP). In: 2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). *Semarang, Indonesia. IEEE,* pp. 7-10(2018).[DOI: 10.1109/ICITACEE.2018.8576895]
[148] Wang, H.X., Li, K.X., et al., Research on improvement of temperature stability of optical voltage transducer. In: 2014 7th International Conference on Intelligent Computation Technology and Automation. Changsha, China. *IEEE,* pp. 534-538(2014).[DOI: 10.1109/ICICTA.2014.135]
[149] Lee, C.Y., Weng, F.B., Huang, et al., Real-time monitoring of internal temperature and voltage of high-temperature fuel cell stack. *Electrochim. Acta* 161, 413-419 (2015).[DOI: 10.1016/j.electacta.2015.02.135]
[150] Lee, C.-Y., Peng, H.C., et al., A flexible three-in-one microsensor for real-time monitoring of internal temperature, voltage and current of lithium batteries. *Sensors* 15, 11485-11498 (2015).[PMID: 25996509 PMCID: PMC4481915 DOI: 10.3390/s150511485]

[151] You, Y., Zhu, G.Q., Li, et al., 17-Cell battery monitoring analog front end with high sampling accuracy for battery pack applications. *Microelectron. J.* 151, 106327 (2024).[DOI: 10.1016/j.mejo.2024.106327]

[152] Shi, Q.Q., Tang, W., Lv, Q.S.: A high precision voltage detection circuit for multiple lithium batteries. In: 2022 7th International Conference on Integrated Circuits and Microsystems (ICICM). Xi'an, China. *IEEE*, pp. 60-66(2022).[DOI: 10.1109/ICICM56102.2022.10011305]

[153] Mao, B.B., Huang, P.F., Chen, H.D., et al., Self-heating reaction and thermal runaway criticality of the lithium ion battery. *Int. J. Heat Mass Transf.* 149, 119178 (2020).[DOI: 10.1016/j.ijheatmasstransfer.2019.119178]

[154] Fleming, J., Amietszajew, T., Charmet, J., et al., The design and impact of in situ and operando thermal sensing for smart energy storage. *J. Energy Storage* 22, 36-43 (2019).[DOI: 10.1016/j.est.2019.01.026]

[155] Zhu, S.X., Han, J.D., An, H.Y., et al., A novel embedded method for in situ measuring internal multi-point temperatures of lithium ion batteries. *J. Power Sources* 456, 227981 (2020).[DOI: 10.1016/j.jpowsour.2020.227981]

[156] Drake, S.J., Martin, M., Wetz, D.A., et al., Heat generation rate measurement in a Li-ion cell at large C-rates through temperature and heat flux measurements. *J. Power Sources* 285, 266-273 (2015).[DOI: 10.1016/j.jpowsour.2015.03.008]

[157] Zhang, G.S., Cao, L., Ge, S.H., et al.: In situ measurement of radial temperature distributions in cylindrical Li-ion cells. *J. Electrochem. Soc.* 161, A1499-A1507 (2014).[DOI: 10.1149/2.0051410jes]

[158] Mutyala, M.S.K., Zhao, J.Z., et al., In situ temperature measurement in lithium ion battery by transferable flexible thin film thermocouples. *J. Power Sources* 260, 43-49 (2014).[DOI: 10.1016/j.jpowsour.2014.03.004]

[159] Liu, P., Yang, L.Y., Xiao, B.W., et al., Revealing lithium battery gas generation for safer practical applications. *Adv. Funct. Mater.* 32, 2208586 (2022).[DOI: 10.1002/adfm.202208586]

[160] Cai, T., Valecha, P., Tran, V., et al., Detection of Li-ion battery failure and venting with Carbon Dioxide sensors. *eTransportation* 7, 100100(2021).[DOI: 10.1016/j.etran.2020.100100]

[161] Mateev, V., Marinova, I., Kartunov, Z., et al., Gas Leakage Source Detection for Li-Ion Batteries by Distributed Sensor Array. *Sensors* 19, 13 (2019).[PMID: 31262052 PMCID: PMC6651848 DOI: 10.3390/s19132900]
[162] Jin, Y., Zheng, Z.K., Wei, D.H., et al., Detection of micro-scale Li dendrite via -H2 gas capture for early safety warning. *Joule* 4, 1714-1729 (2020).[DOI: 10.1016/j.joule.2020.05.016]
[163] Cheng, K.W.E., Divakar, B.P., Wu, H.J., et al., Battery management system (BMS) and SOC development for electrical vehicles. *IEEE Trans. Veh. Technol.* 60, 76-88 (2011).[DOI: 10.1109/TVT.2010.2089647]
[164] Yi, F., Jiaqiang, E., Zhang, B., et al., Effects analysis on heat dissipation characteristics of lithium-ion battery thermal management system under the synergism of phase change material and liquid cooling method. *Renew. Energy* 181, 472-489 (2022).[DOI: 10.1016/j.renene.2021.09.073]
[165] Xiong, R., Yang, R.X., Chen, Z.Y., et al., Online fault diagnosis of external short circuit for lithium-ion battery pack. *IEEE Trans. Ind. Electron.* 67, 1081-1091 (2020).[DOI: 10.1109/TIE.2019.2899565]
[166] Kim, S., Park, H.J., Choi, J.H., et al., A novel prognostics approach using shifting kernel particle filter of Li-ion batteries under state changes. *IEEE Trans. Ind. Electron.* 68, 3485-3493 (2021).[DOI: 10.1109/TIE.2020.2978688]
[167] Li, C.R., Xiao, F., Fan, Y.X., et al., A recurrent neural network with long short-term memory for state of charge estimation of lithiumion batteries. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). Chongqing, China. *IEEE,* pp. 1712-1716 (2019).
[168] Chowdhury, M.A., Al-Wahaibi, S.S.S., Lu, Q.G.: Adaptive safe reinforcement learning-enabled optimization of battery fast-charging protocols. *AlChE. J.* 71, e18605 (2025).[DOI: 10.48550/arXiv.2406.12309]
[169] Wang, T.J., Dong, Z.Y., Xiong, H.B.: Adaptive multipersonalized federated learning for state of health estimation of multiple batteries. *IEEE Internet Things J.* 11, 39994-40008 (2024).[DOI: 10.1109/JIOT.2024.3448626]
[170] Cho, G., Wang, M.Q., Kim, Y., et al., A physics-informed machine learning approach for estimating lithium-ion battery temperature. *IEEE Access* 10, 88117-88126 (2022).[DOI: 10.1109/ACCESS.2022.3199652]

[171] Ojo, O.J., Lin, X.K., Lang, H.X., et al., A voltage fault detection method enabled by A recurrent neural network and residual threshold monitor for lithium-ion batteries. In: 2021 IEEE Transportation Electrification Conference & Expo (ITEC). Chicago, IL, USA. *IEEE,* pp. 813-820(2021).[DOI: 10.1109/ITEC51675.2021.9490102]
[172] Lu, Y., Wang, X.D., Mao, S.Y., et al., Smart batteries enabled by implanted flexible sensors. *Energy Environ. Sci.* 16, 2448-63 (2023).[DOI: 10.1039/D3EE00695F]
[173] Duan, J., Tang, X., Dai, H.F., et al., Building safe lithium-ion batteries for electric vehicles: A review. *Electrochem. Energy Rev.* 3, 1-42 (2020).[DOI: 10.1007/s41918-019-00060-4]
[174] Yang, S.J., Zhang, C.P., Jiang, J.C., et al., Review on state-of-health of lithium-ion batteries: characterizations, estimations and applications. *J. Clean Prod.* 314, 128015 (2021).[DOI: 10.1016/j.jclepro.2021.128015]
[175] Li, A., Weng, J.W., Yuen, A.C.Y., et al., Machine learning assisted advanced battery thermal management system: a state-of-the-art review. *J. Energy Storage* 60, 106688 (2023).[DOI: 10.1016/j.est.2023.106688]