IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Lifelong Reinforcement Learning for Health-Aware Fast Charging of Lithium-Ion Batteries: A Comprehensive Review

[1]H A Shruti, [2]Dr. Sreenath k

[1]PG Scholar, Department of Electrical & Electronics, Sri Siddhartha Institute of Technology, SSAHE, Tumkur, Karnataka

[2]Associate Professor, Department of Electrical & Electronics, Sri Siddhartha Institute of Technology, SSAHE, Tumkur, Karnataka, India

Doi: https://doi.org/10.56975/ijcrt.v14i7.309891

Abstract: Lithium-ion batteries remain foundational to electric vehicle propulsion and stationary energy storage, yet the persistent tension between fast charging throughput and long-term electrochemical health continues to constrain their practical deployment. Conventional constant-current constant-voltage protocols, while computationally inexpensive, impose rigid current and voltage boundaries that fail to adapt to progressive cell degradation, accelerating solid-electrolyte interphase growth and lithium plating over repeated charge cycles. Advanced control paradigms—spanning model predictive control, side-reaction overpotential regulation, and increasingly, deep reinforcement learning—have emerged as promising alternatives capable of balancing charging speed against battery longevity. Among these, lifelong or continual reinforcement learning frameworks represent the most ambitious frontier, enabling charging agents that continuously adapt their policies as cells age without discarding prior learning.. A structured literature synthesis table is presented to facilitate direct cross-study comparison across topology, method, key contribution, and identified limitation.

Index Terms - Battery degradation, deep reinforcement learning, electric vehicle, fast charging, lifelong learning, lithium-ion battery, model predictive control, single particle model, state of charge, state of health.

I. INTRODUCTION

The accelerating global transition toward electrified mobility and distributed renewable energy integration has placed lithium-ion batteries (LiBs) at the center of modern energy infrastructure [20], [21]. Traction battery packs in battery-electric vehicles must satisfy contradictory demands: drivers require charge replenishment within time intervals competitive with conventional refueling, while original equipment manufacturers and fleet operators demand that cells retain usable capacity across hundreds of thousands of kilometers of service life.

At the electrochemical level, fast charging imposes elevated current densities that perturb lithium-ion concentration gradients within electrode particles and electrolyte phases beyond the self-regulating limits of the intercalation reaction. Two degradation mechanisms dominate the fast-charging regime. First, solid-electrolyte interphase (SEI) growth on the graphite anode surface consumes cyclable lithium inventory and increases internal resistance progressively with each charge cycle through a diffusion-limited reaction pathway [24]. Second, and more acutely damaging, lithium plating occurs when the electrochemical potential at the graphite anode surface falls below the lithium reference potential, causing metallic lithium to deposit rather than intercalate. Plated lithium may subsequently form electrically isolated "dead lithium" or react exothermically with electrolyte constituents, creating safety hazards alongside capacity loss [18], [25].

Conventional constant-current constant-voltage (CC-CV) charging protocols, despite their universal adoption in commercial battery management systems (BMS), were not designed with electrochemical health awareness. The fixed current and voltage thresholds of CC-CV remain static throughout the battery's service life, meaning that the protocol applied to a fresh cell at the beginning of life is identical to the one applied at 80% state of health (SoH), even though the optimal charging parameters change substantially as internal resistance rises and lithium inventory depletes [1], [6]. This rigidity produces two undesirable consequences: an overly conservative charging rate that extends charge duration as the battery ages, or an overly aggressive one that accelerates further degradation. CC-CV variants that dynamically adjust cut-off voltage based on a precomputed SoH mapping partially address this limitation [1], but static parameter mappings cannot fully capture the nonlinear interplay of concurrent degradation mechanisms [24].

Model predictive control (MPC) frameworks have advanced health-aware charging by explicitly incorporating electrochemical constraint sets into the optimization horizon. Lu et al. [5] demonstrated real-time lithium plating detection integrated within an MPC framework alongside adaptive parameter updates tracking cell aging. Hwang et al.

[11] extended a single particle model with first-principles chemical and mechanical degradation terms to minimize intra-cycle capacity fade through optimal current scheduling. Yuan et al. [19] further applied robust MPC to battery pretreatment for safe recycling.

Learning-based approaches have emerged as a compelling complement to, and in some cases substitute for, model-based controllers. Deep reinforcement learning (RL) enables charging agents to discover optimal current profiles through direct interaction with a battery environment, bypassing the need for explicit plant models [2], [3]. Wei et al. [2] embedded Multiphysics constraints implicitly within the reward structure of a deep deterministic policy gradient (DDPG) agent, while Park et al. [3] constructed a comprehensive RL framework evaluating performance with and without direct overpotential measurement. Hao et al. [4] introduced Gaussian process-based adaptive models within an RL loop to enforce operational constraints without requiring a first-principles electrochemical model. Critically, however, the majority of prior RL-based charging works reset the battery model to a fresh initial state at the beginning of each training episode, thereby failing to expose the agent to the slow-timescale degradation dynamics that determine lifelong performance [1].

The work by Yuan and Zou [1], published in IEEE Transactions on Transportation Electrification in October 2025, represents a landmark departure from this episodic paradigm. By training a twin delayed deep deterministic policy gradient (TD3) agent on a single-particle model with electrolyte (SPMe) that ages continuously throughout the training process, their lifelong learning framework produces a charging policy explicitly optimized for the full battery service life. The SoH-dependent voltage constraint integrated into the reward function through a precomputed charge cut-off voltage versus SoH mapping—derived from constant current constant overpotential (CC-COP) experiments—provides an indirect but measurable surrogate for the unmeasurable side-reaction overpotential in commercial cells [1].

.

II. REVIEW METHODOLOGY

This review adheres to a structured and reproducible literature search protocol. Primary database searches were conducted across IEEE Xplore, Scopus, and Web of Science using combinations from the following keyword clusters: "fast charging lithium-ion battery reinforcement learning," "health-aware battery charging," "lifelong reinforcement learning battery," "lithium plating detection fast charging," "model predictive control battery degradation," "single particle model electrolyte," "state of health estimation neural network," and "deep deterministic policy gradient battery." Secondary searches targeted specific algorithmic terms including "TD3 battery," "SPMe aging model," and "constant overpotential charging."

III. BATTERY MODELLING APPROACHES FOR HEALTH-AWARE CHARGING

A. Equivalent Circuit and Reduced-Order Models

Equivalent circuit models (ECMs) approximate battery terminal dynamics through networks of resistors and capacitors parameterized from impedance spectroscopy or pulse characterization data. Their computational efficiency, with typical sampling periods in the millisecond range, makes ECMs the dominant implementation choice for production BMS embedded in automotive microcontrollers. However, ECMs lack the physical grounding required to predict localized phenomena such as concentration polarization, lithium plating overpotential, and SEI film growth that are essential for health-aware charging algorithms. Consequently, ECMs are generally unsuitable as simulation environments for training and validating degradation-aware charging controllers, though they remain valuable for SoC and SoH estimation at the system level [15].

The single particle model (SPM) reduces the Doyle-Fuller-Newman (DFN) framework by representing each electrode as a single representative spherical particle, eliminating the need to resolve spatial gradients in the electrode thickness direction. Li et al. [26] surveyed model order reduction techniques for physics-based LiB management, evaluating the accuracy-complexity trade-off across SPM, SPMe, and various reduced-order DFN implementations. The survey concluded that SPMe represents the most favorable balance for control-oriented applications, providing accurate terminal voltage prediction across C-rates up to approximately 3C with a computational cost amenable to receding-horizon optimization [26]. Khalik et al. [22] addressed the practical challenge of DFN parameter estimation, proposing normalization and sensitivity grouping strategies that reduce ill-conditioning in the nonlinear least-squares fitting problem, while a companion study [23] quantified the computational complexity reductions achievable through systematic model simplification.

B. Single Particle Model with Electrolyte (SPMe)

The SPMe augments the SPM by retaining a lumped electrolyte concentration variable that captures concentration polarization effects responsible for significant voltage prediction errors at elevated C-rates. Yuan and Zou

[1] selected SPMe as the simulation environment for their TD3 training loop, implemented within the PyBaMM open-source platform [27], because it simultaneously captures the electrochemical dynamics relevant to lithium plating and SEI growth while remaining computationally tractable for the multi-thousand episode training process. The SPMe resolves the solid-phase diffusion through a spherical diffusion partial differential equation and computes the terminal voltage as the difference between positive and negative electrode potentials, accounting for electrolyte ohmic losses and SEI film overpotential terms [1].

Hwang et al. [11] extended the SPM to incorporate first-principles chemical (solvent reduction SEI growth) and mechanical (particle cracking) degradation mechanisms within an MPC cost function, enabling optimization of charging profiles that explicitly minimize intra-cycle capacity fade in addition to tracking speed. Huang et al. [12] took an orthogonal approach, proposing machine-in-the-loop network (MINN) architectures that learn the dynamics of differential-algebraic equation systems directly from data, bypassing the need for explicit electrochemical model derivation. This data-driven approach offers potential advantages when first-principles parameters are unknown, though the SPMe with identified parameters remains more interpretable and physically consistent for aging-aware control design [12].

IV. BATTERY DEGRADATION MECHANISMS AND CHARACTERISATION

A. Solid-Electrolyte Interphase Growth

SEI layer formation on graphite anode particles is a thermodynamically spontaneous process driven by the chemical incompatibility between the litigated graphite surface and organic electrolyte solvents. The growing SEI layer consumes cyclable lithium inventory irreversibly, manifesting as a gradual capacity fade that progresses monotonically from the first charge cycle onward. Yuan and Zou [1] employed a two-layer diffusion-limited SEI growth model for the LG M50 cell parameterized by O'Kane et al. [24], with an SEI solvent diffusivity of 2.5×10^{-22} m²/s characterizing the slow kinetics of this process. O'Kane et al. [24] provided a comprehensive mechanistic review of LiB degradation modelling, cataloguing SEI growth, lithium plating, particle cracking, cathode dissolution, and mechanical contact loss as the dominant life-limiting processes and evaluated the fidelity of available mathematical models for each mechanism. Their analysis confirmed that diffusion-limited SEI models, though simplified, capture the dominant capacity fade trends observed experimentally across standard cycling conditions.

B. Lithium Plating and Its Electrochemical Signature

Lithium plating is the dominant fast-charging degradation mechanism and represents the primary constraint against which health-aware charging controllers must guard. Plating initiates when the graphite anode overpotential becomes sufficiently negative to reduce the plating equilibrium potential below zero, causing metallic lithium to deposit on particle surfaces rather than intercalate into the graphite lattice [25]. The side-reaction overpotential, defined as the difference between the graphite electrode potential, the electrolyte potential, and the equilibrium side-reaction potential, serves as the key electrochemical indicator of plating risk [1]. When this quantity becomes negative, further reduction in its magnitude drives progressively faster lithium metal deposition, establishing a thermodynamic threshold below which charging acceleration is counterproductive from a lifetime perspective.

Gao et al. [25] conducted a single-particle operando study revealing the interplay between intercalation and plating kinetics within individual graphite particles, demonstrating that plating initiates preferentially at particle surfaces and propagates inward as local current density increases. This finding motivates the use of particle-level electrochemical models—rather than empirical voltage thresholds—as the basis for plating-aware control design. Wang et al. [16] developed an onboard in-situ lithium plating warning and detection system using deep learning applied to voltage signatures, validating detection in commercial cells without reference electrode instrumentation. Huang et al. [18] similarly demonstrated early non-invasive plating detection in fast-charging cells and integrated the detection signal into a closed-loop BMS. Zhao et al. [17] extended the detection paradigm to lithium-ion capacitors under high-rate charging, underscoring the universality of the plating challenge across cell chemistries. Zhang et al. [7] proposed machine-learning-based lifelong estimation of lithium plating potential from measurable signals, providing a practical surrogate for the unmeasurable side-reaction overpotential needed by overpotential-regulated controllers.

V. State Estimation: SoC and SoH

A. State of Charge Estimation

Accurate real-time SoC estimation is a prerequisite for any charging controller that operates on SoC-referenced cost functions or state vectors. The extended Kalman filter (EKF) and its sigma-point variants have been the dominant embedded SoC estimation algorithms for automotive BMS. Chen et al. [15] advanced this tradition with an adaptive cubature Kalman filter based on a generalized minimum error entropy criterion, demonstrating improved robustness to non-Gaussian disturbances and sensor noise compared with standard EKF implementations. The adaptive covariance adjustment mechanism in their formulation is particularly relevant for aging batteries whose internal impedance characteristics drift from the original filter model. Yuan and Zou [1] leveraged the robustness of such estimators to argue that SoC, as a state input to their TD3 agent, is reliably obtainable from production BMS hardware without additional specialized sensors.

B. State of Health Estimation

SoH estimation presents fundamentally greater challenges than SoC estimation because capacity fade evolves on timescales of hundreds to thousands of charge cycles rather than within a single cycle. Zhang et al. [13] proposed a multi-objective optimization with multi-task learning (CMMOG-MTL) framework for SoH estimation that transfers knowledge across cell types, improving sample efficiency when training data from a new cell chemistry is limited. Zhang et al. [14] developed a hybrid attention network that fuses multi-source data streams—voltage, current, temperature, and impedance—through attention-weighted feature extraction, achieving accurate SoH estimation under non-stationary operating conditions representative of real-world EV usage. Gao et al. [8] addressed a related challenge of recovering degraded battery datasets from real-world EV logs with missing or corrupted segments, enabling retrospective SoH trend analysis from sparse field data. Gao et al. [9] contributed a global parameter sensitivity analysis of electrochemical models for aging batteries, identifying the parameters whose variation most influences model predictions, which guides both model reduction and targeted estimation effort.

The SoH information plays a pivotal role in the health-aware fast-charging framework proposed by Yuan and Zou [1]. Rather than incorporating SoH as an explicit penalty term in the reward function—which would require online SoH estimation with high temporal resolution—their framework embeds SoH influence indirectly through the dynamic adjustment of the cut-off voltage upper bound $V_{ma}\chi(SoH)$. This design choice decouples the charging control loop from the SoH estimation loop, requiring SoH updates only at the inter-cycle timescale where estimation accuracy is highest [1], [13], [14].

VI. FAST-CHARGING CONTROL STRATEGIES

A. Constant Current Constant Voltage and Its Variants

CC-CV remains the industrial standard for LiB charging due to its operational simplicity and compatibility with low-cost charging hardware. In the initial constant-current phase, a fixed charging current flows until the cell terminal voltage reaches a specified upper limit, after which a constant-voltage phase maintains that voltage while the current decays exponentially toward a cut-off threshold. Fast charging via CC-CV is achieved by increasing the constant-current magnitude or raising the constant-voltage target, at the cost of proportionally accelerated lithium plating [1], [6]. Wassiliadis et al. [6] quantified this trade-off rigorously in a model-based study showing that unconstrained high-current CC-CV reduces battery cycle life markedly relative to health-aware alternatives. Zhang et al. [21] demonstrated that scheduling CC-CV charging periods in coordination with grid demand response can reduce peak power draw by 18%, though without addressing cell-level degradation in the current profile design.

Yuan and Zou [1] implemented a CC-CV variant termed CC-CV with varying voltage (CC-CV-V), in which the constant-voltage set point is dynamically adjusted each cycle based on the precomputed $V_{m}{}^{5}{}_{to\varphi}{}^{i}$-SoH mapping. This modification yields a modest 6.82% improvement in equivalent full cycles (EFCs) over baseline CC-CV while maintaining near-identical average charge time. The limited improvement confirms that static parameter mappings, however well calibrated, are insufficient to capture the complex nonlinear interactions among concurrent degradation mechanisms across the full battery service life [1].

B. Constant Overpotential and Side-Reaction-Based Control

The constant current constant overpotential (CC-COP) control paradigm, proposed conceptually in prior work and benchmarked extensively by Yuan and Zou [1], replaces the voltage-regulated second phase of CC-CV with a feedback loop that maintains the side-reaction overpotential at a prescribed reference value. When the reference overpotential is set slightly positive (CC-COP-slow), the controller stringently avoids any electrochemical condition conducive to lithium plating, achieving the longest battery lifespan of all tested strategies (1316 EFCs) at the cost of the longest average charge time (36.02 minutes). A more aggressive negative reference (CC-COP-fast) accepts marginal plating risk in exchange for faster charging (22.40 minutes) while still achieving substantially longer life than CC-CV (1009 EFCs). The fundamental limitation of CC-COP is its dependence on direct side-reaction overpotential measurement, which requires reference electrode instrumentation available only in laboratory-grade three-electrode cells and not in commercial two-electrode battery packs [1], [7].

Wassiliadis et al. [6] presented a PID-based side-reaction overpotential controller specifically designed for automotive BMS implementation, addressing the practical measurement challenge through a model-based observer that estimates the overpotential from measurable terminal signals. Li et al. [10] developed a nonlinear model inversion controller that achieves fast SoC tracking while respecting electrochemical constraints derived from the SPM, providing a complementary approach that does not require explicit overpotential regulation.

C. Model Predictive Control

MPC formulates the fast-charging problem as a finite-horizon constrained optimization solved iteratively at each sampling instant. The control input—typically the charging current—is selected to minimize a cost function penalizing SoC tracking error, constraint violations (terminal voltage, temperature, and plating overpotential), and control effort, subject to explicit inequality constraints on all states and inputs. Lu et al. [5] demonstrated the most complete health-aware MPC framework reviewed herein, integrating real-time lithium plating detection, adaptive electrochemical parameter updates tracking cell aging, and a receding-horizon optimizer that adapts the charging trajectory dynamically throughout battery life. Hwang et al. [11] extended the SPM with first-principles degradation sub-models within an MPC cost function, targeting intra-cycle capacity fade minimization through current shaping. Yuan et al. [19] applied robust MPC to battery pretreatment for recycling, addressing model uncertainty in heavily aged cells. The principal limitation of all MPC variants is the computational burden of solving the optimization problem at every control step, which becomes particularly challenging when the prediction model includes electrochemical degradation sub-models with stiff differential-algebraic equation structures [22], [23], [26].

VII. REINFORCEMENT LEARNING ARCHITECTURES FOR HEALTH-AWARE CHARGING

A. Deep Deterministic Policy Gradient Foundations

Deep reinforcement learning operates through an agent-environment interaction loop in which the agent selects charging current actions based on the observed battery state,receives scalar reward signals encoding the quality of the action, and updates its policy parameters to maximize the expected cumulative reward. For fast-charging control, where the action space is continuous and bounded, off-policy actor-critic algorithms are preferred. The deep deterministic policy gradient (DDPG) algorithm introduced the foundational actor-critic architecture for continuous control, pairing a deterministic policy network with a Q-value critic network and a replay buffer for sample-efficient off-policy learning. Wei et al. [2] deployed DDPG with Multiphysics-motivated reward shaping that implicitly constrained thermal and electrochemical states during fast charging. Park et al. [3] built on DDPG to construct a comprehensive framework evaluating RL charging performance under two sensing scenarios: one with direct side-reaction overpotential feedback and one using only measurable terminal voltage and SoC signals.

B. Twin Delayed Deep Deterministic Policy Gradient (TD3)

The TD3 algorithm addresses three systematic weaknesses of DDPG that reduce stability in continuous control applications. First, TD3 employs twin critic networks and uses the minimum of their Q-value estimates to form the target, substantially reducing the positive bias in Q-value estimates that causes DDPG to overfit to overestimated rewards and destabilize training. Second, TD3 delays actor network updates relative to critic updates, performing multiple critic gradient steps per actor step to allow the critic to converge to a more accurate value function before it is used to guide policy improvement. Third, TD3 adds truncated normal noise to target policy actions during critic updates, smoothing the target Q-value surface and acting as a regularize that prevents overfitting to narrow spikes in the action space [1].

Yuan and Zou [1] identified TD3 as the appropriate algorithm for their health-aware fast charging task based on the requirement for smooth, bounded charging current actions, the high sample cost of full-lifetime simulations in PyBaMM, and the need for stable long-horizon learning across approximately 2000 training episodes. The training process, executed on a consumer-grade desktop with an NVIDIA RTX 3060 GPU, required approximately 22 hours to converge, confirming feasibility without access to high-performance computing infrastructure. The TD3 agent converged to a cumulative reward near −60, a cut-off voltage near 4.2 V, and a charging time near 27 minutes after approximately 1000 episodes, with stable behavior thereafter.

C. Lifelong Learning: The Critical Innovation

The defining characteristic distinguishing the framework of Yuan and Zou [1] from all prior RL-based charging works is the lifelong training procedure. In episodic training protocols (used in [2], [3], [4]), the battery model is reset to its initial fresh state at the beginning of each episode, exposing the agent exclusively to the electrochemical dynamics of a healthy cell and producing a policy that is near-optimal for fresh batteries but increasingly suboptimal as the cell ages.This mismatch between training distribution and deployment conditions is the fundamental reason episodically trained agents fail to preserve battery longevity over the full-service life. In the lifelong framework, the battery model ages continuously across all training episodes, accumulating SEI layer growth and lithium plating capacity loss organically. Consequently, the TD3 agent encounters the full SoH spectrum from 100% at the beginning of training to 80% at end-of-life in the correct temporal order, learning a policy that generalizes across all aging states rather than overfitting to the fresh-cell condition [1].

The SoH-dependent voltage constraint $V_{ma}\chi$(SoH) derived from CC-COP experiments provides the mechanistic link between observable terminal signals and the latent degradation state. By adjusting the upper voltage boundary as a function of SoH, the reward function directs the agent to apply progressively gentler charging constraints such as the cell ages, mirroring the physiological change in the optimal plating-safe current profile. This mechanism enables the proposed method to achieve 703 EFCs compared to 572 EFCs for CC-CV—a 22.9% lifespan extension—while maintaining a competitive average charge time of 24.12 minutes, placing it favorably relative to CC-CV-V (611 EFCs) and V-constrained RL (664 EFCs) benchmarks [1].

D. Adaptive and Gaussian Process-Based RL

Hao et al. [4] proposed an alternative to physics-based environment models by using Gaussian processes (GPs) to learn a probabilistic surrogate of the battery charging environment from data. The GP captures uncertainty in the environment dynamics and propagates it through the RL value function, enabling the agent to enforce probabilistic safety constraints on terminal voltage and temperature without a first-principles electrochemical model. The primary advantage of this approach is model agnosticism: the controller can in principle be applied to any battery chemistry for which charging data is available. The primary limitation is that GP scalability deteriorates as the dimensionality of the state space grows or the number of data points increases, making it computationally challenging to extend to long-horizon lifelong scenarios where the battery state space evolves continuously over thousands of cycles [4].

VIII. LITERATURE SYNTHESIS: STRUCTURED COMPARISON TABLE

Table I presents a structured synthesis of all twenty-seven reviewed publications, enabling direct cross-study comparison of methodology, contribution, and limitation. The table is ordered by reference number, which follows the primary citation sequence throughout the review body. Abbreviations are defined in the footnote following the table.

Abbreviations — BMS: Battery Management System; CC-COP: Constant Current Constant Overpotential; CC-CV: Constant Current Constant Voltage; CKF: Cubature Kalman Filter; DAE: Differential-Algebraic Equation; DDPG: Deep Deterministic Policy Gradient; DFN: Doyle-Fuller-Newman;

EFC: Equivalent Full Cycle; GP: Gaussian Process; MPC: Model Predictive Control; MTL: Multi-Task Learning; PyBaMM: Python Battery Mathematical Modelling; RL: Reinforcement Learning; SEI: Solid-Electrolyte Interphase;

SoC: State of Charge; SoH: State of Health; SPM: Single Particle Model; SPMe: Single Particle Model with Electrolyte; TD3: Twin Delayed Deep Deterministic Policy Gradient.

TABLE I. STRUCTURED LITERATURE SYNTHESIS: HEALTH-AWARE FAST CHARGING OF LITHIUM-ION BATTERIES (2021–2026)

Ref.	Authors	Year	Method / Model	Key Contribution	Limitation
[1]	Yuan & Zou	2025	TD3 RL + SPMe + Lifelong Training	First lifelong RL formulation for LiB fast charging; 22.9% lifespan extension over CC-CV with comparable 24.12 min charge time	Requires offline three-electrode characterization for V-SoH map; single temperature tested
[2]	Wei et al.	2022	DDPG + Multiphysics Constraints	Multiphysics reward design (thermal + electrochemical); improved safety under dynamic fast charging	Battery model reset per episode; long-term aging dynamics not explicitly modelled
[3]	Park et al.	2022	Deep RL Framework (DDPG/TD3)	Comprehensive RL charging framework; two cases compared	Aging modelled only as film resistance growth; no lifelong training protocol
[4]	Hao et al.	2023	Adaptive GP-based Model RL	Gaussian processes capture uncertainty and enforce constraints online	Focuses on charging time minimization; no explicit health extension
[5]	Lu et al.	2024	MPC + Li Plating Detection + Lifelong Param	Real-time lithium plating detection integrated in MPC; adaptive updates	High computational cost; limits embedded deployment

[6]	Wassiliadis et al.	2023	Model-Based Health-Aware Fast Charging	Overpotential control reduces plating risk while accelerating charging	Requires laboratory-grade reference electrodes
[7]	Zhang et al.	2025	ML Lifelong Li Plating Estimation	ML estimates plating potential across lifetime	Mainly simulation-based; hardware latency untested
[8]	Gao et al.	2024	Model-Based Dataset Recovery	Recovers EV battery datasets considering aging	Limited chemistry validation
[9]	Gao et al.	2021	Parameter Sensitivity	Identifies parameters sensitive to aging	Offline analysis only
			Analysis		
[10]	Li et al.	2023	Nonlinear Model Inversion Control	Fast SoC tracking with electrochemical constraints	Depends on accurate model; aging reduces accuracy
[11]	Hwang et al.	2022	MPC + Enhanced SPM	Minimizes intra-cycle capacity fade	Only short-term optimization
[12]	Huang et al.	2024	MINN (DAE Learning)	Learns battery dynamics from data	Needs further BMS integration
[13]	C. Zhang et al.	2025	Multitask SOH Estimation	Cross-cell knowledge transfer for SOH	Requires large datasets
[14]	C. Zhang et al.	2024	Attention-based SOH Estimation	Accurate SOH using multi-source data	Real-time performance not benchmarked
[15]	C. Chen et al.	2025	Adaptive CKF SoC Estimation	Robust SoC estimation under noise	Issues under extreme temperature
[16]	Wang et al.	2024	DL Li Plating Detection	Real-time onboard plating detection	Latency and false positives not fully addressed
[17]	Zhao et al.	2025	Li Plating Detection (Capacitors)	High-rate plating detection	Different chemistry limits transferability
[18]	Huang et al.	2022	Early Plating Detection	Non-invasive detection in commercial cells	Limited ageing scenarios
[19]	Yuan et al.	2025	MPC for Recycling	Safe battery pretreatment	Not for active charging
[20]	Ahmad et al.	2025	Techno-Economic Analysis	Cost-benefit of ultra-fast charging	No cell-level degradation analysis
[21]	L. Zhang et al.	2023	EV Charging Optimization	Reduces grid peak load by 18%	Ignores battery degradation
[22]	Khalik et al.	2021	DFN Parameter Estimation	Reduces ill-conditioning	Computationally intensive
[23]	Khalik et al.	2021	Model Simplification	Reduces computational complexity	Accuracy loss at high C-rates
[24]	O'Kane et al.	2022	Degradation Modelling Review	Comprehensive degradation	No unified real-time model

				mechanisms	
[25]	Gao et al.	2021	Intercalation vs Plating Study	Reveals plating mechanisms	Single-particle limitation
[26]	Y. Li et al.	2021	Model Reduction Survey	Compares BMS models	No experimental validation

IX. OPEN CHALLENGES AND FUTURE RESEARCH DIRECTIONS

A. Hardware Validation and Embedded Deployment

The overwhelming majority of reviewed works validate their charging controllers within high-fidelity simulation environments—primarily PyBaMM with SPMe-aging models—rather than on physical battery hardware [1], [2], [3], [4]. While simulation validation is a necessary first step, the translation of RL charging policies to real cells introduces several non-trivial challenges: sensor noise in voltage and current measurements degrades the quality of state inputs to the policy network; real battery cells exhibit cell-to-cell variability not captured by a single parameterized model; and the latency of neural network inference on automotive microcontrollers must remain within the millisecond-scale control update period [1], [15]. Yuan and Zou [1] argued that their TD3 controller's robustness to moderate sensor noise—attributable to target-policy smoothing and twin-critic minimization—makes it compatible with production BMS hardware, but experimental demonstration on an actual LG M50 or equivalent cell remains an important validation step for the community.

B. Multi-Cell Pack-Level Generalization

All reviewed works characterize individual cell charging behaviour; none addresses the complexities introduced by series-parallel cell assemblies in traction battery packs. Cell-to-cell SoH variability within a pack means that a uniform charging current distributed across series-connected cells will stress the weakest cell disproportionately, accelerating its degradation relative to healthier cells and degrading overall pack capacity [21]. Pack-level health-aware charging requires either per-cell current control through active balancing circuitry or intelligent scheduling of bypass or redistribution switches. Extending the lifelong RL framework to multi-cell environments where the state space scales with pack size and SoH variability is an open and high-priority research challenge.

C. Temperature and Chemistry Generalisation

Yuan and Zou [1] noted that the $V_{m}{}^{5}{}_{to\varphi}{}^{i}$-SoH mapping is temperature- and chemistry-dependent, demonstrating through PyBaMM simulations that the mapping shifts measurably between 25°C and 30°C and differs between LG M50 (NMC 811) and NMC 532 cells. For practical deployment in EVs operating across a −20°C to +45°C temperature range, a condition-aware prediction model for $V_{m}{}^{5}{}_{to\varphi}{}^{i}$ that takes SoH, temperature, and chemistry as inputs is required. Transfer learning—specifically, adapting a pre-trained mapping or policy from one cell type to another with minimal new experimental data—was identified by Yuan and Zou [1] as a promising direction for reducing the experimental characterization burden. Zhang et al. [7] and Zhang et al. [13], [14] demonstrated the viability of transfer learning for lifetime estimation and SoH estimation respectively, providing foundational evidence that this direction is technically feasible.

D. Integration with V2G and Second-Life Applications

The lifetime extension achievable through health-aware fast charging is particularly valuable in the context of vehicle-to-grid (V2G) applications, where batteries serve dual roles as EV range providers and grid ancillary service assets, accumulating additional degradation from bidirectional power flows [20], [21]. Ahmad et al. [20] analyzed the techno-economic implications of ultra-fast charging station integration with grid-scale storage, noting that battery degradation cost is the dominant variable in economic feasibility assessments. A charging controller that simultaneously manages degradation under fast-charging and V2G cycling conditions would substantially improve the economic case for combined EV-grid service deployments. Similarly, batteries transitioning from first-life EV service to second-life stationary storage retain residual capacity that health-aware charging during first life can maximize. Yuan et al. [19] addressed the end-of-life boundary through recycling pretreatment control, but the continuum from in-service health-aware charging through second-life deployment represents a systems-level challenge that no single reviewed work addresses holistically.

E. Real-Time SoH Estimation for Dynamic Constraint Updating

The lifelong TD3 framework requires SoH information to dynamically update the $V_{ma}\chi$(SoH) constraint. In simulation, SoH is directly computable from the electrochemical model state. In hardware deployment, SoH must be estimated from measurable signals using algorithms like those proposed by Zhang et al. [13], [14]. The accuracy and latency of onboard SoH estimation directly determine the quality of the dynamic voltage constraint and, consequently, the effectiveness of the health-aware charging policy. Developing co-designed SoH estimators and charging controllers—where the estimator uncertainty is propagated into the constraint update to ensure safety under estimation error—represents a technically deep and practically important research agenda [8], [9].

X. CONCLUSION

This review has synthesized twenty-seven IEEE-indexed and high-impact peer-reviewed publications spanning 2021 to 2026, delivering a comprehensive critical evaluation of the research landscape for health-aware fast charging of lithium-ion batteries through reinforcement learning and related intelligent control paradigms. The principal conclusions are as follows.

First, conventional CC-CV charging is electrochemically insufficient for managing battery degradation over the full-service life, as its fixed parameter boundaries fail to adapt to progressive SoH decline. Variants that dynamically adjust cut-off voltage via precomputed SoH mappings yield modest improvements (6.82% EFC extension) but cannot fully capture nonlinear multi-mechanism aging interactions [1], [6].

Second, CC-COP control represents the current performance ceiling for health-aware charging (1316 EFCs in the slow variant), but its dependence on direct side-reaction overpotential measurement—unavailable in commercial two-electrode cells—limits its practical deplorability. Machine learning-based surrogates for plating potential estimation represent the most promising route to democratizing overpotential-based control [7], [16], [18].

Third, model predictive control achieves health-aware charging through explicit electrochemical constraint incorporation and real-time adaptive parameter estimation, as demonstrated by Lu et al. [5] and Hwang et al. [11], but the computational burden of coupled electrochemical-degradation optimization remains a significant barrier to automotive embedded deployment [22], [23].

Fourth, deep reinforcement learning—particularly the TD3 algorithm trained through a lifelong protocol on a continuously aging SPMe simulation—achieves a compelling balance between charging speed (24.12 min average) and battery longevity (703 EFCs, 22.9% improvement over CC-CV) without requiring overpotential sensors. The SoH-dependent dynamic voltage constraint, derived offline from CC-COP characterization experiments, provides the essential mechanistic link between observable terminal signals and latent degradation state [1].

Fifth, the most critical open challenges are hardware experimental validation, multi-cell pack-level generalization, temperature and chemistry-robust policy transfer, co-design of SoH estimation and charging control under estimation uncertainty, and integration with V2G and second-life deployment scenarios. Addressing these challenges will require sustained interdisciplinary collaboration spanning electrochemistry, control theory, and machine learning, supported by open-access datasets and reproducible simulation platforms such as MATLAB R2024a.

REFERENCES

[1] M. Yuan and C. Zou, “Lifelong reinforcement learning for health-aware fast charging of lithium-ion batteries,” IEEE Trans. Transp. Electrify., Oct. 2025, doi: 10.1109/TTE.2025.3625421.

[2] Z. Wei, Z. Quan, J. Wu, Y. Li, J. Pou, and H. Zhong, “Deep deterministic policy gradient-DRL enabled Multiphysics-constrained fast charging of lithium-ion battery,” IEEE Trans. Ind. Electron., vol. 69, no. 3, pp. 2588–2598, 2022.

[3] S. Park et al., “A deep reinforcement learning framework for fast charging of Li-ion batteries,” IEEE Trans. Transp. Electrify., vol. 8, no. 2,pp. 2770–2784, 2022.

[4] Y. Hao, Q. Lu, X. Wang, and B. Jiang, “Adaptive model-based reinforcement learning for fast charging optimization of lithium-ion batteries,” IEEE Trans. Ind. Inform., 2023.

[5] Y. Lu et al., “Health-aware fast charging for lithium-ion batteries: Model predictive control, lithium plating detection, and lifelong parameter updates,” IEEE Trans. Ind. Appl., 2024.

[6] N. Wassiliadis et al., “Model-based health-aware fast charging to mitigate the risk of lithium plating and prolong the cycle life of lithium-ion batteries in electric vehicles,” J. Power Sources, vol. 561, p. 232586, 2023.
[7] Y. Zhang et al., “Machine learning-based lifelong estimation of lithium plating potential: A path to health-aware fastest battery charging,” Energy Storage Mater., vol. 74, p. 103877, 2025.
[8] Y. Gao et al., “A model-based battery dataset recovery method considering cell aging in real-world electric vehicles,” IEEE Trans. Ind. Inform., vol. 20, no. 5, pp. 7904–7914, 2024.
[9] Y. Gao et al., “Global parameter sensitivity analysis of electrochemical model for lithium-ion batteries considering aging,” IEEE/ASME Trans. Mechatronics, vol. 26, no. 3, pp. 1283–1294, 2021.
[10] Y. Li et al., “Nonlinear model inversion-based output tracking control for battery fast charging,” IEEE Trans. Control Syst. Technol., vol. 32, no. 1, pp. 225–240, 2023.
[11] G. Hwang et al., “Model predictive control of lithium-ion batteries: Development of optimal charging profile for reduced intracycle capacity fade using an enhanced single particle model with first-principled chemical/mechanical degradation mechanisms,” Chem. Eng. J., vol. 435, p. 134768, 2022.
[12] Y. Huang et al., “MINN: Learning the dynamics of differential-algebraic equations and application to battery modeling,” IEEE Trans. Pattern Anal. Mach. Intell., 2024.
[13] C. Zhang et al., “A CMMOG-based lithium-battery SOH estimation method using multitask learning framework,” J. Energy Storage, vol. 107, p. 114884, 2025.
[14] C. Zhang et al., “Estimating lithium-ion battery health using hybrid attention networks and multi-source data,” IEEE Sens. J., 2024.
[15] C. Chen et al., “State of charge estimation for lithium-ion batteries using an adaptive cubature Kalman filter based on improved generalized minimum error entropy criterion,” Green Energy Intell. Transp., p. 100292, 2025.
[16] H. Wang et al., “Onboard in-situ warning and detection of Li plating for fast-charging batteries with deep learning,” Energy Storage Mater., vol. 71, p. 103585, 2024.
[17] S. Zhao et al., “Lithium plating accurate detection of lithium-ion capacitors upon high-rate charging,” Green Energy Intell. Transp., p. 100268, 2025.
[18] W. Huang et al., “Onboard early detection and mitigation of lithium plating in fast-charging batteries,” Nat. Commun., vol. 13, no. 1, p. 7091, 2022.
.[19] M. Yuan, A. Burman, and C. Zou, “Robust model predictive control of fast lithium-ion battery pretreatment for safe recycling,” arXiv preprint arXiv:2503.11857, 2025.
[20] A. Ahmad et al., “Techno-economic analysis of energy storage systems integrated with ultra-fast charging stations: A Dutch case study,” transportation, p. 100411, 2025.
[21] L. Zhang et al., “Charging and discharging optimization strategy for electric vehicles considering elasticity demand response,” transportation, vol. 18, p. 100262, 2023.
[22] Z. Khalik et al., “Parameter estimation of the Doyle–Fuller–Newman model for lithium-ion batteries by parameter normalization, grouping, and sensitivity analysis,” J. Power Sources, vol. 499, p. 229901, 2021.
[23] Z. Khalik, M. Donkers, and H. J. Bergveld, “Model simplifications and their impact on computational complexity for an electrochemistry-based battery modeling toolbox,” J. Power Sources, vol. 488, p. 229427, 2021.
[24] S. E. O’Kane et al., “Lithium-ion battery degradation: how to model it,” Phys. Chem. Chem. Phys., vol. 24, no. 13, pp. 7909–7922, 2022.
[25] T. Gao et al., “Interplay of lithium intercalation and plating on a single graphite particle,” Joule, vol. 5, no. 2, pp. 393–414, 2021.
[26] Y. Li et al., “Model order reduction techniques for physics-based lithium-ion battery management: A survey,” IEEE Ind. Electron. Mag., vol. 16, no. 3, pp. 36–51, 2021.