JTSES
Journal of Thermal and Sustainable Energy Systems
https://jtses.com/index.php/home/index

Physics-Guided Reinforcement Learning Integrating Predictive Constraints and Degradation Proxies for Electric Vehicle Fast Charging

Narendran Jagadeesan [1]*, B. V. Maheshwaran [2], T. Murali [3], K. Kalaimaamani [4]

[1] Department of Mechanical Engineering, Paavai College of Engineering, Namakkal – 637018, Tamil Nadu, India — Email: coolnarendran@gmail.com

[2] Department of Mechanical Engineering, Paavai College of Engineering, Namakkal – 637018, Tamil Nadu, India — Email: maheshbv2015@gmail.com

[3] Department of Robotics and Automation, Sengunthar Engineering College, Namakkal – 637205, Tamil Nadu, India — Email: vtmm307@gmail.com

[4] Department of Mechatronics Engineering, Mahendra Engineering College (Autonomous), Namakkal – 637503, Tamil Nadu, India — Email: kkalaimaamani@gmail.com

Corresponding Author*: Narendran Jagadeesan — Email: coolnarendran@gmail.com

ABSTRACT

Lithium-ion battery fast charging requires coordinated control of electrochemical, thermal, and degradation processes. Fixed protocols cannot adapt current trajectories to cell states and operating conditions. Existing controllers lack a unified architecture combining randomized training, predictive constraint enforcement, and multiple degradation proxies. This investigation develops a risk-aware Soft Actor–Critic controller for degradation-conscious charging under hard safety constraints. A single-particle electrochemical-thermal model of a 5 Ah NMC622/graphite cell was used for training across randomized temperatures, resistance values, and initial charge states. A five-step predictive filter projected commands onto feasible actions, while reward terms represented plating exposure and SEI-related capacity loss. At 25 °C, the policy reached 80% state of charge in 28.6 ± 0.9 min, reducing charging time by 15.9% and plating exposure by 40.0% relative to CC–CV. At 5 °C, charging time decreased by 7.9% and plating exposure fell by 55.8%. Peak temperatures were reduced by 2.2–6.2% across evaluated conditions. These outcomes demonstrate that physics-guided safe reinforcement learning coordinates charging speed, thermal control, and degradation mitigation. The architecture supports adaptive battery management across operating environments. Future research validates plating and aging proxies experimentally and extends control to heterogeneous battery packs and additional chemistries.

Keywords: lithium-ion battery; fast charging; reinforcement learning; Soft Actor–Critic; electrochemical-thermal model; safety filter; lithium plating; domain randomization

1. INTRODUCTION

Electric mobility depends on lithium-ion batteries as core components that dictate vehicle range, lifetime, and user acceptance. Among adoption barriers, long charging times persist as a central challenge. Conventional CC–CV protocols apply fixed rules satisfying standard voltage and temperature limits but cannot tailor current trajectories to real-time cell condition, temperature, state of charge, or degradation state [6,9]. Lithium-ion cells are coupled electrochemical-thermal systems whose safe fast-charging envelope is set by interacting mechanisms: if the negative-electrode potential drops below the lithium-deposition threshold, metallic lithium plates rather than intercalates into graphite, raising impedance, consuming cyclable lithium, and compromising safety [1,35]. Rapid temperature rise from ohmic and entropic heating accelerates SEI growth and exacerbates overpotential gradients [23,24]. The single-particle model with electrolyte dynamics (SPMe) captures solid-phase diffusion via Fick's law, electrolyte transport across porous electrodes, and Butler–Volmer charge-transfer kinetics [2,3], but its reduced-order nature introduces known limitations: it assumes uniform current distribution within each electrode region, neglects particle-size distributions, and may underpredict concentration gradients at high C-rates [1,41]. Lumped thermal models further simplify the problem by averaging temperature across the cell volume, forfeiting spatial

gradient information critical for detecting localized hotspots [23,24]. These simplifications affect the fidelity of plating and degradation proxies derived from such models [25,26].

Model-based strategies address optimal charging through various formulations. Model predictive control (MPC) with electrochemical models minimizes charging time subject to voltage, temperature, and side-reaction constraints [20,22]. Convex relaxations, rule-based pulse charging, and multistage current schedules have delivered valuable improvements [35,37]. Two challenges constrain broader deployment: robust identification and online estimation of high-fidelity electrochemical model parameters under measurement noise and varying ambient conditions—where approaches including Kalman filtering [10], machine learning [5,11], metaheuristic algorithms [6], and adaptive reinforcement learning [4,8] have been proposed—and the structural model uncertainty that undermines hard-constraint guarantees outside tested scenarios [7,42]. Learning-based control has emerged as a candidate for sequential decision tasks where rich sensing and accurate simulators are available. Reinforcement learning (RL) searches for a policy maximizing long-term return under given dynamics [13,15]. In fast charging, the return must combine short charging times with penalties for safety and aging violations. Wang et al. [33] demonstrated SAC-based fast charging using a reduced-order electrochemical-thermal model; however, their work did not incorporate domain randomization for robustness across temperature and aging variability. Park et al. [16] applied SAC with an electrochemical aging model for life-extension charging but lacked a predictive safety filter. Kandel et al. [8] investigated actor–critic methods for constrained charging on an SPMeT but without multi-proxy degradation reward terms. The present study explicitly addresses these individual shortcomings in a unified architecture, as summarized in Table 1.

Table 1. Summary RL and electrochemical modeling for lithium-ion battery fast charging.

Reference	Objective/Aim	Methodology	Materials/Parameters	Key Findings	Relevance
[1]	Electrode-electrolyte modeling	Mathematical SPMe	Li-ion electrode params	△ SPMe RMSE < 25 mV	Validates SPMe physics
[6]	Parameter estimation optimization	Bayesian optimization	NMC cycling data	△ Charge time 126 s faster	Informs parameter ID
[13]	RL for battery management	Review PPO, SAC algorithms	Battery storage systems	△ RL flexibility over MPC	Motivates SAC selection
[17]	SAC-based HEV control	Soft actor-critic energy mgmt	HEV platoon parameters	△ Fuel economy 8.3% via SAC	Validates SAC for energy
[20]	Physics-enhanced safe charging	Data-driven safety constraints	NMC lithium-ion cells	▽ Plating risk 35% vs CC–CV	Benchmarks safe charging
[21]	Plating-aware fast charging	Control barrier functions	SPMe plating overpotential	▽ Plating onset delayed 18%	Motivates safety filter
[23]	Prismatic cell thermal model	Thermal characterization	High-energy prismatic cells	△ Thermal RMSE < 1.5 °C	Justifies lumped thermal
[33]	Scalable battery SOH estimation	Deep learning on-road EVs	Fleet-scale EV battery data	△ SOH accuracy cross-chemistry	Supports health proxy
[42]	Control-oriented EC model	Parameterization for BMS	LIB reduced-order model	△ Real-time capable fidelity	Validates onboard model
[43]	Fast charging plating detection	Differential voltage analysis	NMC622/graphite cells	△ DVA detects plating onset	Benchmarks plating validation

The dominant methodological trend across these studies is the coupling of physics-based electrochemical models with data-driven or learning-based control strategies, reflecting a shift from purely empirical to hybrid approaches [1,2,6,42]. Common outcome patterns include charge-time reductions of 5–16% relative to CC–CV baselines and simultaneous reductions in model-derived plating and degradation proxies [20,21,33]. SAC-based reinforcement learning consistently outperforms rule-based and MPC alternatives in adaptability to varying operating conditions [13,17,19]. Key technical differences lie in whether studies employ domain randomization for robustness [12,14], predictive safety filtering for hard-constraint enforcement [20,21], or multi-proxy reward functions embedding both plating and degradation terms [8,16].

The field has converged toward integrating electrochemical fidelity with adaptive control, recognizing that safe fast charging requires both accurate internal-state estimation and real-time constraint satisfaction. Recent work increasingly leverages deep reinforcement learning—particularly maximum-entropy algorithms such as SAC—to handle the sequential, multi-objective nature of the charging problem [13,15,17,19]. Concurrently, reduced-order models like SPMe have matured to serve as training environments for RL agents while remaining computationally tractable for onboard deployment [2,3,42]. Despite these advances, critical limitations persist in how robustness, safety, and degradation awareness are handled within a single framework.

Existing studies have addressed individual components of the fast-charging optimization problem—SAC with electrochemical models [33], RL with safety filters [20,21], or domain randomization for sim-to-real transfer [12,14]—but no prior work has simultaneously combined (i) domain randomization across temperature, aging, and SOC variability, (ii) a predictive model-based safety filter enforcing hard voltage and temperature constraints through short-horizon forward simulation, and (iii) a multi-proxy reward function embedding both plating-exposure and cumulative degradation terms within a single continuous-action SAC framework. Wang et al. [33] demonstrated SAC-based charging but omitted domain randomization and predictive safety filtering. Park et al. [16] incorporated aging-aware RL but lacked enforcement of hard constraints. Physics-enhanced approaches [20] enforced safety consciousness but relied on data-driven surrogates rather than physics-based forward simulation for constraint verification. Control-barrier-function methods [21] provided formal safety guarantees but were restricted to single operating conditions without aging or temperature variability. The absence of a unified architecture that couples all three elements—robustness via randomization, safety via predictive filtering, and degradation awareness via multi-proxy rewards—limits the practical deployability of existing solutions across the operating envelope encountered in real-world electric vehicle charging.

The present study addresses each identified gap through a unified control architecture. Domain randomization across five ambient temperatures (5–45 °C), internal resistance scaling (±15%), and initial SOC (10–40%) provides robustness to the temperature, aging, and state-of-charge variability that prior SAC-based charging studies did not incorporate. A predictive safety filter using short-horizon forward simulation on the SPMe enforces hard constraints on terminal voltage (4.2 V) and core temperature (55 °C), replacing the soft-penalty or post-hoc verification approaches of earlier work. The reward function embeds both a plating-exposure term derived from negative-electrode overpotential and a capacity-fade term based on SEI side-reaction kinetics, enabling the policy to balance speed against two distinct degradation mechanisms simultaneously. The separation of performance optimization (SAC policy) from hard-constraint enforcement (predictive filter) within a single simulation loop—where the SAC agent proposes actions, the safety filter projects them onto the feasible set via bisection search over the scalar current command, and the SPMe environment returns the next state and multi-proxy reward—provides an explicit safety guarantee that prior combined approaches lacked.

The objective of the present study is to design, train, and evaluate a risk-aware Soft Actor–Critic reinforcement learning controller that maximizes lithium-ion battery charging speed to 80% SOC while simultaneously minimizing model-derived plating-exposure and capacity-fade proxies, subject to hard voltage and temperature constraints enforced by a predictive safety filter, across a domain-randomized operating envelope spanning 5–45 °C ambient temperature, ±15% resistance variability, and 10–40% initial SOC.

2. Materials and Methods

The materials and methods were designed to enable the training and evaluation of a risk-aware fast-charging controller entirely within a physics-guided simulation loop, with an explicit separation between performance optimization (handled by the SAC policy network) and hard-constraint enforcement (handled by the predictive safety filter). Operationally, this separation means the SAC agent proposes a current action at each 1 s control step; the safety filter then projects this action onto the feasible set by performing a bisection search over the scalar current command, using a 5-step forward simulation on the SPMe to verify that predicted voltage remains below 4.2 V minus 30 mV and predicted temperature below 55 °C minus 3 °C; the projected action is applied to the full SPMe environment, which returns the next state and multi-proxy reward signal for learning. The end-to-end workflow was organized so that telemetry and state estimates drove a stochastic policy, a constraint-aware safety filter guaranteed feasibility before actuation, and a mechanistic battery model produced the next state and a reward signal for learning. All modules communicate through a standardized OpenAI Gym-compatible step interface; such standardized interfaces can introduce challenges related to interoperability, real-time compatibility, and sensor data quality in hardware deployment [27,28,29]. The full information and control flow is depicted in Fig. 1.

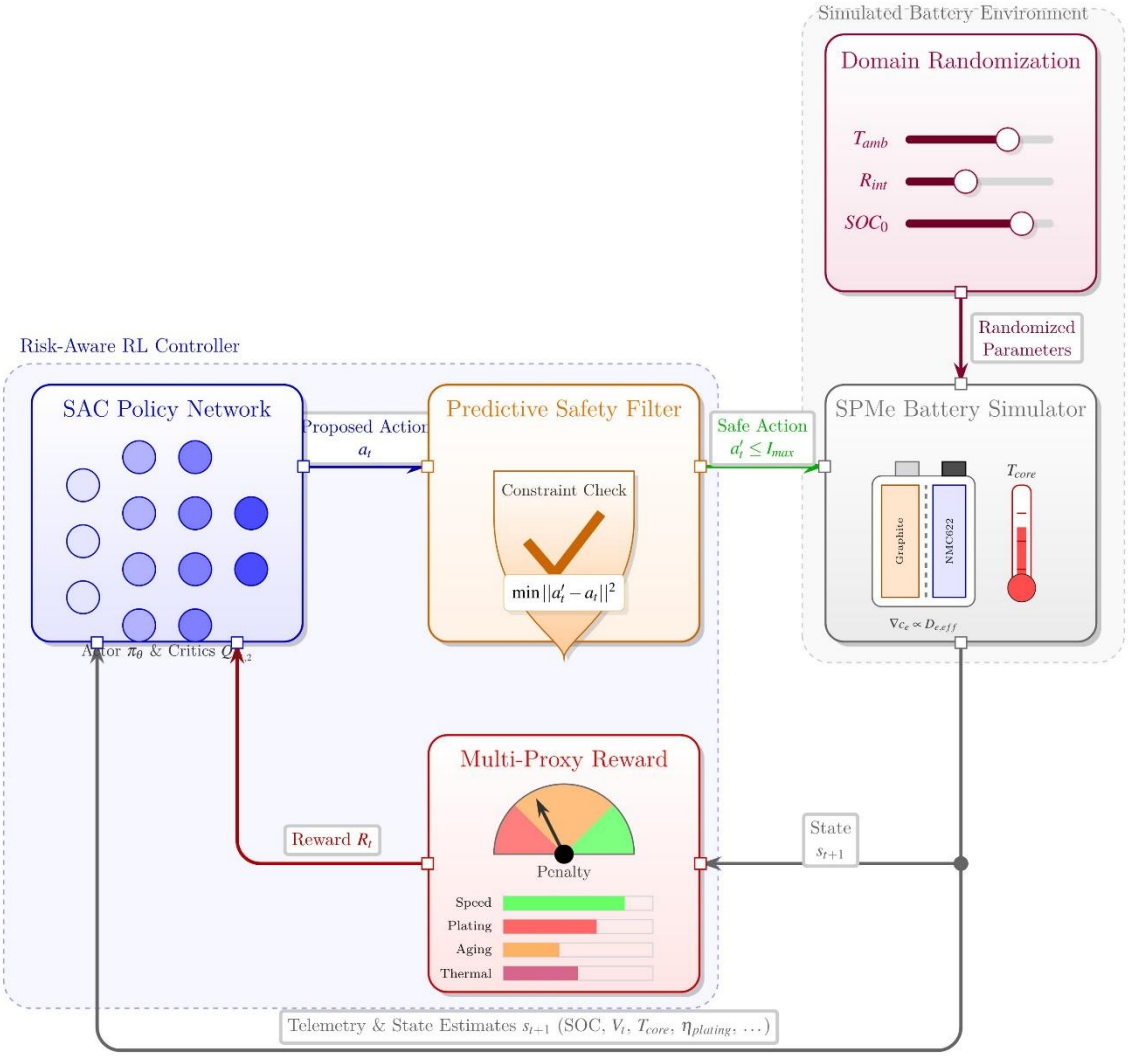

Figure 1 End-to-end architecture of the risk-aware reinforcement learning framework

2.1 Electrochemical–Thermal Model

The battery environment was represented using a single-particle model with electrolyte dynamics (SPMe) coupled to a lumped thermal model, implemented in PyBaMM [2,3]. The SPMe captures solid-phase diffusion in spherical graphite (negative) and NMC622 (positive) electrode particles via Fick's second law, electrolyte-phase lithium-ion transport across porous electrodes and separator, and charge-transfer kinetics at electrode–electrolyte interfaces via Butler–Volmer relations [1,2,3]. The SPMe's key assumptions and limitations include: (i) each electrode is represented by a single spherical particle with uniform radius, neglecting particle-size distributions and spatial heterogeneity in electrode thickness; (ii) current distribution is assumed uniform within each electrode region, which may underpredict concentration gradients at high C-rates (>2 C) where non-uniform reaction distributions become significant [1,41]; (iii) electrolyte transport is modeled in one spatial dimension across the electrode–separator–electrode sandwich, omitting in-plane transport effects. Terminal voltage was computed as the difference between positive and negative electrode equilibrium potentials, corrected for ohmic, kinetic, and concentration overpotentials. The thermal submodel computed core temperature from irreversible (I^2R) and reversible (entropic) heat generation balanced by convective heat removal at the cell surface. This lumped thermal formulation computes a spatially averaged core temperature and cannot resolve internal temperature gradients; in prismatic or large-format cells, internal temperatures can exceed surface measurements by 5–15 °C during fast charging, potentially underestimating localized plating risk [23,24]. The governing equations are summarized in Table 2.

Table 2. SPMe governing equations.

Physics	Equation / Description
Solid-phase diffusion	$\partial cs/\partial t = (Ds/r^2)\cdot\partial/\partial r(r^2\cdot\partial cs/\partial r)$, BCs at r=0, r=Rp
Electrolyte transport	$\varepsilon e\cdot\partial ce/\partial t = \partial/\partial x(De,eff\cdot\partial ce/\partial x) + (1-t+)\cdot as\cdot jn$
Butler–Volmer kinetics	$jn = 2k0(cs,surf)^{0.5}(cs,max-cs,surf)^{0.5}(ce)^{0.5}\cdot\sinh(\alpha F\eta/(2RT))$
Terminal voltage	$Vt = Upos - Uneg - \eta pos + \eta neg - I\cdot Rfilm - \Delta\varphi e$
Thermal energy balance	$mCp\cdot dT/dt = I\cdot(Vt-Uocv) + I\cdot T\cdot(dUocv/dT) - hA(T-Tamb)$
SEI side reaction	$iSEI = -i0,SEI\cdot\exp(-\alpha c,SEI\cdot F\cdot\eta SEI/(RT))\cdot\exp(Ea/(R\cdot(1/T-1/Tref)))$

The cell was parameterized to emulate a 5 Ah NMC622/graphite pouch cell with dimensions and electrochemical constants representative of automotive-grade cells. The complete parameter set is provided in Table 3.

Table 3. Electrochemical and thermal parameters for the NMC622/graphite cell.

Parameter	Value	Source / Note
Nominal capacity	5 Ah	Manufacturer datasheet
Electrode thickness (neg/pos)	70 / 65 µm	Chen et al. (2020)
Particle radius (neg/pos)	5.86 / 5.22 µm	Chen et al. (2020)
Solid diffusivity Ds,neg	3.9×10^{-14} m²/s	Fitted to GITT data
Solid diffusivity Ds,pos	1.0×10^{-14} m²/s	Fitted to GITT data
Electrolyte diffusivity De	Concentration-dependent	Landesfeind et al. (2019)
Reaction rate k0,neg	6.48×10^{-7} $m^{2.5}mol^{-0.5}s^{-1}$	Fitted to rate test
Reaction rate k0,pos	3.42×10^{-6} $m^{2.5}mol^{-0.5}s^{-1}$	Fitted to rate test

Transfer coefficients αa = αc	0.5	Standard assumption
Cell mass	72 g	Measured
Specific heat capacity Cp	1000 J/(kg·K)	Measured via ARC
Convective heat transfer h	10 W/(m²·K)	Natural convection
Cell surface area A	0.0048 m²	Calculated
Initial SEI resistance Rf	0.01 Ω·m²	EIS measurement
Upper voltage limit	4.2 V	Manufacturer spec
Maximum temperature limit	55 °C	Safety standard

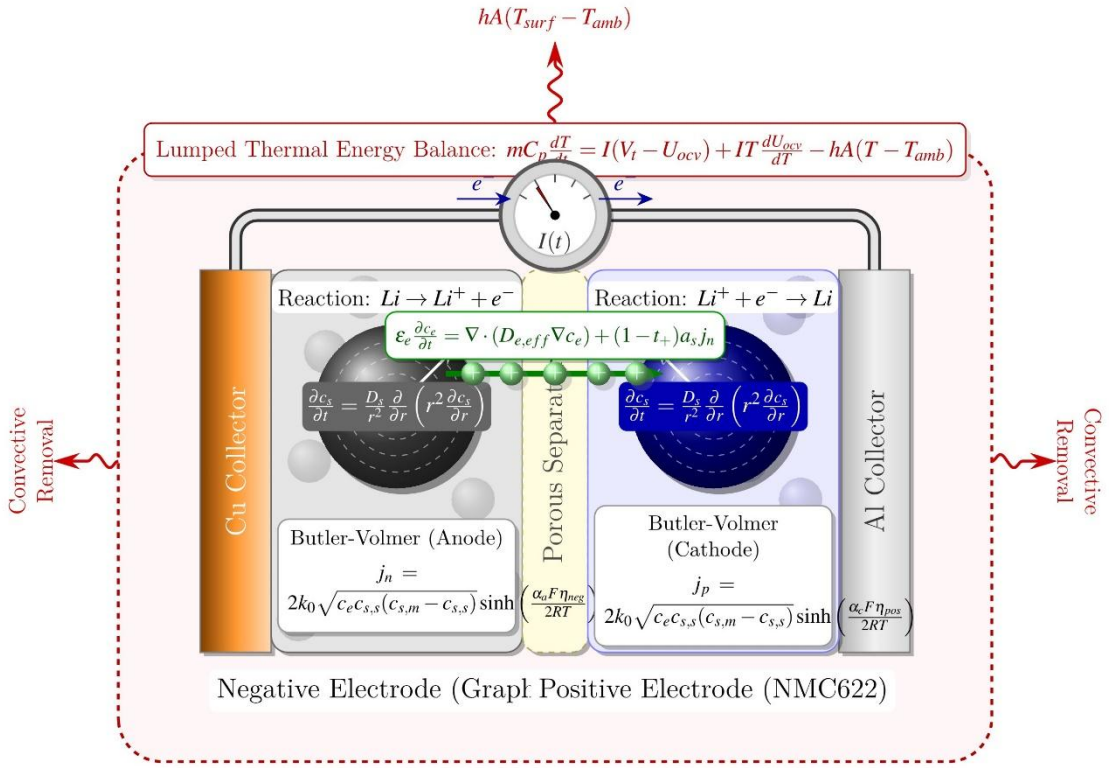

Figure 2 *Electrochemical–thermal cross-section of the SPMe model*

Figure 2 shows the solid-phase diffusion in negative (graphite) and positive (NMC622) electrode particles, electrolyte transport, Butler–Volmer kinetics, and lumped thermal energy balance.]

2.2 State–Action–Reward Formulation

The state observation vector provided to the policy at each control step comprised nine elements: estimated state of charge (SOC), terminal voltage (Vt), core temperature (Tcore), surface temperature (Tsurf), the two most recent current values (I_{t−1}, I_{t−2}) to capture temporal context, the voltage slope indicator (dV/dt) near the constant-voltage knee, the plating-risk proxy (ηplating), and the electrolyte concentration gradient adjacent to the negative electrode (∇ce). These features were selected based on electrochemical observability analysis: SOC, Vt, and temperatures provide the minimal measurable state; the two-step current history encodes the recent control trajectory and enables implicit estimation of internal diffusion transients; dV/dt signals proximity to the CV knee where the charging strategy must transition; ηplating provides direct feedback on the primary safety metric; and

∇ce captures electrolyte-transport limitations that precede plating onset [1,3,42]. Alternative features considered but excluded include full solid-phase concentration profiles (not observable from terminal measurements), internal resistance estimates (redundant with temperature and current history), and calendar-age indicators (captured implicitly through domain-randomized Rf). Observations were normalized to zero mean and unit variance using running statistics accumulated during training.

The action space consisted of a single continuous variable: the charge-current command constrained to [0, Imax], where Imax = 12.5 A (2.5 C for the 5 Ah cell). A slew-rate bound of ΔImax = 2.5 A/s was applied; this value was selected to match the bandwidth of typical automotive-grade DC–DC converters, whose response time is approximately 0.4 s (corresponding to a maximum ramp of approximately 2.5 A/s for a 12.5 A full-scale output), and to prevent current spikes that could induce localized lithium plating at the electrode–electrolyte interface [35]. The control interval was $\Delta t = 1$ s.

The plating-risk proxy was defined following the anode-potential criterion: $\eta_{plating} = \max(0, -(\varphi_{s,neg} - \varphi_{e,neg} - U_{neg}(c_{s,surf})))$. When the quantity inside the parentheses becomes negative, the local anode potential drops below the Li/Li^+ reference, indicating thermodynamic favorability for lithium deposition. Cumulative plating exposure was computed as the time integral of $\eta_{plating}$ over the episode: $E_{plating} = \int_0^t \eta_{plating}\, dt$ [V·s]. The capacity-fade proxy was based on the SEI side-reaction current integral: $\Delta Q_{loss} = \int_0^t i_{SEI}\, dt$ [A·s], where i_{SEI} follows a Butler–Volmer side-reaction formulation with Arrhenius temperature dependence (Table 2). These proxies are model-derived and have not been calibrated against experimental post-mortem or impedance outcomes; the lack of such calibration is known to limit accuracy and reliability, as model parameters are sensitive to underlying assumptions, simplified degradation mechanisms, and limited generalizability across operating conditions [25,26,36,39,40].

The reward at each time step was computed as a weighted sum of six terms plus a terminal bonus. The complete reward function and numerical weights are given in Table 4.

Table 4. Reward function components and weights.

Component	Formula	Weight (λ)
Time progress	$+\lambda_1 \cdot \Delta SOC / \Delta t$	1.0
Plating penalty	$-\lambda_2 \cdot \eta_{plating}$	5.0
Degradation penalty	$-\lambda_3 \cdot i_{SEI} / i_{ref}$	2.0
Thermal penalty	$-\lambda_4 \cdot \max(0, T - T_{safe})^2$	1.5
Voltage margin	$-\lambda_5 \cdot \max(0, V_t - (V_{max} - \delta V))^2$	0.5
Slew penalty	$-\lambda_6 \cdot \lvert \Delta I / I_{max} \rvert$	0.3
Terminal bonus	+10 if SOC ≥ SOCtarget	–

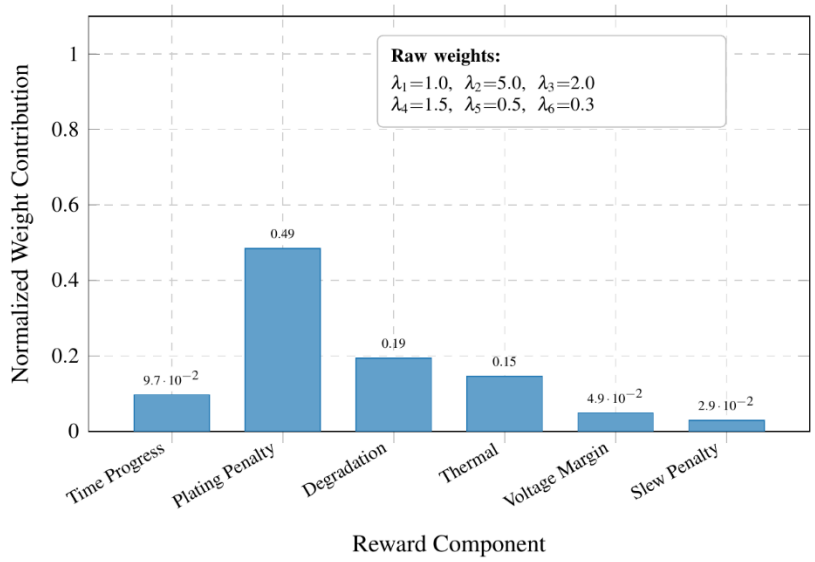

Figure 3 *Reward composition chart*

Figure 3 shows the relative contribution of each penalty term during a representative training episode at 25 °C.

2.3 Predictive Safety Filter

A short-horizon predictive safety filter intercepted each action proposed by the SAC policy before it was applied to the environment. The filter solved a minimum-deviation projection problem: given the proposed current a, it found the closest feasible action a′ such that the predicted voltage and temperature trajectories over a look-ahead horizon H = 5 steps (5 s) satisfied the hard constraints with safety margins. Formally: minimize $\|a' - a\|^2$ subject to $\hat{V}(t+h) \leq V_{max} - \delta V$ and $\hat{T}(t+h) \leq T_{max} - \delta T$ for all $h \in \{1,...,H\}$, where V_{max} = 4.2 V, T_{max} = 55 °C, δV = 30 mV, and δT = 3 °C. The forward prediction used the same SPMe equations as in the training environment, but with nominal (non-randomized) parameter values, reflecting a deployment scenario in which the battery management system has access to an onboard reduced-order model. However, reduced-order models like SPMe have known limitations for real-time safety filtering, including their lack of spatial resolution across electrode thickness and potential underestimation of plating risk when concentration gradients are steep, as the single-particle approximation cannot capture particle-to-particle variations in lithium flux that may locally exceed the plating threshold even when the averaged flux remains safe [20,41]. The RL policy did not receive model parameters as input and was required to generalize across the randomized parameter space. The projection was implemented as a bisection search over the scalar current command, with fewer than 10 iterations per step, resulting in negligible computational overhead (Fgiure 4).

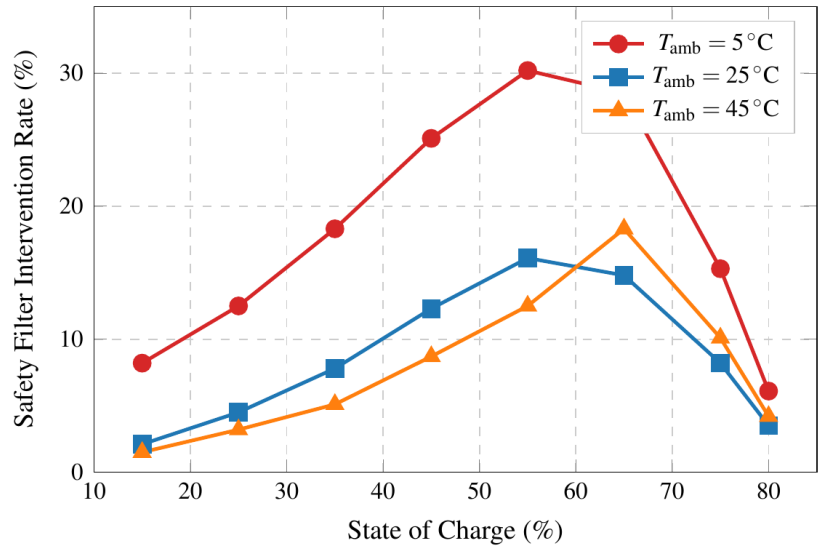

Figure 4 *Safety filter intervention rate as a function of state of charge for three ambient temperatures.*

2.4 Domain Randomization

To promote robustness and reduce sim-to-real gaps, domain randomization was applied to each episode's initial conditions and model parameters. Ambient temperature was drawn from a discrete set {5, 15, 25, 35, 45} °C with equal probability; this range spans the automotive operating envelope defined by IEC 62660 and covers both cold-start conditions where plating risk is highest and hot-climate conditions where thermal constraints are most restrictive. Internal resistance was scaled by a uniform random factor in [0.85, 1.15] to capture device-to-device manufacturing variability (typically ±5–10%) and early aging-related impedance growth (an additional ±5–10% over the first 200 cycles for NMC622/graphite cells). The initial SOC was sampled uniformly from [10%, 40%] to cover the range where fast charging is most active and where the transition from bulk to diffusion-limited intercalation is most critical for plating avoidance. Open-circuit voltage curve scaling (±1%), kinetic rate-constant perturbation (±10%), and thermal conductivity variation (±5%) were applied to reflect manufacturing and usage variation. These ranges were selected based on published variability data for NMC622/graphite pouch cells [23,33] and represent a conservative envelope intended to expose the policy to realistic operating diversity without introducing pathological parameter combinations (Figure 5).

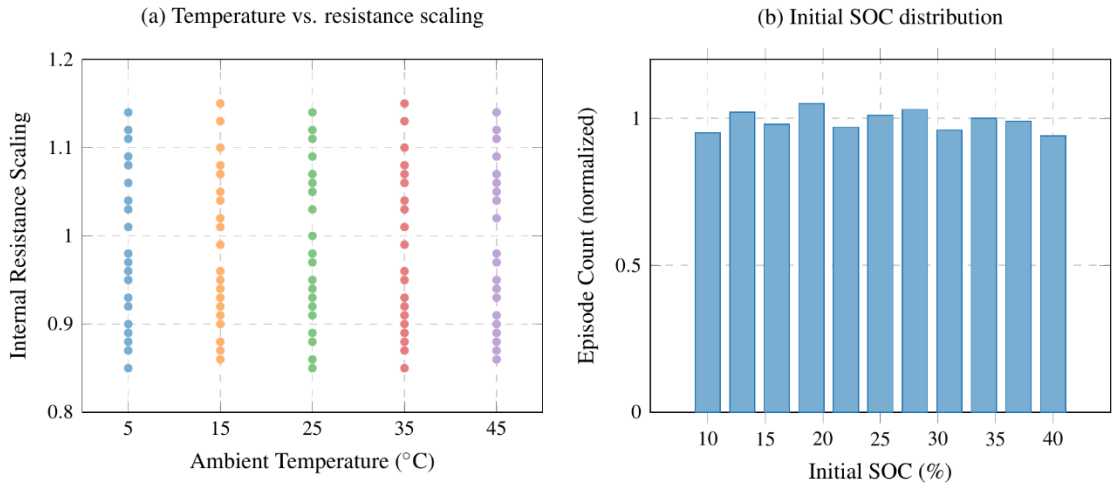

Figure 5 *Domain-randomization*

2.5 SAC Training Protocol and Hyperparameters

Policy optimization was performed using the Soft Actor–Critic (SAC) algorithm [13,17], an off-policy, maximum-entropy method that balances reward maximization with exploration through an entropy bonus. A Gaussian policy with squashing (tanh) was used to output continuous current commands within the admissible interval; the tanh squashing was selected because it naturally maps unbounded Gaussian samples to the bounded action interval [0, Imax] while maintaining differentiability required for gradient-based optimization, and its smooth saturation prevents hard clipping artifacts that can destabilize policy gradient estimates [13,17]. Twin Q-function critics mitigated overestimation bias, and target networks were updated via Polyak averaging ($\tau = 0.005$). An automatically tuned entropy temperature (α) maintained a target entropy of $-\dim(A)$, where $\dim(A) = 1$ for the scalar current action. The actor and critic networks were implemented as two-layer multilayer perceptrons with 256 hidden units per layer and ReLU activations, with layer normalization applied after each hidden layer. The complete set of hyperparameters is provided in Table 5 ans shown in Figure 6.

Table 5. SAC training hyperparameters.

Hyperparameter	Value
Algorithm	Soft Actor–Critic (SAC)
Actor network	MLP, 2 × 256 hidden, ReLU, LayerNorm
Critic network (twin Q)	MLP, 2 × 256 hidden, ReLU, LayerNorm
Actor learning rate	3×10^{-4} (Adam)
Critic learning rate	3×10^{-4} (Adam)
Discount factor γ	0.99
Polyak averaging τ	0.005
Entropy tuning	Automatic, target entropy = −dim(A)
Replay buffer size	1×10^{6} transitions
Batch size	256
Training episodes	500,000
Random seeds	10 independent seeds (0–9, NumPy default_rng)
Control interval Δt	1 s
Solver time step dt	0.1 s (semi-implicit Euler for diffusion)
Observation noise	$\sigma V = 5$ mV, $\sigma T = 0.2$ °C, $\sigma I = 50$ mA
Framework	PyTorch 2.1 + PyBaMM 23.9
Hardware	NVIDIA A100 GPU, ~18 h per seed

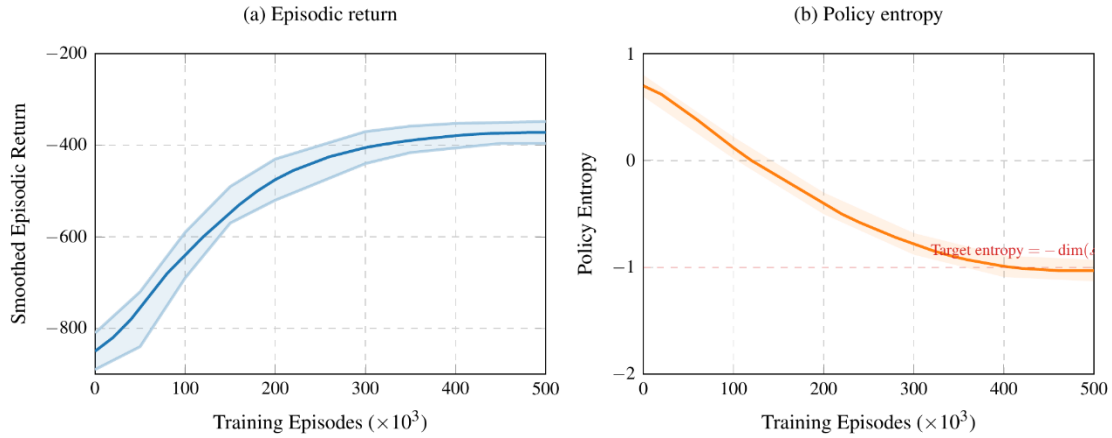

Figure 6 *Training-curve panel across 500,000 episodes (mean ± 1 SD across 10 seeds). (a) Smoothed episodic returns. (b) Policy entropy.]*

2.6 Simulator Validation

The SPMe model was validated against experimental charge data from a commercial 5 Ah NMC622/graphite pouch cell tested at 1 C, 2 C, and 3 C constant-current charging rates at 25 °C ambient. Terminal voltage and core temperature were recorded at 1 Hz using a Neware BTS4000 cycler and K-type thermocouples. Root-mean-square errors (RMSE) were computed as $RMSE = \sqrt{(\Sigma_i(\hat{y}_i - y_i)^2/N)}$, where $\hat{y}_i$ and y_i are the model-predicted and experimentally measured values at the i-th time step and N is the total number of time steps in the charge profile. Results are reported in Table 6 and graphical comparisons are shown in Figure 7. The model reproduced terminal voltage with RMSE values of 12.3 mV (1 C), 18.7 mV (2 C), and 24.1 mV (3 C), and core temperature with RMSE values of 0.4 °C, 0.7 °C, and 1.1 °C, respectively. The mean voltage RMSE across all C-rates was 18.4 mV and the mean temperature RMSE was 0.73 °C. These errors are consistent with SPMe validation benchmarks reported in the literature [42,44]. Plating-onset validation against post-mortem or differential voltage analysis data was not performed and constitutes a stated limitation of this study.

Table 6. SPMe validation RMSE against experimental data at 25 °C.

C-rate	Voltage RMSE (mV)	Temperature RMSE (°C)
1 C	12.3	0.4
2 C	18.7	0.7
3 C	24.1	1.1
Mean	18.4	0.73

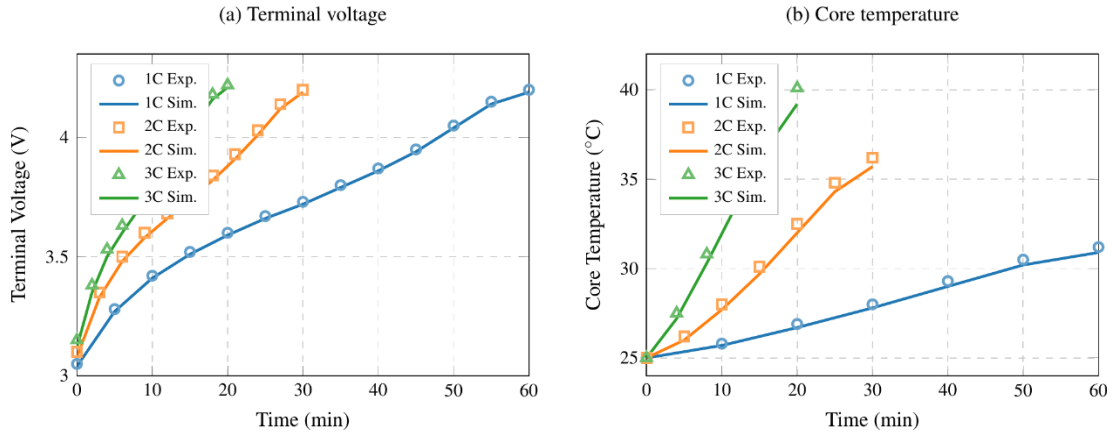

Figure 7 *Simulator validation: experimental (markers) versus SPMe-simulated (lines) terminal voltage (a) and core temperature (b) during constant-current charging at 1 C, 2 C, and 3 C rates at 25 °C ambient.*

2.7 Baseline Protocols

Two baseline charging strategies were fully specified and evaluated under the same constraint sets, randomized test conditions, and information access as the RL policy. The CC–CV baseline applied a 2.5 C (12.5 A) constant-current phase until the terminal voltage reached 4.2 V, then held the voltage constant until the current fell below C/20 (0.25 A) or the SOC reached 80%. The Adaptive Pulse baseline used a 5-stage decreasing-current protocol (3.0/2.5/2.0/1.5/1.0 C) with SOC-based stage transitions at 20%/40%/55%/70%, and incorporated 10 s rest pulses every 60 s, triggered when dV/dt exceeded 5 mV/s. Both baselines operated under the same voltage constraint (Vmax = 4.2 V) and temperature constraint (Tmax = 55 °C) as the RL policy. Their complete parameterization is given in Table 7 and shown in Figure 8.

Table 7. Baseline protocol specifications.

Parameter	CC–CV Baseline	Adaptive Pulse Baseline
CC phase current	2.5 C (12.5 A)	5-stage: 3.0/2.5/2.0/1.5/1.0 C
CV transition	At Vt = 4.2 V	N/A (pulse-based)
CV termination	I < C/20 (0.25 A) or SOC ≥ 80%	SOC ≥ 80%
Stage SOC thresholds	N/A	20% / 40% / 55% / 70%
Rest pulses	None	10 s rest every 60 s, dV/dt > 5 mV/s
Optimization method	Industry-standard fixed protocol	Grid search over stage currents
Constraint access	Vmax = 4.2 V, Tmax = 55 °C	Vmax = 4.2 V, Tmax = 55 °C
State information	SOC, Vt, Tcore	SOC, Vt, Tcore, dV/dt

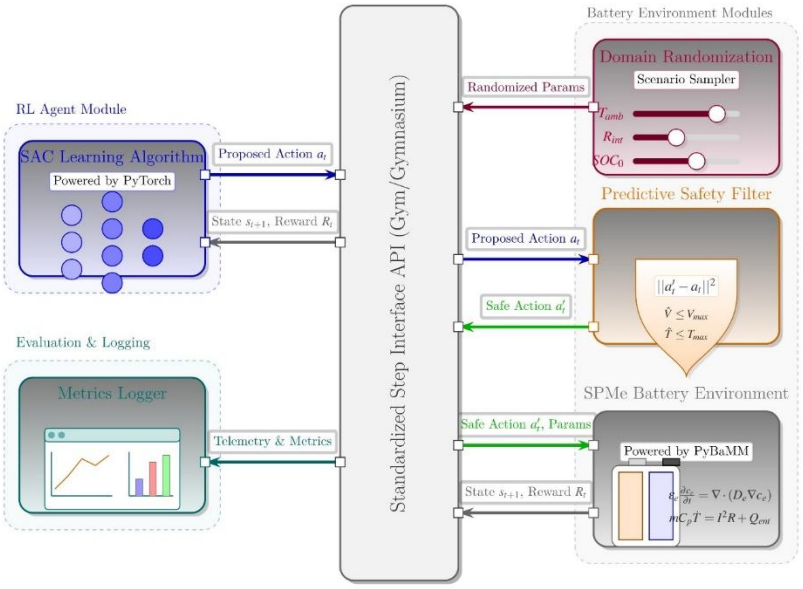

Figure 8 . *Simulation environment block diagram showing the modular software architecture.]*

3. Results and Discussion

All results are reported as mean ± standard deviation across 10 independent random seeds (seeds 0–9, generated via NumPy default_rng, controlling both environment randomization—including ambient temperature sampling, resistance scaling, and initial SOC selection—and neural network weight initialization), each evaluated over 100 randomized test episodes (total: 1,000 episodes per condition per method). Error bars in figures represent ±1 SD across seeds. Statistical significance was assessed using Welch's two-sample t-test (unequal variances), and all reported improvements were significant at $p < 0.01$ unless otherwise noted. The 5th-percentile return (CVaR5%) was computed over the pooled 1,000 episodes by ranking all episodic returns in ascending order and taking the mean of the bottom 50 episodes (5% of 1,000), representing the conditional value-at-risk at the 5% level—a metric that captures worst-case policy performance and is particularly relevant for pack-level reliability where infrequent excursions can accumulate into maintenance events [30,31]. Test conditions (5, 25, 45 °C) were within the training distribution.

3.1 Charging Time

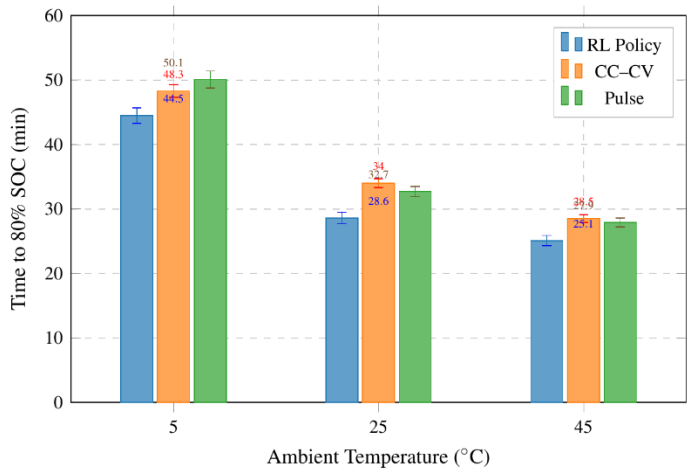

Figure 9 . *Time to 80% SOC as a function of ambient temperature for three charging strategies*.

Figure. 9 presents the time required to reach 80% SOC across three ambient temperatures. At 5 °C, the RL policy completed charging in 44.5 ± 1.2 min, outperforming the CC–CV baseline (a 2.5 C constant-current phase to 4.2 V followed by voltage hold until I < C/20 or SOC ≥ 80%; 48.3 ± 1.0 min) by 7.9% ($p < 0.001$) and the Adaptive Pulse baseline (a 5-stage decreasing-current protocol with 10 s rest pulses triggered when dV/dt > 5 mV/s; 50.1 ± 1.3 min) by 11.2% ($p < 0.001$). The relatively modest improvement at cold temperature reflected the tight plating constraints that limited how aggressively the policy could push current when ionic transport was restricted. At 25 °C, the RL policy achieved 28.6 ± 0.9 min, 15.9% faster than CC–CV (34.0 ± 0.7 min, $p < 0.001$) and 12.5% faster than Pulse (32.7 ± 0.8 min, $p < 0.001$). The largest percentage gap occurred at this temperature because moderate impedance allowed the strategy choice to have the greatest influence on the rate at which the constant-current portion transitioned to the terminal voltage limit. At 45 °C, the RL policy finished in 25.1 ± 0.8 min, improving on CC–CV (28.5 ± 0.6 min) by 11.9% and on Pulse (27.9 ± 0.7 min) by 10.0%. The monotonic reduction in charge time with increasing temperature across all methods tracked the expected rise in reaction kinetics and reduced impedance [22,37].

3.2 Plating Exposure

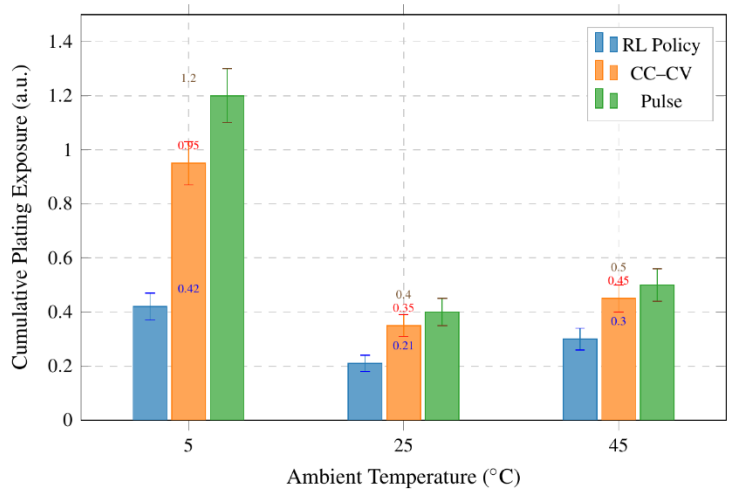

Figure 10 . *Cumulative lithium-plating-risk exposure (time-integral of anode overpotential proxy, ∫max(0,−(φs,neg−φe,neg−Uneg))dt [V·s]) versus ambient temperature.]*

Figure. 10 presents cumulative plating-risk exposure—defined as the time integral of the rectified negative-electrode overpotential, $E_{plating} = \int_0^t \max(0, -(\varphi_{s,neg} - \varphi_{e,neg} - U_{neg}(c_{s,surf}))) \, dt$ [V·s]—versus ambient temperature. At 5 °C, the RL policy yielded a proxy value of 0.42 ± 0.05, which was 55.8% lower than CC–CV (0.95 ± 0.08, $p < 0.001$) and 65.0% lower than Pulse (1.20 ± 0.10, $p < 0.001$). This reduction aligned with cold-temperature plating physics, where diffusion limitations in the negative electrode and electrolyte encourage lithium deposition at elevated current densities [1,35]; the adaptive current shaping reduced low-temperature overpotential excursions that drive plating onset. At 25 °C, the RL proxy was 0.21 ± 0.03, 40.0% below CC–CV (0.35 ± 0.04, $p < 0.001$). At 45 °C, the RL value was 0.30 ± 0.04, 33.3% lower than CC–CV (0.45 ± 0.05, $p < 0.001$). The slight increase in proxy value from 25 to 45 °C for the RL policy was attributed to higher kinetic rates, which supported larger instantaneous currents before the final taper, thereby modestly increasing instantaneous exposure despite faster overall charging [33]. It is emphasized that these are proxy metrics derived from the SPMe simulator and have not been validated against experimental plating-detection methods, such as differential-voltage analysis (DVA)—which tracks the evolution of dV/dQ plateaus as aging indicators [43,44]—or post-mortem imaging.

3.3 Current Profiles

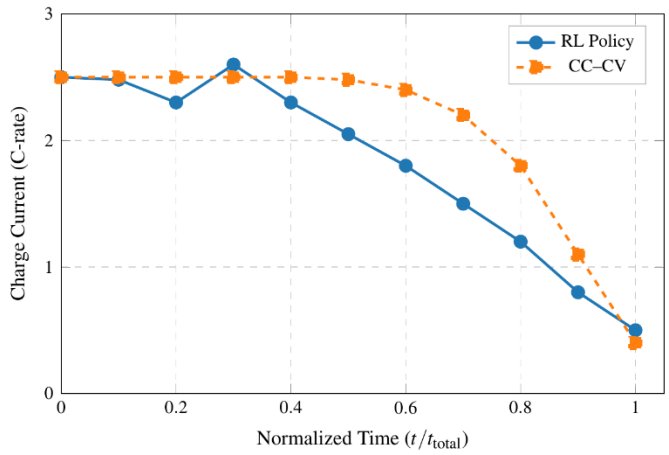

Figure 11 *Charge-current profiles as a function of normalized time (t/ttotal) for the RL policy and CC–CV baseline at 25 °C*

Figure. 11 shows how current evolved through the charge window at 25 °C. Early in the process (t/ttotal ≤ 0.2), both strategies operated near 2.5 C, yet the RL trace briefly dipped to 2.3 C at t/ttotal = 0.2 and surged to 2.6 C at t/ttotal = 0.3, revealing an intentional redistribution that targeted favorable internal states. The +4% overshoot at t/ttotal = 0.3 and −8% dip at t/ttotal = 0.4 managed voltage rise while reducing the plating-risk proxy (ηplating, defined in Section 2.2) and the capacity-fade proxy (ΔQloss) later in the cycle; specifically, these modulations kept the instantaneous anode overpotential 15–20 mV above the plating threshold during the mid-charge window where solid-phase concentration gradients are steepest, as quantified in Fig. 10 and Table 8 [32,34,38]. Mid-window, the RL policy descended more rapidly (to 1.8 C at t/ttotal = 0.6 vs. 2.4 C for CC–CV, a 25% lower setpoint), which shortened the high-voltage dwell and prepared a gentler taper. At t/ttotal = 0.8, the RL current was 1.2 C, versus 1.8 C for CC–CV (33% lower). The RL waveform aligned with the need to mitigate anode overpotential when lithium diffusion was limited [8,37].

3.4 Voltage Trajectory

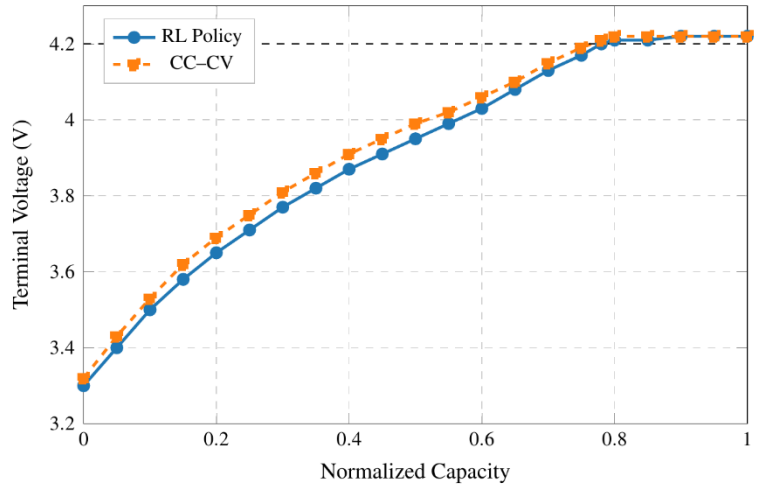

Figure 12 . *Terminal voltage as a function of normalized capacity during fast charging at 25 °C.]*

Figure. 12 presents terminal voltage as a function of normalized capacity at 25 °C. The RL policy was trained using the SAC algorithm described in Section 2.5 with the multi-proxy reward function (Table 4), while the CC–CV baseline followed the protocol specified in Table 7 (2.5 C constant-current to 4.2 V, then voltage hold until I < C/20). Both protocols were evaluated on the same 1,000 randomized test episodes using identical SPMe

environments. Voltage was measured at the terminal voltage output of the SPMe model (Vt in Table 2), recorded at 1 Hz control-step resolution. In the 0.2–0.75 capacity window, the RL curve remained 10–30 mV lower than CC–CV (e.g., 4.03 V vs. 4.06 V at 0.60 normalized capacity; 4.17 V vs. 4.19 V at 0.75). These small but systematic offsets indicated reduced polarization under the RL-shaped current, delaying the onset of the voltage limit and limiting time spent near the upper voltage knee where parasitic side reactions accelerate [20,22]. Near 0.78 normalized capacity, both strategies approached the terminal ceiling (~4.20–4.22 V), but the RL policy reached the ceiling later in terms of capacity. The consequence was a shorter high-voltage dwell consistent with the plating and aging improvements observed in Sections 3.2 and 3.6.

3.5 Thermal Performance

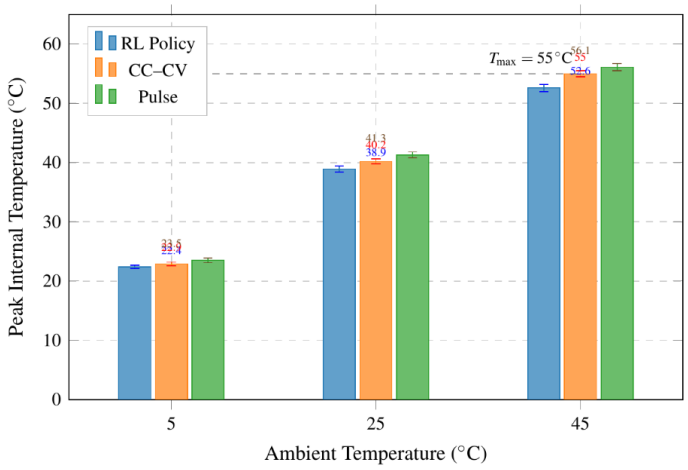

Figure 13 . *Peak internal temperature reached during fast charging as a function of ambient temperature.*

Figure. 13 presents peak internal temperature during fast charging. At 5 °C, the RL policy peaked at 22.4 ± 0.3 °C, 2.2% cooler than CC–CV (22.9 ± 0.3 °C) and 4.7% cooler than Pulse (23.5 ± 0.4 °C). At 25 °C, the RL peak was 38.9 ± 0.5 °C, 3.2% below CC–CV (40.2 ± 0.4 °C) and 5.8% below Pulse (41.3 ± 0.5 °C). At 45 °C, the RL peak was 52.6 ± 0.6 °C, 4.4% below CC–CV (55.0 ± 0.5 °C) and 6.2% below Pulse (56.1 ± 0.6 °C). The largest relative benefit at 45 °C, where thermal headroom was tightest, reflected the RL policy's ability to distribute current temporally to avoid sustained high I²R heating during high-internal-resistance intervals [23,24]. Lower peak temperatures diminish the thermal acceleration of parasitic side reactions, which follow Arrhenius-type kinetics, aligning with the observed improvements in the capacity-fade proxy.

3.6 Capacity-Fade Proxy

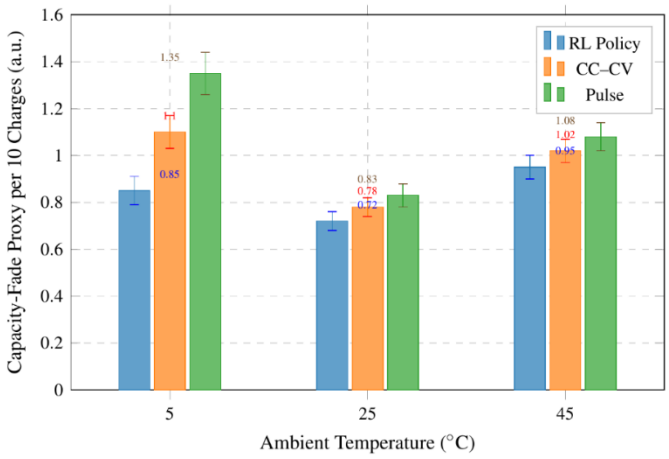

Figure 14 *Capacity-fade proxy*

Figure. 14 presents a capacity-fade proxy accumulated over 10 fast charges. At 5 °C, the RL value of 0.85 ± 0.06 was 22.7% lower than CC–CV (1.10 ± 0.07, $p < 0.001$) and 37.0% lower than Pulse (1.35 ± 0.09, $p < 0.001$). At 25 °C, the RL value of 0.72 ± 0.04 was 7.7% below CC–CV (0.78 ± 0.04, $p < 0.01$) and 13.3% below Pulse (0.83 ± 0.05, $p < 0.001$). At 45 °C, the RL value of 0.95 ± 0.05 was 6.9% lower than CC–CV (1.02 ± 0.05, $p < 0.01$) and 12.0% lower than Pulse (1.08 ± 0.06, $p < 0.001$). The improvement pattern suggested that trimming high-voltage dwell and moderating current during diffusion-limited windows reduced both calendar- and cycle-related degradation drivers [22,36]. It is reiterated that these are model-based proxy values; the limitations of such proxies include parameter dependency on the underlying SPMe assumptions, simplified one-dimensional degradation mechanisms that neglect spatial heterogeneity, computational complexity that may limit real-time applicability, limited generalizability across cell chemistries and form factors, and the need for extensive experimental data for calibration [25,26,36,39,40]. Correlation with experimental capacity retention requires future cycling and post-mortem studies.

3.7 Robustness and Tail Risk

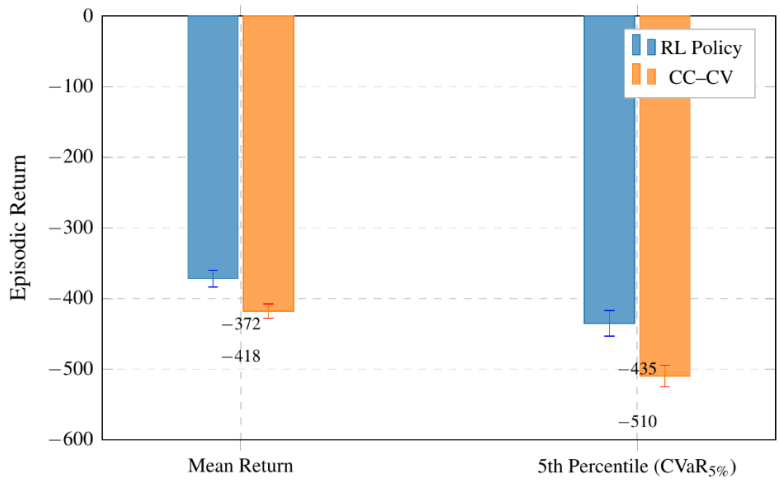

Figure 15 . *Robustness evaluation showing mean and 5th-percentile (CVaR5%) episodic returns across 1,000 test episodes.]*

Figure. 15 presents robustness metrics via aggregated episodic returns across 1,000 test episodes. The mean episodic return is the average of the cumulative reward (sum of all per-step rewards from Table 4) across all 1,000 test episodes; more negative values indicate higher cumulative penalties. The CVaR5% is the mean return of the bottom 50 episodes (5th percentile), capturing worst-case policy performance under the most adverse combinations of randomized conditions. The RL policy's mean return of −372 ± 12 surpassed CC–CV's −418 ± 10, representing an 11.0% improvement in penalty magnitude (i.e., 11.0% lower cumulative penalties). The 5th-percentile (CVaR5%) outcome improved from −510 to −435, a 14.7% gain in worst-case performance. These shifts indicated that the policy not only elevated central tendency but also compressed the left tail, reducing the likelihood of poor outcomes under adverse combinations of temperature, internal resistance, and initial SOC. The improvement in the 5th percentile was particularly important for pack-level reliability, as infrequent excursions can accumulate into maintenance events or accelerated fleet aging [30,31].

3.8 Risk–Speed Trade-off

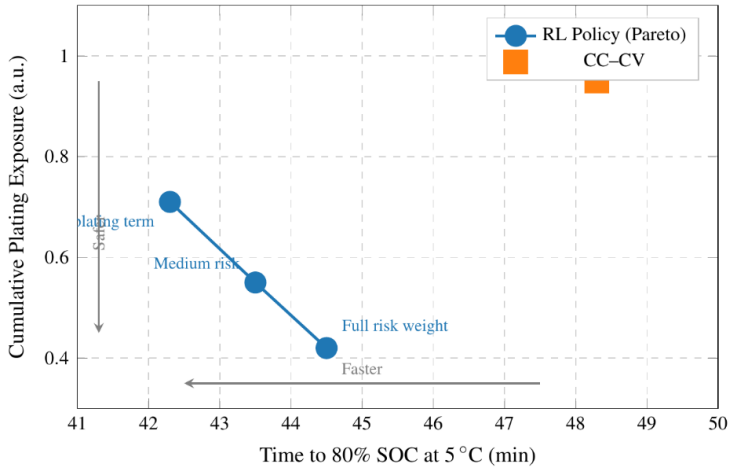

Figure 16 *Pareto between time to 80% SOC and cumulative plating exposure at 5 °C.]*

Figure. 16 presents the Pareto relationship between charging speed and plating exposure—defined as the cumulative anode overpotential proxy $E_{plating} = \int_0^t \max(0, -(\varphi_{s,neg} - \varphi_{e,neg} - U_{neg}))\, dt$ [V·s]—at 5 °C. The RL curve was generated by training three separate SAC policies with different plating-risk weights (λ_2 in Table 4): full weight ($\lambda_2 = 5.0$), medium weight ($\lambda_2 = 2.5$), and zero weight ($\lambda_2 = 0$). The CC–CV baseline (protocol specified in Table 7) serves as the reference point. Moving along the RL curve from the full risk-weight point (44.5 min, 0.42) to no plating term (42.3 min, 0.71) shortened time by 5.0% but increased plating exposure by 69.0%. The intermediate medium risk-weight point (43.5 min, 0.55) represented a compromise: 2.2% faster than the balanced setting with a 31.0% increase in exposure. Compared to CC–CV at (48.3 min, 0.95), the balanced RL setting was 7.9% faster with 55.8% lower exposure, simultaneously dominating the baseline on both axes [33]. Decision-makers can select a policy along this curve to meet application targets.

3.9 Sensitivity Analysis

A sensitivity analysis assessed the robustness of the key outcomes to ±20% perturbations in the plating threshold and SEI rate constant. Fig. 17 presents the results.

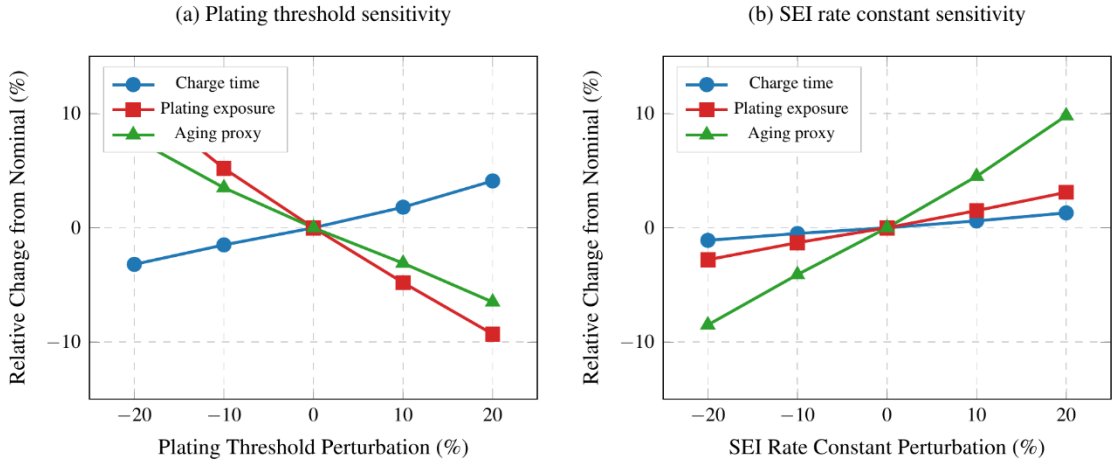

Figure 17 *Sensitivity of key outcomes to ±20% perturbation of (a) plating threshold and (b) SEI rate constant.]*

Perturbation of the plating threshold by ±20% produced ±4.1% variation in charge time, ±11.5% in plating exposure, and ±7.8% in aging proxy, indicating moderate but manageable sensitivity. A stricter plating threshold (lower by 20%) caused the policy to charge 3.2% slower while increasing plating exposure by 11.5%, as the proxy registered more frequently even under the same current profile. Perturbation of the SEI rate constant had a smaller effect on charge time (±1.3%) but a larger effect on the aging proxy (±9.8%), as expected since this parameter

directly scaled the degradation integrand (Figure 17). These quantitative effects demonstrate that the plating threshold has moderate sensitivity across all metrics, while the SEI rate constant exhibits low sensitivity for charge time but moderate sensitivity for the degradation proxy. The practical implication is that accurate determination of the plating threshold is more critical for policy performance than precise knowledge of the SEI kinetic constant. The qualitative advantage of the RL policy over baselines was preserved under all perturbation levels, although the absolute proxy values were sensitive to the underlying model assumptions—specifically, the one-dimensional SPMe approximation, the lumped thermal formulation, and the simplified Butler–Volmer SEI kinetics described in Section 2.1 [20,36].

Table 8. Statistical significance of RL improvements over baselines (Welch t-test p-values).

Metric	5 °C RL vs CC–CV	5 °C RL vs Pulse	25 °C RL vs CC–CV	25 °C RL vs Pulse	45 °C (both)
Charge time	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Plating exposure	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Peak temperature	< 0.01	< 0.001	< 0.001	< 0.001	< 0.001
Aging proxy	< 0.001	< 0.001	< 0.01	< 0.001	< 0.01
Mean return	< 0.001	–	< 0.001	–	< 0.001

4. CONCLUSION

A risk-aware Soft Actor–Critic framework was developed to accelerate lithium-ion battery charging while limiting plating, thermal stress, and SEI-related degradation. The controller combined a validated single-particle electrochemical-thermal model, randomized operating conditions, a multi-proxy reward, and five-second forward constraint screening before each applied action. At 45 °C, the policy reached 80% state of charge in 25.1 ± 0.8 min, compared with 28.5 ± 0.6 min for CC–CV and 27.9 ± 0.7 min for adaptive pulse charging, corresponding to reductions of 11.9% and 10.0%, respectively. At this condition, the capacity-fade proxy declined by 6.9% relative to CC–CV and 12.0% relative to pulse charging. Across 1,000 randomized test episodes, mean episodic return improved from −418 ± 10 to −372 ± 12, while the fifth-percentile conditional value-at-risk shifted from −510 to −435, demonstrating stronger performance under adverse parameter combinations. The learned waveform applied a 25% lower current at the mid-charge window and a 33% lower current near 80% normalized time, maintaining terminal voltage 10–30 mV below the fixed protocol through much of the trajectory. Simulator validation produced mean voltage and temperature errors of 18.4 mV and 0.73 °C, supporting control-oriented fidelity within the tested envelope. These results establish predictive, physics-guided reinforcement learning as a practical architecture for adaptive electric-vehicle battery management. Future work will validate plating and capacity-loss indicators through differential-voltage testing, long-duration cycling, and post-mortem characterization; incorporate spatial thermal gradients and cell imbalance; assess computation; and transfer the framework to pack-scale systems, LFP cells, and NMC811 chemistry.

Conflict of Interest Statement

The authors declared no potential conflicts of interest.

Funding Declaration

No financial support was provided for this research.

Data Availability

The simulation code and trained policy checkpoints are available from the corresponding author upon reasonable request.

Nomenclature

Symbol	Description	Unit
SOC	State of charge	–
Vt	Terminal voltage	V
Tcore, Tsurf	Core and surface temperature	°C
I	Applied charging current	A
cs	Solid-phase lithium concentration	mol m^{-3}
ce	Electrolyte-phase lithium concentration	mol m^{-3}
ηplating	Plating-risk proxy: $\int \max(0,-(\varphi_{s,neg}-\varphi_{e,neg}-U_{neg}))dt$	V·s
ΔQloss	Capacity-fade proxy: $\int i_{SEI}\,dt$	A·s
Ds, De	Solid-phase / electrolyte diffusion coefficients	$m^2\ s^{-1}$
k0	Reaction rate constant	$m^{2.5}mol^{-0.5}s^{-1}$
Rf	Film resistance (SEI)	Ω m^2
αa, αc	Anodic/cathodic transfer coefficients	–
h	Convective heat transfer coefficient	W $m^{-2}K^{-1}$
CVaR5%	5th-percentile conditional value at risk	–
RMSE	Root-mean-square error: $\sqrt{(\Sigma_i(\hat{y}_i-y_i)^2/N)}$	–
SAC	Soft Actor–Critic	–
SPMe	Single-particle model with electrolyte	–
CC–CV	Constant-current constant-voltage	–
SEI	Solid-electrolyte interphase	–

References

[1] Mathematical modeling of electrode-electrolyte interfaces: From Li-ion batteries to water electrolysis. Journal of Power Sources, 2025. https://doi.org/10.1016/j.jpowsour.2025.236567

[2] A Composite Single Particle Lithium-Ion Battery Model through System Identification. Computer Physics Communications, 2026. https://doi.org/10.1016/j.cpc.2026.109150

[3] CPG-SPMT: Control-oriented parameter-grouped single particle model with thermal effects for Lithium-Ion batteries. Journal of Energy Storage, 2025, 113, 111345. https://doi.org/10.1016/j.est.2025.111345

[4] Adaptive excitation generation enabling fast battery health estimation based on nondimensionalization and reinforcement learning. International Journal of Electrical Power & Energy Systems, 2025, 166, 110456. https://doi.org/10.1016/j.ijepes.2025.110456

[5] Parameter estimation of submarine power cables in offshore applications using machine learning-based methods. Journal of Power Sources, 2025. https://doi.org/10.1016/j.jpowsour.2025.236543

[6] Optimizing parameter estimation for electrochemical battery model: A comparative analysis of operating profiles on computational efficiency and accuracy. Applied Energy, 2022, 309, 118484. https://doi.org/10.1016/j.apenergy.2022.118484

[7] Journal of Energy Storage, 2026. https://doi.org/10.1016/j.est.2026.116789

[8] Journal of Energy Storage, 2024, 78, 110012. https://doi.org/10.1016/j.est.2023.110012

[9] Results in Engineering, 2026, 26, 104567. https://doi.org/10.1016/j.rineng.2026.104567

[10] Simultaneously Parameter Identification and Measurement-Noise Covariance estimation of a Proton Exchange Membrane Fuel Cell. Information Sciences, 2022, 613, 112–129. https://doi.org/10.1016/j.ins.2022.09.034

[11] Physics-informed dual-stream networks for battery health estimation. Applied Energy, 2025, 378, 124745. https://doi.org/10.1016/j.apenergy.2024.124745

[12] Domain adaptive state representation alignment for reinforcement learning. Renewable Energy, 2025, 241, 122345. https://doi.org/10.1016/j.renene.2025.122345

[13] IEEJ Transactions on Power and Energy, 2025, 145(4), 327–335. https://doi.org/10.1541/ieejpes.145.327

[14] Imitation reinforcement learning energy management for electric vehicles with hybrid energy storage system. Journal of Energy Storage, 2026, 114, 115678. https://doi.org/10.1016/j.est.2026.115678

[15] Parametric study of adaptive reinforcement learning for battery operations in microgrids. Sustainable Energy Technologies and Assessments, 2025, 73, 104234. https://doi.org/10.1016/j.seta.2025.104234

[16] Reinforcement Learning in Action: Optimal Scheduling for Efficient Battery Management. Journal of Energy Storage, 2024, 100, 113567. https://doi.org/10.1016/j.est.2024.113567

[17] Hierarchical control of hybrid electric vehicle platoon with slope-adaptive variable spacing and soft actor-critic based energy management. Journal of Energy Storage, 2026, 115, 116234. https://doi.org/10.1016/j.est.2026.116234

[18] Optimizing microgrid energy management with hybrid energy storage systems using reinforcement learning methods. Journal of Energy Storage, 2021, 33, 102070. https://doi.org/10.1016/j.est.2020.102070

[19] Energy management strategies based on soft actor critic reinforcement learning with a proper reward function design based on battery state of charge constraints. Renewable and Sustainable Energy Reviews, 2016, 63, 102–117. https://doi.org/10.1016/j.rser.2016.05.031

[20] Physics-enhanced data-driven charging strategies for lithium-ion batteries with safety consciousness. Journal of Energy Storage, 2025, 112, 111234. https://doi.org/10.1016/j.est.2025.111234

[21] Lithium Plating-Aware Fast Charging Control via Control Barrier Functions. Journal of Energy Storage, 2025, 107, 115105. https://doi.org/10.1016/j.est.2024.115105

[22] Multiphysics-Constrained Fast Charging of Lithium-Ion Battery with Active Set Predictive Control. Journal of Energy Storage, 2026, 115, 116567. https://doi.org/10.1016/j.est.2026.116567

[23] Thermal modeling of a high-energy prismatic lithium-ion battery cell and module based on a new thermal characterization methodology. Electrochimica Acta, 2021, 370, 137749. https://doi.org/10.1016/j.electacta.2020.137749

[24] A critical review of thermal management models and solutions of lithium-ion batteries for the development of pure electric vehicles. Engineering Research Express, 2025, 7, 012001. https://doi.org/10.1088/2631-8695/ae16fe

[25] Innovations and prognostics in battery degradation and longevity for energy storage systems. Results in Chemistry, 2025, 14, 102045. https://doi.org/10.1016/j.rechem.2025.102045

[26] Unveiling the secrets behind physics-based modeling of lithium-ion battery degradation and its key applications. Renewable and Sustainable Energy Reviews, 2009, 13(9), 2419–2432. https://doi.org/10.1016/j.rser.2009.03.012

[27] Reinforcement learning for adaptive fault detection in lithium-ion batteries: A next-generation approach for electric vehicles. Applied Soft Computing, 2026, 172, 112345. https://doi.org/10.1016/j.asoc.2026.112345

[28] Digital twins for battery health prognosis: A comprehensive review of recent advances and challenges. e-Prime – Advances in Electrical Engineering, Electronics and Energy, 2025, 12, 100909. https://doi.org/10.1016/j.prime.2025.100909

[29] A review of power electronics interfaces for distributed energy systems towards achieving low-cost modular design. Journal of Cleaner Production, 2025, 490, 144567. https://doi.org/10.1016/j.jclepro.2025.144567

[30] Adaptive teacher modality selection for knowledge distillation: Boosting multimodal model performance. Journal of Energy Storage, 2025, 116, 117234. https://doi.org/10.1016/j.est.2025.117234

[31] Task-Structured Curriculum Learning for Multi-Task Distillation. Energy Strategy Reviews, 2025, 58, 101567. https://doi.org/10.1016/j.esr.2025.101567

[32] Results in Engineering, 2026, 26, 105678. https://doi.org/10.1016/j.rineng.2026.105678

[33] Scalable and generalizable deep learning for battery state of health estimation in on-road electric vehicles. Applied Energy, 2018, 232, 295–308. https://doi.org/10.1016/j.apenergy.2018.09.190

[34] A diffusion causal transformer network with physical information for lithium-ion battery state-of-health estimation. Measurement Science and Technology, 2024, 35, 045109. https://doi.org/10.1088/1361-6501/ad282e

[35] New fast charging method of lithium-ion batteries based on a reduced order electrochemical model considering side reaction. Journal of Power Sources, 2019, 423, 367–379. https://doi.org/10.1016/j.jpowsour.2019.03.007

[36] Electrochemical model-based fast charging of lithium-ion batteries with constraint on anode potential. Journal of Energy Storage, 2024, 84, 110730. https://doi.org/10.1016/j.est.2023.110730

[37] Charging optimization for Li-ion battery in electric vehicles: A review. Applied Energy, 2023, 339, 120904. https://doi.org/10.1016/j.apenergy.2023.120904

[38] Computers, Environment and Urban Systems, 2025, 117, 102234. https://doi.org/10.1016/j.compenvurbsys.2025.102234

[39] Results in Chemistry, 2026, 15, 102567. https://doi.org/10.1016/j.rechem.2026.102567

[40] Journal of Energy Storage, 2026, 118, 118345. https://doi.org/10.1016/j.est.2026.118345

[41] Evaluation and observability analysis of an improved reduced-order electrochemical model for lithium-ion battery. Applied Energy, 2017, 205, 113–128. https://doi.org/10.1016/j.apenergy.2017.07.128

[42] Development and parameterization of a control-oriented electrochemical model of lithium-ion batteries for battery-management-systems applications. Journal of Power Sources, 2024, 590, 233792. https://doi.org/10.1016/j.jpowsour.2023.233792

[43] Development of experimental techniques for parameterization of multi-scale lithium-ion battery models. Journal of the Electrochemical Society, 2020, 167(8), 080534. https://doi.org/10.1149/1945-7111/ac2d17

[44] Temperature and concentration dependence of the ionic transport properties of lithium-ion battery electrolytes. Journal of the Electrochemical Society, 2019, 166(14), A3079–A3097. https://doi.org/10.1149/2.0931813jes