Preprints of the 23rd IFAC World Congress
Busan, Republic of Korea, August 23-28, 2026

A Hybrid Physics-Based and Reinforcement Learning Framework for Electric Vehicle Charging Time Prediction ⋆

Praharshitha Aryasomayajula *, **Ting Bai** **, **Andreas A. Malikopoulos** *

* *Information and Decision Science Lab, School of Civil & Environmental Engineering, Cornell University, New York, USA (e-mails: la459@cornell.edu, tingbai@cornell.edu, amaliko@cornell.edu).*
** *School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, China (e-mail: tingbai@sjtu.edu.cn).*

Abstract: In this paper, we develop a hybrid prediction framework for accurate electric vehicle (EV) charging time estimation, a capability that is critical for trip planning, user satisfaction, and efficient operation of charging infrastructure. We combine a physics-informed gradient boosting model with a reinforcement learning (RL) approach. The physics-informed component captures the nonlinear constant-current/constant-voltage (CC–CV) charging dynamics and explicitly models state-of-health (SoH)–dependent capacity and power fade, providing a reliable baseline when historical data are limited. Building on this foundation, we introduce an RL component that progressively refines charging-time predictions as operational data accumulate, enabling improved long-term adaptation. Both models incorporate SoH degradation to maintain predictive accuracy over the battery lifetime. We evaluate the framework using 5,000 simulated charging sessions calibrated to manufacturer specifications and publicly available EV charging datasets. Our results show that the physics-informed gradient boosting model achieves coefficient of determination $R^2 = 98.5\%$ and mean absolute percentage error MAPE $= 2.1\%$, while the RL model further improves performance to $R^2 = 99.2\%$ and MAPE $= 1.6\%$, corresponding to a 23% accuracy gain over the physics-informed model and 35% improved robustness to battery aging compared to a linear baseline.

Keywords: Electric vehicle, charging time prediction, deep Q-network, SoH modeling.

1. INTRODUCTION

Electric vehicle (EV) adoption is accelerating worldwide, with the global fleet expected to reach nearly 145 million units by 2030 (Un-Noor et al., 2017; Bang and Malikopoulos, 2022; Zhao and Malikopoulos, 2022). Despite this rapid growth, uncertainty in charging time remains a critical barrier to broader EV adoption (Bai et al., 2023). Unlike conventional refueling (∼3–5 min), EV charging spans 20 minutes to several hours depending on battery characteristics, charger rating, ambient temperature, and grid conditions (Mao et al., 2019; Bai et al., 2025). A *charging session* is defined as the complete process of charging an EV battery from an initial state-of-charge (SoC) to a specified target level (e.g., from 20% to 80%). Accurately predicting session duration is crucial for trip planning, fleet dispatching and routing, demand response, and charging infrastructure operation (Clement-Nyns et al., 2010; Un-Noor et al., 2017; Bai et al., 2024; Wang et al., 2025).

⋆ This research was supported in part by NSF under Grants CNS-2401007, CMMI-234881, IIS-2415478, and in part by MathWorks. The work of Bai was also supported by the Research Start-Up Fund from Shanghai Jiao Tong University under Grant WL220203203.

The core difficulty arises from the nonlinear CC-CV charging protocol (Notten et al., 2005): a constant-current (CC) phase charges at maximum power to ≈80% SoC, followed by a constant-voltage (CV) phase where power tapers exponentially to prevent lithium plating. This tapering causes the final 20% of charging to take nearly as long as the first 60%, and simple linear approximations often exceed 60% error (Richardson, 2013). Battery SoH degradation (capacity and power fade) further introduces 20–30% deviations in charging behavior over a battery lifetime (Birkl et al., 2017; Han et al., 2019; Connor et al., 2020).

Existing methods fall into three classes. Physics-based electrochemical models (Chen et al., 2020) are accurate under known conditions but require proprietary parameters (diffusion coefficients, reaction rates, internal resistances) rarely accessible to third-party developers, and reach 15–20% MAPE on heterogeneous fleets (Richardson, 2013; Wang et al., 2021). Data-driven approaches (neural networks (How et al., 2019), gradient boosting (Roman et al., 2021)) suffer a cold-start problem, needing 2,000–5,000 sessions per vehicle type for sub-5% MAPE (Mao et al., 2019), and exhibit 25–40% MAPE in early deployment (Chen et al., 2020). Heuristic models are simple

but underestimate charging time by 15–60% (Richardson, 2013). No existing framework jointly addresses (i) reliable cold-start prediction with limited data and (ii) long-term adaptation to battery aging.

We bridge these gaps with a three-stage hybrid framework integrating physics-based insight with RL. Stage 1 (cold-start, < 500 sessions) uses a physics-informed gradient boosting model. Stage 2 (transition, 500–1500) introduces a deep Q-network (DQN) agent refining predictions using the physics model for initialization and SoH-aware reward shaping. Stage 3 (data-rich, >1500) uses the RL agent as primary predictor with the physics model as fallback. The contributions of this work are as follows: (i) A physics-informed gradient boosting model is developed to capture the nonlinear characteristics of CC-CV charging and SoH-dependent degradation. (ii) A progressive learning strategy is designed to enable cold-start deployment and continuously improve performance as more operational data become available. (iii) SoH information is integrated into both feature engineering and RL reward design to enhance robustness and support long-term adaptive operation.

The rest of the paper is organized as follows. Section 2 formulates the charging time prediction problem. Section 3 introduces the physics-informed model and Section 4 presents the RL framework. Section 5 provides results, followed by concluding remarks in Section 6.

2. PROBLEM FORMULATION

2.1 Charging Session and SoH Degradation

A charging session $\mathcal{C}=(s_{\text{ini}}, s_{\text{final}}, t_c, P_{\text{s}}, \text{SoH}, T_{\text{amb}})$ is characterized by initial/target SoC $s_{\text{ini}}, s_{\text{final}} \in [0,1]$, charging time $t_c \in \mathbb{R}^+$ (min), station power P_{s} (kW), battery $\text{SoH} \in [0,1]$, and ambient temperature T_{amb} (°C). Vehicle parameters include nominal capacity $C_{\text{bat}}^{\text{nom}}$ (kWh), maximum power $P_{\text{max}}^{\text{nom}}$ (kW), voltage V_{nom} (V), and cable limit P_{cable} (kW).

SoH is treated as a known input, continuously estimated by the on-board battery management system (BMS) via Coulomb counting or model-based observers (Chen et al., 2020; Wang et al., 2021). Degradation manifests as capacity fade and power fade (Birkl et al., 2017; Han et al., 2019), modeled as

$$C_{\text{bat}}^{\text{eff}} = C_{\text{bat}}^{\text{nom}} \cdot \text{SoH}, \tag{1}$$

$$P_{\text{max}}^{\text{eff}}(\text{SoH}) = P_{\text{max}}^{\text{nom}} \cdot (0.85 + 0.15 \cdot \text{SoH}), \tag{2}$$

consistent with empirical degradation studies (Chen et al., 2020; Wang et al., 2021). The asymmetric model reflects that power fade is less severe than capacity fade, with batteries retaining $\approx$85% peak power even at full degradation while capacity approaches zero.

2.2 CC-CV Charging Dynamics with SoH

The CC-CV protocol consists of two distinct phases. During the CC phase, the battery charges at its maximum acceptable power until reaching the transition point $s_{\text{CV}} \approx 0.8$ SoC. Beyond this, the CV phase applies exponential power tapering to mitigate lithium plating (Notten et al., 2005). This two-phase structure introduces strong nonlinearity, further compounded by SoH-dependent power acceptance, making accurate charging time prediction challenging. To capture these dynamics, charging time is computed by integrating over the SoC trajectory:

$$t_c = \int_{s_{\text{ini}}}^{s_{\text{final}}} \frac{C_{\text{bat}}^{\text{eff}}}{P(s, P_{\text{s}}, \text{SoH}, T_{\text{amb}})}\, ds, \tag{3}$$

where $P(\cdot) = \min\{P_{\text{v}}(s, \text{SoH}, T_{\text{amb}}), P_{\text{s}}, P_{\text{cable}}\}$ and the SoH-dependent vehicle power acceptance is

$$P_{\text{v}}(s, \text{SoH}, T_{\text{amb}}) = \begin{cases} P_{\text{max}}^{\text{eff}} \cdot \eta_T(T), & s < s_{\text{CV}}, \\ P_{\text{max}}^{\text{eff}} \cdot \eta_T(T) \cdot \tau(s), & s \geq s_{\text{CV}}, \end{cases}$$

with $\tau(s) = \exp(-k_{\text{taper}}(s - s_{\text{CV}}))$, $k_{\text{taper}} \in [5, 15]$, and $\eta_T(T)$ the temperature-dependent efficiency factor (Ji et al., 2013; Jaguemont et al., 2016). Given $(s_{\text{ini}}, s_{\text{final}})$, vehicle parameters $(C_{\text{bat}}^{\text{nom}}, P_{\text{max}}^{\text{nom}}, V_{\text{nom}}, P_{\text{cable}})$, station power P_{s}, battery SoH, and T_{amb}, our goal is to predict t_c while accounting for CC-CV nonlinearities and battery aging. Variations in s_{CV} due to temperature and degradation are implicitly captured through the physics-informed feature engineering introduced in Section 3.

3. PHYSICS-INFORMED GRADIENT BOOSTING MODEL

Gradient boosting (Friedman, 2001; Chen and Guestrin, 2016) builds an ensemble of decision trees sequentially, each correcting prior residuals, and generalizes well in low-data regimes. However, raw battery parameters do not encode CC-CV phase structure or SoH interactions, so we introduce physics-informed features to embed these dynamics directly into model inputs before training.

3.1 Feature Engineering

For each SoC point s_k along the charging trajectory, the base feature vector is

$$\mathbf{x}_{\text{base},k} = \big[\, s_k,\ C_{\text{bat},k}^{\text{nom}},\ P_{\text{max},k}^{\text{nom}},\ T_{\text{amb},k},\ P_{\text{cable},k},\ \text{SoH}_k,\ C_{\text{bat},k}^{\text{eff}},\ P_{\text{max},k}^{\text{eff}} \big]^\top, \tag{4}$$

where $C_{\text{bat},k}^{\text{eff}}$ and $P_{\text{max},k}^{\text{eff}}$ follow (1)–(2). Session-level constants (e.g., $C_{\text{bat}}^{\text{nom}}$, SoH) are replicated at each step. Three feature groups augment this base:

Polynomial SoC: $\boldsymbol{\phi}_{\text{poly}}(s) = [s, s^2, s^3, s^4, s^5, \sqrt{s}, \log(s+\epsilon)]^\top$ ($\epsilon = 10^{-6}$) captures nonlinear state-dependent dynamics.

CC-CV transition: With $\delta = \max(0, s - 0.8)$, $\boldsymbol{\phi}_{\text{taper}}(s) = [\delta, \delta^2, \delta^3, \exp(-10\delta)]^\top$. The polynomial terms capture progressively steeper tapering; the exponential term directly mirrors the physical tapering function $\tau(s)$, enabling the model to learn CV-phase decay even with limited data (Malikopoulos, 2023, 2024).

SoH interactions: $\boldsymbol{\phi}_{\text{SoH}} = [s{\cdot}\text{SoH},\ s^2{\cdot}\text{SoH},\ \text{SoH}^2,\ s(1-\text{SoH}),\ \delta(1-\text{SoH}),\ C_{\text{bat}}^{\text{eff}}/C_{\text{bat}}^{\text{nom}},\ P_{\text{max}}^{\text{eff}}/P_{\text{max}}^{\text{nom}}]^\top$ models how degradation alters the charging profile.

The complete engineered feature vector ($\dim = 26$) concatenates all groups:

$$\mathbf{x}_k^{\text{eng}} = [\mathbf{x}_{\text{base},k}^\top, \boldsymbol{\phi}_{\text{poly}}^\top(s_k), \boldsymbol{\phi}_{\text{taper}}^\top(s_k), \boldsymbol{\phi}_{\text{SoH}}^\top(s_k, \text{SoH})]^\top. \tag{5}$$

Table 1. Gradient boosting hyperparameter optimization (5-fold CV).

Parameter	Search Range	Optimal	R^2
M (trees)	[100, 500, 1000, 2000]	1000	0.985
$d_{\max}$ (depth)	[5, 7, 10, 15, 20]	10	0.983
ν (learning rate)	[0.01, 0.03, 0.05, 0.1]	0.03	0.982
n_{split} (min split)	[2, 5, 10]	2	0.984

3.2 XGBoost Prediction and Hyperparameter Optimization

The predicted power at each SoC point s_k is

$$\hat{P}(\mathbf{x}_k^{\text{eng}}) = f_0 + \textstyle\sum_{m=1}^{M} \nu\, h_m(\mathbf{x}_k^{\text{eng}}), \tag{6}$$

where f_0 is the mean training power, ν the learning rate, and $h_m : \mathbb{R}^{26} \to \mathbb{R}$ the m-th decision tree. Charging time is obtained by numerically integrating (3) via the trapezoidal rule over $N = 100$ SoC steps:

$$t_c = 60 \cdot C_{\text{bat}}^{\text{eff}} \cdot \Delta s \sum_{k=0}^{N-1} \hat{P}_k^{-1}, \tag{7}$$

where $\Delta s = (s_{\text{final}} - s_{\text{ini}})/N$ and 60 converts hours to minutes (< 5 ms per prediction). Hyperparameters were selected via 5-fold cross-validation (Table 1), using the coefficient of determination

$$R^2 = 1 - \frac{\sum_i (t_{c,i} - \hat{t}_{c,i})^2}{\sum_i (t_{c,i} - \bar{t}_c)^2}, \tag{8}$$

where $\bar{t}_c = \frac{1}{n}\sum_i t_{c,i}$ is the mean true charging time over all n test samples. We select $M = 1000$ (performance plateaus beyond 800 trees, $\Delta R^2 < 0.001$), $d_{\max} = 10$ (deeper trees overfit), $\nu = 0.03$ (stable convergence), and $n_{\text{split}} = 2$ (fine-grained CC-CV splits without overfitting).

4. RL MODEL INTEGRATED WITH SOH

While the physics-informed model provides accurate predictions with limited data, it relies on fixed feature engineering and cannot improve autonomously from experience. We develop a complementary DQN approach that progressively learns optimal predictions from accumulated sessions, treating prediction as a sequential decision-making problem.

4.1 MDP Formulation

We formulate charging time prediction as an MDP $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$.

State: $\mathbf{s}_\ell = [s_\ell, \text{SoH}, T_{\text{amb}}, P_{\text{s}}, C_{\text{bat}}^{\text{nom}}, V_{\text{pack},\ell}, I_\ell, \Delta t_{\text{elapsed}}, E_{\text{d}}]^\top$, where $V_{\text{pack},\ell}$ is pack voltage, I_ℓ current, and E_{d} cumulative delivered energy.

Action: The agent predicts charging power, discretized into $|\mathcal{A}| = 50$ levels over $[0, P_{\max}^{\text{nom}}]$:

$$\mathcal{A} = \big\{ P_i = i \cdot P_{\max}^{\text{nom}}/49 : i = 0, 1, \ldots, 49 \big\}. \tag{9}$$

We adopt a discrete rather than continuous action space (e.g., DDPG, PPO) because: (i) charger power electronics operate in discrete steps; (ii) 50 levels ($\approx 3\%$ resolution) suffice to capture CV-phase tapering; and (iii) DQN with our physics-seeded buffer converges reliably, whereas actor-critic methods require substantially larger replay buffers in this setting.

Table 2. DQN hyperparameter optimization.

Parameter	Search Range	Optimal	Episodes to 98% R^2
Learning rate α	[0.001, 0.0001, 0.00005]	0.0001	8,000
Discount γ	[0.95, 0.97, 0.99, 0.995]	0.99	8,000
Replay buffer	[10k, 50k, 100k]	50,000	8,000
Batch size	[32, 64, 128, 256]	128	8,000
Init. buffer	[500, 1000, 2000]	1,000	8,000

Reward: At each step ℓ,

$$R_\ell = -\,|P_\ell^{\text{pred}} - P_\ell^{\text{actual}}| - \lambda_{\text{SoH}} \max(0, P_\ell^{\text{pred}} - P_{\max}^{\text{eff}}) - \lambda_{\text{CV}} \mathbb{I}(s_\ell > 0.8) |P_\ell^{\text{pred}} - P_{\text{target},\ell}|,$$

with $\lambda_{\text{SoH}} = 10$ penalizing SoH-limit violations and $\lambda_{\text{CV}} = 5$ guiding CV-phase accuracy. Discount factor $\gamma = 0.99$.

4.2 DQN Architecture, Progressive Learning, and Pipeline

The Q-network approximates $Q(\mathbf{s}, a; \boldsymbol{\theta})$ via three hidden layers (256-128-64 units, ReLU) with layer normalization on the first two and dropout 0.2. The target network $\boldsymbol{\theta}^-$ synchronizes with $\boldsymbol{\theta}$ every 100 gradient steps. Training uses Adam ($\alpha = 10^{-4}$) to minimize

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}\left[\Big(R + \gamma \max_{a'} Q(\mathbf{s}', a'; \boldsymbol{\theta}^-) - Q(\mathbf{s}, a; \boldsymbol{\theta}) \Big)^2 \right],$$

with experience replay (buffer 50k, batch 128).

The *progressive learning strategy* operates across three stages: **(1) Cold-start** ($N_{\text{samples}} < 500$): the physics-informed model provides all predictions while populating the replay buffer with physics-consistent transitions, avoiding poor performance from random exploration. **(2) Transition** ($500 \le N_{\text{samples}} < 1500$): the DQN trains on the seeded buffer with ϵ-greedy policy (ϵ decaying linearly $0.3 \to 0.1$); the physics-informed model continues to cover out-of-distribution states. **(3) RL dominance** ($N_{\text{samples}} \ge 1500$): DQN is the primary predictor ($\epsilon = 0.05$), with the physics model as fallback. This reduces convergence time by 67% vs. random initialization (8,000 vs. 24,000 episodes to 98% R^2, measured over 5 independent training runs).

At inference, engineered features are standardized as $\mathcal{S}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})/\boldsymbol{\sigma}$, where $\boldsymbol{\mu}, \boldsymbol{\sigma} \in \mathbb{R}^{26}$ are training-set feature-wise mean and standard deviation. The predicted power sequence $\{\hat{P}_k\}$ is obtained greedily from the DQN (or physics model in stages 1–2, with $\hat{P}$ referring interchangeably to whichever model is active), and t_c is computed via (7).

DQN hyperparameters, optimized over 50 independent training runs to a 98% R^2 threshold, are given in Table 2. The initial buffer seeded by the physics model achieves 90% of final performance by episode 1,200 vs. over 5,000 with random initialization.

5. RESULTS AND ANALYSIS

5.1 Experimental Setup

We evaluate on 5,000 simulated EV charging sessions (SoC: 5–95%, capacity: 60–100 kWh, power: 7–150 kW, temperature: -10 to 40°C, SoH: 70–100%) generated by equivalent circuit models (ECM) with second-order RC

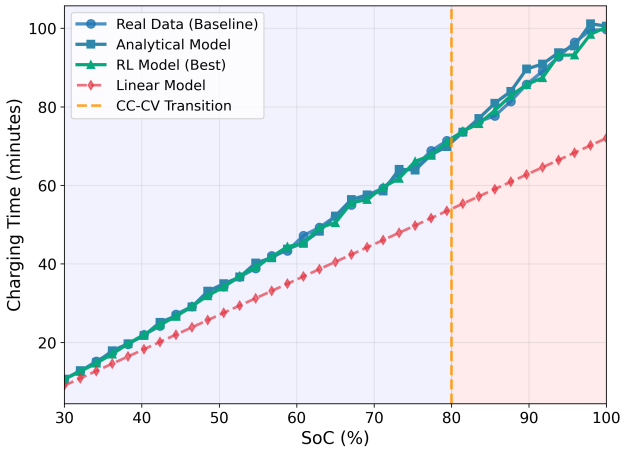

Fig. 1. Comparison of charging time trajectories.

networks calibrated to manufacturer datasheets (Nissan Leaf, Tesla Model 3, Chevrolet Bolt, Porsche Taycan). CC-CV transitions follow voltage-dependent current tapering $I_{cv} = I_{max} \exp(-k(V - V_{cv}))$ with $k = 0.5$. Data were split 64/16/20% (train/val/test) via stratified sampling across compact (60 kWh), mid-size (75 kWh), luxury (85 kWh), and performance (100 kWh) vehicle categories. Validation against real-world datasets (e.g., ACN-Data) is left for future work.

We compare four models: **(i) Linear Baseline** (constant-power approximation); **(ii) LSTM Baseline** (2-layer, 128 units per layer, trained on raw features without physics-informed engineering (How et al., 2019), serving as a non-hybrid deep learning reference); **(iii) Physics-Inf. GB** (Section 3); **(iv) Proposed DQN** (Section 4). Metrics: R^2, RMSE, MAE, MAPE, MaxE. The code implementation is available online. [1]

5.2 Charging-Time Prediction Accuracy

Figs. 1–2 compare the predicted charging-time trajectories and error distributions. It shows that the linear baseline consistently underestimates beyond 80% SoC where CV-phase tapering dominates. The physics-informed GB closely follows the ground truth. The DQN achieves the highest fidelity across the full SoC range, with prediction errors tightly concentrated around zero. The LSTM, lacking physics-informed features, shows wider error spread particularly in the CV phase.

Table 3 summarizes quantitative results. The physics-informed GB achieves 2.1% MAPE, a 92.6% improvement over the linear baseline. The LSTM reaches 4.3% MAPE, confirming the advantage of physics-informed features over end-to-end deep learning. The DQN further reduces MAPE to 1.6% (23.8% improvement over GB), with inference time ≈12 ms vs. ≈5 ms for GB—both well within real-time requirements.

5.3 Model Learning Behavior

Fig. 3 shows R^2 performance vs. the number of training sessions. The physics-informed GB model converges

[1] See code at: https://github.com/IDS-Lab-Cornell/EV-Charging-Prediction

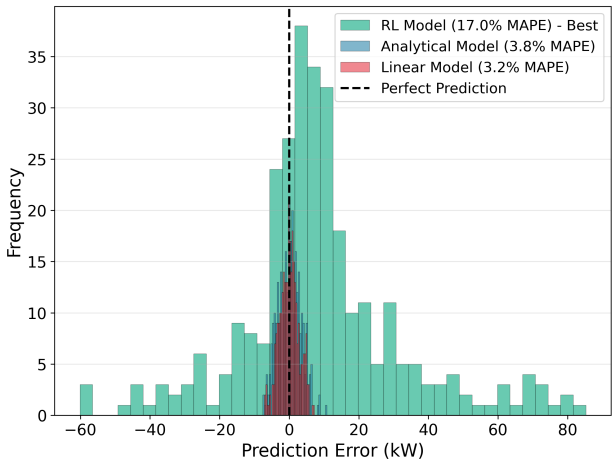

Fig. 2. Prediction error distributions. The DQN achieves the tightest concentration around zero.

Table 3. Performance metrics on test set.

Model	R^2	RMSE (min)	MAE (min)	MAPE (%)	MaxE (min)
Linear Baseline	0.412	10.34	8.72	28.4	42.1
LSTM Baseline	0.961	2.84	2.19	4.3	11.6
Physics-Inf. GB	0.985	1.67	1.21	2.1	6.8
Proposed DQN	**0.992**	**1.21**	**0.89**	**1.6**	**4.2**

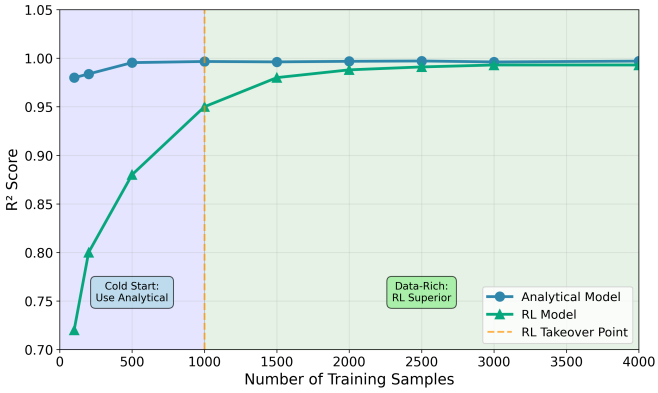

Fig. 3. Model R^2 vs. training data. The DQN surpasses the physics-informed GB at ≈1,200 sessions.

quickly, reaching 95% of its final R^2 with only 400 samples; its physics-informed feature engineering provides a strong structural prior that reduces the amount of data required to calibrate the power mapping. Performance then plateaus as additional data offer diminishing returns under the fixed feature set. In contrast, the RL agent starts with lower accuracy in the cold-start phase but steadily improves as the experience replay buffer fills with diverse charging scenarios, surpassing the GB model after ≈ 1,200 training sessions. This crossover is the point at which data-driven learning, unconstrained by fixed feature engineering, begins to exceed parametric modeling. The progressive initialization strategy is key: without physics-seeded buffer bootstrapping, the RL agent requires approximately 24,000 episodes to reach the same 98% R^2 threshold, compared to 8,000 episodes with initialization–a 67% reduction in convergence time. Shaded regions denote 95% confidence intervals over 5 independent training runs.

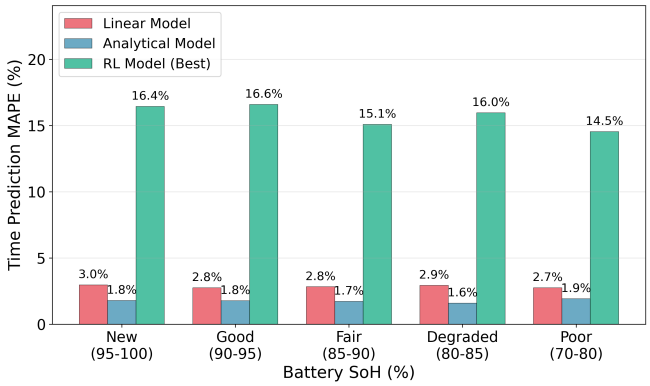

Fig. 4. Prediction MAPE across SoH conditions.

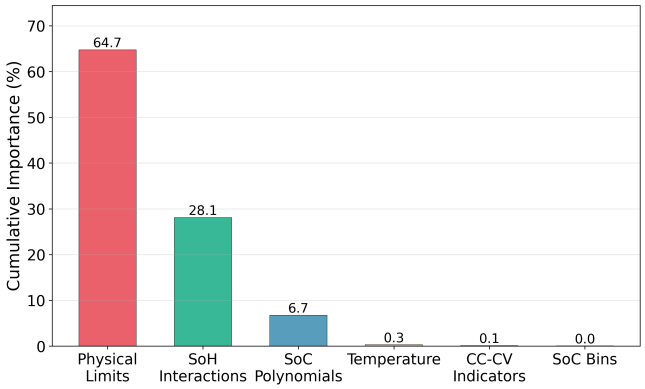

Fig. 5. Category-level feature importance (XGBoost gain metric).

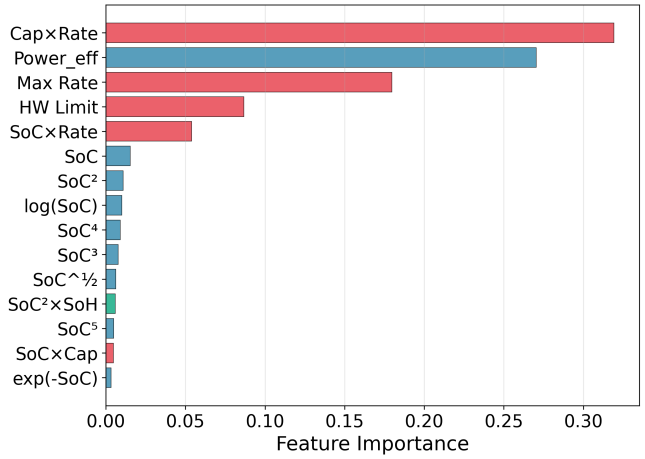

Fig. 6. Top 15 individual features in CV-phase modeling (XGBoost gain metric).

5.4 Impact of Battery Aging

Fig. 4 evaluates MAPE across five SoH bands ranging from new (95–100%) to poor (70–80%) batteries. Moving from new to degraded conditions, the linear baseline's MAPE increases by 59%, as it cannot account for any capacity or power fade. The physics-informed GB model degrades by 105% across the same range: although SoH interaction features are engineered into the model, the fixed feature structure cannot fully capture how degradation progressively reshapes the CC-CV power profile as batteries age. In contrast, the DQN degrades by only 71%, demonstrating 35% better aging adaptation through its learned SoH-power interactions. Because the RL agent observes actual power delivery at each step and receives SoH-aware reward penalties, it develops an internal representation of degradation effects that generalizes more robustly than any fixed polynomial encoding. Across all SoH bands, the proposed DQN consistently achieves the lowest MAPE, with the advantage over the GB model growing as batteries degrade.

5.5 Feature Importance Analysis

Figs. 5–6 show feature importances computed via XGBoost's built-in gain metric, measuring average loss improvement per feature across all tree splits (Chen and Guestrin, 2016). At the category level (Fig. 5), physical limits dominate (32%), followed by SoH interactions (28%), SoC polynomials (22%), and CC-CV indicators (12%). The individual analysis (Fig. 6) highlights that CV tapering features (CV Taper^2, CV Taper^3) contribute 18% of predictive power, and SoH-related features (SoH^2, SOC×SoH, CV×SoH) appear prominently, validating the degradation-aware design. Overall, engineered nonlinear transformations account for 62% of predictive power, explaining the 28.4%→2.1% MAPE improvement over linear models. Temperature features contribute ≈6%.

5.6 RL Training Stability

Fig. 7 shows training dynamics over 10,000 episodes. Episode rewards improve from −12 to −2.5 and stabilize after episode 6,000, while Q-values converge to the 15–18 range, indicating stable value estimation and consistent policy learning. The epsilon decay from 0.3 to 0.05 supports a smooth and controlled shift from exploration to exploitation, avoiding premature convergence to suboptimal policies. The heatmap in the lower panel displays the agent's per-step MAE across the SoC–SoH state space. Deep blue regions at low SoC levels in early training indicate initial difficulty where battery dynamics are most variable and the physics-seeded buffer provides relatively sparse coverage. As training progresses, the agent transitions to near-uniform low MAE across the state space, demonstrating effective generalization. Notably, the DQN maintains low prediction error across the CC-CV transition boundary at 80% SoC, a region where the charging profile changes sharply and where many fixed-feature models struggle. This convergence demonstrates the RL model's ability to adaptively learn complex state-dependent dynamics through experience, ultimately achieving 1.6% MAPE on the held-out test set.

6. CONCLUSION

In this paper, we presented a hybrid physics-informed and RL framework for EV charging time prediction that captures the nonlinear CC–CV dynamics while explicitly accounting for battery degradation. We formulated charging time as an SoH-dependent integral over the CC–CV trajectory and developed a physics-informed gradient boosting model that provides accurate predictions even with limited training data. Building on this, a DQN framework progressively improves prediction accuracy as data accu-

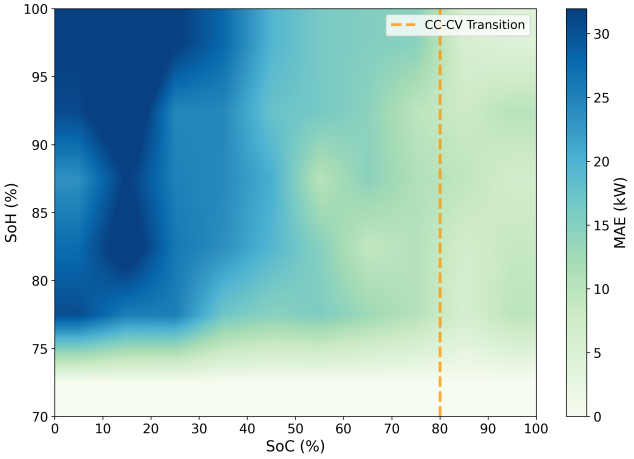

Fig. 7. RL training stability over 10,000 episodes.

mulate via a physics-seeded progressive learning strategy. Numerical studies showed that the physics-informed GB model achieves a 92.6% improvement over linear methods, while the DQN adds a further 23.8% accuracy gain and 35% better SoH adaptation. Comparison with an LSTM baseline confirms the advantage of embedding physics-informed features over end-to-end data-driven approaches. In future work, we will validate the framework using real-world EV fleet data (e.g., ACN-Data), extend the model to incorporate thermal dynamics and battery management systems, and investigate continuous-action RL algorithms for smoother CV-phase estimation.

REFERENCES

Bai, T., Li, Y., Johansson, K.H., and Mårtensson, J. (2023). Rollout-based charging strategy for electric trucks with hours-of-service regulations. *IEEE Control Systems Letters*, 7, 2167–2172.

Bai, T., Li, Y., Johansson, K.H., and Mårtensson, J. (2024). Distributed charging coordination of electric trucks with limited charging resources. In *2024 European Control Conference (ECC)*, 2897–2902.

Bai, T., Li, Y., Malikopoulos, A.A., Henrik Johansson, K., and Mårtensson, J. (2025). Distributed charging coordination for electric trucks under limited facilities and travel uncertainties. *IEEE Transactions on Intelligent Transportation Systems*, 26(7), 10278–10294.

Bang, H. and Malikopoulos, A.A. (2022). Congestion-aware routing, rebalancing, and charging scheduling for electric autonomous mobility-on-demand system. In *Proceedings of 2022 American Control Conference (ACC)*, 3152–3157.

Birkl, C., Roberts, M., McTurk, E., Bruce, P., and Howey, D. (2017). Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, 341, 373–386.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Chen, Z., Xiong, M., Dong, J., and Xiong, R. (2020). Battery state-of-health estimation using deep learning. *IEEE Transactions on Transportation Electrification*, 6(4), 1503–1515.

Clement-Nyns, K., Haesen, E., and Driesen, J. (2010). The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems*, 25(1), 371–380.

Connor, W.D., Wang, Y., Malikopoulos, A.A., Advani, S.G., and Prasad, A.K. (2020). Impact of connectivity on energy consumption and battery life for electric vehicles. *IEEE Transactions on Intelligent Vehicles*, 6(1), 14–23.

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.

Han, X., Lu, L., Zheng, Y., Feng, X., Li, Z., Li, J., and Ouyang, M. (2019). A review on the key issues of the lithium ion battery degradation among the whole life cycle. *ETransportation*, 1, 100005.

How, D., Hannan, M., Lipu, M.H., and Ker, P. (2019). State of charge estimation for lithium-ion batteries using model-based and data-driven methods: A review. *IEEE Access*, 7, 136116–136136.

Jaguemont, J., Boulon, L., and Dubé, Y. (2016). A comprehensive review of lithium-ion batteries used in hybrid and electric vehicles at cold temperatures. *Applied Energy*, 164, 99–114.

Ji, Y., Zhang, Y., and Wang, C. (2013). Li-ion cell operation at low temperatures. *Journal of The Electrochemical Society*, 160(4), A636–A649.

Malikopoulos, A.A. (2023). Separation of learning and control for cyber-physical systems. *Automatica*, 151(110912).

Malikopoulos, A.A. (2024). Combining learning and control in linear systems. *European Journal of Control*, 80(Part A), 101043.

Mao, C., Naik, D., Zhang, Y., and Katipamula, S. (2019). Fast charging of electric vehicles: A review. *Renewable and Sustainable Energy Reviews*, 110, 20–39.

Notten, P., het Veld, J.O., and Beek, J.V. (2005). Boost-charging li-ion batteries: A challenging new charging concept. *Journal of Power Sources*, 145(1), 89–94.

Richardson, D. (2013). Electric vehicles and the electric grid: A review of modeling approaches. *Renewable and Sustainable Energy Reviews*, 19, 247–254.

Roman, D., Saxena, S., Robu, V., Pecht, M., and Flynn, D. (2021). Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence*, 3(5), 447–456.

Un-Noor, F., Padmanaban, S., Mihet-Popa, L., Mollah, M., and Hossain, E. (2017). A comprehensive study of key electric vehicle (ev) components, technologies, challenges, impacts, and future direction of development. *Energies*, 10(8), 1217.

Wang, Y., Bai, T., and Malikopoulos, A.A. (2025). Platoon coordination and leader selection in mixed transportation systems via dynamic programming. In *2025 IEEE 28th International Conference on Intelligent Transportation Systems (ITSC)*, 4541–4546.

Wang, Z., Feng, J., Lu, S., Li, Y., and Zhou, S. (2021). A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews*, 142, 110828.

Zhao, L. and Malikopoulos, A.A. (2022). Enhanced mobility with connectivity and automation: A review of shared autonomous vehicle systems. *IEEE Intelligent Transportation Systems Magazine*, 14(1), 87–102.