Simulation-Based Deep Reinforcement Learning Framework for EV Charging Protocols Balancing Battery Degradation and Charging Performance

Aarush Nair[1], **Savar Shandilya**[1]

Received May 1, 2026
Accepted September 5, 2026
Electronic access October 15, 2026

Electric vehicle adoption continues to persist, but battery degradation poses a major issue under elevated charging rates, high temperatures, and operation outside optimal State of Charge (SOC). Constant Current-Constant Voltage (CCCV) charging is standard in electric vehicles (EVs), but its rule-based structure does not adapt to changing conditions, potentially accelerating degradation. This study aims to improve upon CCCV by deploying a reinforcement learning (RL) charging strategy that reduces degradation and charging duration. Data was generated using the Single Particle Model (SPM) in PyBaMM through 200-cycle simulations. A hybrid Deep Q-Network (DQN) controller was trained offline and evaluated online in PyBaMM against CCCV; fixed 0.72C, 1.0C, and 1.5C policies; a rule-based heuristic; and greedy controller. All policies used the 25-85% SOC protocol, transferred 3.0 Ah per charge and discharge phase, and ended at 120 EFC. Policy outputs were evaluated using the Composite Degradation Index (CDI), a constructed stress index, alongside discharge capacity, capacity loss to the Solid Electrolyte Interphase (SEI), Loss of Lithium Inventory (LLI), Loss of Active Material (LAM), resistance growth, capacity fade, charging-cycle duration, and battery temperature. Across five seeds, DQN generated the lowest CDI of 0.03238 ± 0.00020, with a mean charging duration of 42.0 ± 0.25 minutes. This corresponded to a mean CDI reduction of 0.00134 and charging-time reduction of 30.0 minutes relative to CCCV, while avoiding the higher degradation of faster policies.

Keywords: Li-ion battery degradation, PyBaMM, deep reinforcement learning, battery charging optimization, Deep Q-Network, CCCV

Introduction

Electric vehicles are rapidly gaining widespread adoption, but this growth brings attention to the issue of long-term degradation of the lithium-ion batteries utilized to power them. Repeated charging cycles, frequent high-current charges, exposure to elevated temperatures, and operation outside moderate bounded SOC ranges have been shown to accelerate battery degradation[1,2]. These behaviors make degradation-aware charging a necessity for EVs long term. Interacting mechanisms, including Solid Electrolyte Interphase (SEI) growth, Loss of Lithium Inventory (LLI), lithium plating, particle cracking, and stress-driven Loss of Active Material (LAM) can lead to an accumulation of battery-aging and overall degradation[3]. In severe cases, frequent operation under aggressive conditions can trigger thermal runaway, where exothermic chain reactions cause rapid heat generation, cell destabilization, and potential fires[4]. The traditional Constant Current-Constant Voltage (CCCV) method is commonly used for EV charging to mitigate these concerns. In this protocol, the battery is charged at a steady current until a certain voltage threshold is met, and then the voltage is held steady while the current tapers[5]. Although the method is reliable and simple, its rule-based system does not adjust to changing battery conditions. As a result, charging-control studies have designed protocols to optimize charging performance, balancing thermal, degradation, and electrochemical limits.

Prior studies using model-based control and reinforcement learning (RL) have been applied to improve fast-charging behavior. Wei et al. combined a model-based state observer with a DDPG optimizer to reduce charging time while incorporating voltage, temperature, and degradation-related constraints[6]. Wassiliadis et al. used a reduced-order electrochemical model to mitigate lithium-plating risk and evaluate cycle-life effects[7]. Lu et al. designed a protocol that incorporated lithium-plating detection and updated model parameters as the battery aged[8].

Other studies have used physics-based simulation environments or safer charging frameworks to support adaptive charging. El Ouazzani et al. trained a PPO-based multistage charging controller in a PyBaMM Single Particle Model with electrolyte dynamics (SPMe) environment with voltage and tem-

[1] *Washington High School, Fremont, California, USA.*

© The National High School Journal of Science 2026

perature limits[9]. Yuan and Zou developed a TD3 controller that accounted for battery aging instead of resetting the battery to a fresh state after each episode[10]. He et al. trained an RL charging controller using real-world experimental battery state data, addressing the limitations of controllers that utilized idealized simulation data[11]. Zhang et al. used three-electrode measurements and ML to estimate anode plating potential during battery aging[12]. Chowdhury et al. used Gaussian-process models to estimate condition limits and project unsafe actions into feasible operating regions[13]. Another study applied DDPG to pack-level charging under voltage, temperature, SOC, and State of Health (SOH) constraints, reducing charging duration relative to CCCV[14]. Park et al. implemented a deep RL controller using battery states derived from a Doyle-Fuller-Newman (DFN) electrochemical model, testing its sensitivity to model-parameter changes[15].

This study builds upon prior EV battery research by developing an adaptive RL model in PyBaMM using the Single Particle Model (SPM). The model responds to degradation-related outputs including LLI, SEI growth, discharge capacity, loss of capacity to SEI, resistance growth, negative and positive LAM, and capacity fade. Author-constructed metrics were included, such as the Composite Degradation Index (CDI) and CDI rate. PyBaMM simulations across different charging rates and ambient temperatures simulations trained the model to make informed charging decisions. Ridge regression was used as an exploratory tool to identify the association of specific variables with CDI and additional degradation metrics. The proposed RL model is a hybrid Deep Q-Network (DQN) controller designed to reduce total battery degradation through dynamic control of charging rate rather than fixed-policy charging. The controller prioritizes preventing sustained operation at high charging rates and limiting substantial thermal rise, as these affect battery health significantly. The DQN used a discrete charging-action space to select from predefined charging rates. Ambient temperature was explicitly incorporated as an input to avoid operation outside safe and optimal external conditions.

This approach was evaluated in comparison to CCCV, fixed charging-rate policies, and dynamic policies over 200 cycles in PyBaMM. All policies followed the same 25%–85% SOC window, 3.0 Ah charge/discharge throughput, and 120 EFC endpoint to ensure a fair comparison and isolate the DQN's adaptiveness. The primary focus of this study was to determine whether RL-guided charging could reduce long-term degradation while maintaining practical charging and safe temperatures.

Methods

The data for this experiment was generated using PyBaMM, a physics-based lithium-ion battery simulation framework for developing and comparing battery models and numerical methods[16]. It resolves internal metrics, enabling the model to evaluate variables that may not be otherwise available without physical cell testing. The primary model used was the Single Particle Model (SPM), a reduced-order electrochemical model that represents each electrode using a representative active-material particle.

SPM was preferred over higher-fidelity models such as SPMe and DFN because it is less computationally expensive. This makes it more practical for repeated 200-cycle simulations, dataset generation, and multi-policy evaluation. SPMe adds leading-order electrolyte behavior, and DFN resolves porous-electrode and electrolyte transport in greater detail[17].

However, this higher electrochemical detail contributes to substantially higher complexity and runtime[17]. Because SPM does not fully resolve electrolyte concentration gradients or porous-electrode transport, it is unable to track issues that occur during fast-charging such as local anode potential or lithium-plating risk. Due to its simplicity, SPMe and DFN were used to validate whether primary charging trends from SPM persisted in more detailed simulations.

Seven policies were evaluated for 200 cycles in SPM and validated in SPMe and DFN: the Deep Q-Network (DQN), bounded 0.5C Constant Current-Constant Voltage (CCCV), fixed 0.72C with the same tapering rule, fixed 1.0C, fixed 1.5C, the SOC and temperature-based heuristic, and greedy controller. The fixed 0.72C policy used a fixed charging rate of 0.72C while tapering the current near the upper SOC boundary. The fixed 1.0C and 1.5C policies represented faster constant-rate charging comparisons. The SOC and temperature-based heuristic reduced the current when the SOC exceeded 75% or temperature eclipsed 30°C. The greedy controller selected the lowest-immediate-penalty action from the same action space as the DQN.

Each policy followed the same 25–85–25% State of Charge (SOC) protocol[1], transferred 3.0 Ah during each charge and discharge phase, and ended at 120 EFC. This intermediate SOC charging window was selected as prior studies have noted that it can prevent cells from aging rapidly, which often occurs during operation in particularly low and high windows[1]. An EFC endpoint was chosen to ensure cumulative cycling exposure didn't affect the end comparison. All simulations are based on the OKane2022 parameter library and a configuration including ec-reaction-limited SEI growth, SEI porosity change, stress-driven Loss of Active Material, and a lumped thermal model[3]. Minimum local anode potential was recorded within both the SPMe and DFN simulations because values below 0 V vs. Li/Li^+ indicate increased lithium-plating risk, which is a degradation issue linked to fast-charging that SPM does not capture. To capture the percentage of time in which this occurred, the anode-below-0-V fraction was recorded.

© The National High School Journal of Science 2026

Simulation output was collected as 30-second time-series data and cycle-level summaries. Time-series data consisted of charging rate, battery temperature, SOC, Loss of Lithium Inventory (LLI), Solid Electrolyte Interphase (SEI) thickness, the Composite Degradation Index (CDI), and CDI rate. Cycle level data included charging duration, Constant Voltage (CV) phase duration, SOC, Effective Full Cycles (EFC), discharge capacity, loss of capacity to SEI, LLI, negative-electrode and positive-electrode Loss of Active Material (LAM), resistance growth, capacity fade (% of nominal capacity), CDI, and end CDI rate. The Composite Degradation Index (CDI) was used as an author-constructed degradation-stress metric. CDI combined normalized LLI with stress multipliers for charging current, high-voltage exposure, and temperature. Current stress represented charge current relative to nominal capacity to allow for higher currents to bring higher stress. Voltage stress represented the extent to which voltage exceeded 3.7 V and was divided by 0.5 to normalize. Thermal stress was measured as the temperature deviation from 25°C and was scaled by 20 to normalize moderate temperature shifts. This enables stress to reflect Arrhenius-type behavior that occurs within batteries. CDI rate was resolved as the time derivative of CDI.

The implementation used weights of 0.40 for current stress, 0.15 for voltage stress, and 0.50 for thermal stress. Thermal stress received the largest weight because elevated temperature accelerates SEI growth, electrolyte decomposition, and side reactions. Current stress was weighted slightly lower because it increases degradation directly and indirectly through heat generation, meaning some of its effect is already reflected through thermal stress. Voltage stress received the lowest weight to avoid overemphasizing high-SOC exposure and the CV phase, as the upper SOC boundary is only 85%. The CDI equation, normalization constants, exponential temperature term, and stress weights are still author-defined components of a constructed index. They were not calibrated as a physical aging law and should not be seen as a direct capacity fade.

Since CDI utilizes manually selected weights for stress terms, a sensitivity analysis was conducted to identify if the chosen weights were significantly altering the policy ranking. Six weighting methods were used, with the weights listed in the order: current stress, voltage stress, and thermal stress. These methods were the default weights listed above, equal weights (0.3, 0.3, 0.3), high current (0.7, 0.3, 0.3), high thermal (0.3, 0.3, and 0.7), high voltage (0.3, 0.7, 0.3), and low (0.15, 0.10, and 0.20).

$$\sigma_I(t) = \frac{I_{ch}(t)}{Q_{nom}}, \qquad \sigma_V(t) = \frac{\max(0, V(t) - 3.7)}{0.5}, \qquad \sigma_T(t) = \exp\left(\frac{T(t) - 25}{20}\right) \tag{1}$$

Equation 1. $I_{ch}(t)$ is charging current, Q_{nom} is nominal cell capacity, $V(t)$ is terminal voltage, and $T(t)$ is battery temperature in degrees Celsius. $\sigma_I(t)$, $\sigma_V(t)$, and $\sigma_T(t)$ represent current, voltage, and thermal stress terms, respectively.

$$\mathrm{CDI}(t) = \mathrm{LLI}_{\mathrm{norm}}(t)\left(1 + 0.40\,\sigma_I(t)\right)\left(1 + 0.15\,\sigma_V(t)\right)\left(1 + 0.50\,\sigma_T(t)\right) \tag{2}$$

Equation 2. $\mathrm{CDI}(t)$ denotes the Composite Degradation Index, a unitless degradation-stress metric. $\mathrm{LLI}_{\mathrm{norm}}(t)$ is normalized loss of lithium inventory, while $\sigma_I(t)$, $\sigma_V(t)$, and $\sigma_T(t)$ represent the current, voltage, and thermal stress terms defined in Equation 1. The normalized LLI term is multiplied by these stress multipliers.

Constant Current-Constant Voltage Protocol

A structured control design with moderate conditions was created in PyBaMM to evaluate the impact of the Constant Current-Constant Voltage (CCCV) protocol on long-term battery degradation. The baseline charging routine follows the conventional CCCV structure, in which the battery is charged at a constant current until an upper-voltage threshold is reached and then held at constant voltage while the current tapers[5].

These protocol values below were author-selected. These included the 4.2 V upper-voltage limit, 0.05 A current-termination value, and 3.0 V discharge cutoff. The ambient temperature was set at 25°C.

In this implementation, the battery was charged at a constant rate of 0.5C until it approached the 85% SOC upper boundary, at which the current was tapered to prevent the protocol from exceeding the SOC limit. The discharge process then proceeded under the same cycling structure until the cell returned to 25% SOC. The SOC window was kept consistent rather than shifted to represent a universal charging standard.

Additional simulations were conducted with both minor and extreme adjustments to isolate primary battery degradation drivers, particularly charging rate and operating temperature during fast charging[2].

Controlled variations were made to the charging rate, including simulating using C/2, 1.5C, and 3.0C to represent moderate and extreme currents, with an initial temperature of 25°C. Thermal variations were also conducted, including simulations under 25°C, 35°C, and 45°C, while maintaining the charging rate at 0.5C. The selected ranges were author-chosen, but are consistent with lithium-ion battery performance modeling studies[18]. These condition-focused simulations were run to train the DQN and regression model.

Ridge Regression

Ridge regression was used to identify exploratory feature associations. It was implemented because it models linear predictor response relationships while applying L2 regularization on coefficient magnitudes to minimize noise sensitivity [19]. Coefficients were attached to each metric, and StandardScaler was used to standardize them and scale them to unit variance prior to fitting so that differences in numerical scales would not factor into the calculation [20]. The resulting coefficient set was interpreted as exploratory weights that existed solely within the tested simulations.

The regression dataset was synthesized from the CCCV baseline, the 35°C and 45°C simulation, and 1.5C and 3C simulation. Cycle-level outputs were used for training Ridge regression instead of adjacent time-series rows to minimize overfitting from similar neighboring data. Since only a small number of thermal and charging-rate conditions were included, the dataset was generated from a limited simulation space.

The predictor set used included mean charging rate, mean cycle battery temperature, charge-end SOC, maximum voltage, and EFC. Charging rate was selected as it reflects current-induced stress within the cell. Battery temperature reflects the thermal acceleration of degradation through side reactions. SOC and voltage encapsulate the electrochemical operating state of the battery, with degradation being enhanced when deeper cycling occurs. EFC was included as it represents cumulative cycle-induced aging. With these features, the model observes both instantaneous and long-term impacts to battery degradation. CDI, LLI, capacity loss to SEI, negative and positive LAM, and resistance growth, were utilized as targets.

Performance was evaluated using condition-held-out validation, in which four of the five simulated conditions were used for training and the fifth condition was used for testing. The coefficient of determination (R^2), normalized root mean squared error (NRMSE), 95% confidence intervals, mean residual, and residual standard deviation were used as metrics for assessment as well. R^2 quantified the proportion of variance for the targets by the model, while NRMSE measured the prediction error for resolved target values in relation to the actual value. Mean residual represented the average prediction error and residual standard deviation measured how scattered the errors were, with both being used to highlight model bias. The trained regression model was applied to policies excluded from training to assess their predictiveness for degradation outputs under unseen charging strategies.

Reinforcement Learning Controller

The adaptive charging RL controller was designed to change the charging behavior of the battery in response to changes in battery condition. The controller was implemented as a hybrid Deep Q-Network (DQN), in which the neural network generated action-value estimates for each charging action and identified the best course of action based on long-term benefit. Charging rate was dynamic, allowing the controller to learn a balanced approach between health and speed.

Training consisted of a hybrid approach of both offline and online strategies. In the offline phase, the DQN used previously generated baseline and CCCV modified simulations to learn degradation-related patterns before interacting with the simulated conditions. During online evaluation, the controller interacted directly with the PyBaMM environment. It selected a charging-rate action, observed the updated battery state, and updated its policy accordingly. In order to stabilize the training strategy, an experience replay buffer, a target network, and epsilon-greedy exploration were utilized. The replay buffer stored past charging states, actions, and the rewards or penalties that ensued, allowing the model to sample broader scale behavior. The separate target network was updated less frequently than the primary policy network to mitigate unstable action-value estimates while training. Epsilon-greedy exploration allowed the model to opt for alternative charging rates to observe the impacts on the battery state under each condition and identify the best action to fulfill the overall goal [21].

$$y_t = r_t + \gamma \max_{a'} Q_{\text{target}}(s_{t+1}, a') \tag{3}$$

Equation 3. y_t is the temporal-difference target, which is the value the DQN is trained to predict by combining the immediate reward with the estimated best future reward. r_t is the immediate reward assigned at time step t, γ is the discount factor that controls how strongly future rewards are valued, and a' represents a possible next action. The max term selects the highest predicted future action value across all possible next charging actions. $Q_{\text{target}}(s_{t+1}, a')$ is the target-network estimate of the action value for the next battery state s_{t+1} and possible next action a', which helps stabilize DQN training.

The DQN was implemented as a feedforward neural network with 15 input battery-state variables and 5 output nodes, which are the charging-rate actions of 0.5C, 0.75C, 1.0C, 1.25C, and 1.5C. The internal structure of the network consists of three hidden layers with 128, 128, and 64 neurons, with Rectified Linear Unit (ReLU) activation functions after each layer [22]. The hyperparameters selected for this study were all author-chosen for implementation. The Adam optimizer was used to train the model with a learning rate of 0.001, a discount factor $\gamma = 0.97$, and a batch size of 128. The learning rate was set to this value to train without instability and the discount factor is set close to 1 to encourage the model to focus on long-term decisions. The replay buffer stored a maximum of 20,000 transitions and training updates began after a minimum of 1,000 transitions were available. Epsilon-greedy

© The National High School Journal of Science 2026

exploration began at $\varepsilon = 1.0$ and decayed by a factor of 0.992 after each episode with a lower bound of $\varepsilon = 0.05$, gradually minimizing the randomness of the model's actions. The target network was updated every 10 episodes, with eight training updates per episode. It was trained for 400 episodes with an expert warm-start to begin with a stronger baseline rather than just exploring randomly.

The observation vector for this RL model consisted of the same metrics as the other protocols to ensure consistency in evaluation. This included SOC, intrinsic and ambient temperature, EFC, LLI, CDI, CDI rate, voltage, charging rate, and other PyBaMM-resolved metrics. Ambient temperature was initialized at 25°C during evaluation to maintain a semi-realistic environment. The controller is able to choose from a discrete charging-rate action space of 0.5C, 0.75C, 1C, 1.25C, and 1.5C, so that the model is able to balance speed and battery health preservation.

DQN training can often vary due to stochastic exploration and initialization for the neural network, which impacts the stability of the results significantly. To evaluate stability, the DQN was retrained and evaluated across five random seeds: 1, 7, 21, 42, and 100. The same DQN architecture and simulation conditions were maintained to keep consistency. Mean values of metrics across all five seeds were reported. Repeated training stability was evaluated using standard deviation, 95% confidence intervals (CIs), and consistent compliance of the expected operating windows, mentioned in the Deep Q-Network Evaluation Criteria section. 95% CIs allowed for a range of values for metrics to be produced rather than just one inconsistent value.

Penalty-Reward Function

All reward and penalty coefficients were manually selected. The coefficients were chosen heuristically to prioritize the reduction of the Composite Degradation Index (CDI) and several other degradation metrics, thermal safety, SOC-window compliance, charging-rate optimization, and charge progression. Incremental LLI and LAM were given penalties of 3000 and 2500, respectively. The strongest penalty was assigned to incremental SEI capacity loss, with a weight of 15000. Increases in CDI were penalized with a weight of 8000 per increment each cycle. Exceeding the predefined Composite Degradation Index (CDI) threshold of 0.035 resulted in a penalty weight of 4000. Temperature deviation from 25–35°C was penalized with a weight of 1150. CDI rate exceeding 1.5×10^{-6} min^{-1} held a penalty of 150. SOC outside the bounds of 25–85% was penalized by a weight of 160 to prevent deep cycling. Deviations from the expected charging rate range were given a penalty of 80 to discourage aggressive currents. When charging duration eclipsed 60 minutes, a penalty with a weight of 0.08 was applied. The same coefficient magnitudes were applied as a reward for the opposite of these conditions.

A penalty weight of 50 was applied when fast charging was conducted within 3°C of the upper thermal limit, training the model to respond before overheating occurred. Fast charging was also penalized when conducted above 75% SOC by a weight of 45. A penalty of 220 was applied if the upper SOC boundary was not reached, indicating to the model that complete oscillation through this window was required. Even if successful charge completion occurred, the model would be penalized by a weight of 80 if it occurred under unsafe thermal conditions.

Rewards were assigned as well. A weight of 100 was applied for a progression in SOC, encouraging the model to avoid stalling to reduce penalties. A weight of 60 was applied when the charging duration goal of 60 minutes was met. After each charging episode was completed, a weight of 120 was granted if the 25–85% SOC window was complied with. A bonus of 20 was also applied when the battery reached the upper SOC boundary within the configured charging-duration limit.

Penalty and reward weight tuning was conducted to test whether reward emphasis altered controller behavior. The health-centric setup increased the CDI group weights by 1.5×, physical degradation group by 2.0×, and temperature group by 1.25×, while reducing the time-focused group to 0.75×. The temperature-prioritized setup increased the temperature group by 2.0×, as well as the charging-rate and CDI-rate group by 1.5×, while reducing the time group to 0.75×. The speed-prioritized setup increased the SOC group to 1.25× and the time group to 2.0×, while reducing the CDI, physical degradation, as well as charging-rate and CDI-rate groups to 0.75×. The CDI-prioritized setup increased the CDI group by 2.0×, as well as the charging-rate and CDI-rate group by 1.25×, while reducing the time group to 0.75×. The degradation-focused setup increased the physical degradation group by 2.0× while reducing the time group to 0.75×. The balanced alternate setup increased the CDI, physical degradation, and temperature groups by 1.25×. Each penalty-reward variation was evaluated across three seeds, which were 1, 42, and 100, and in comparison to the default weighting system.

Deep Q-Network Evaluation Criteria

The learned RL policy was applied over 200 cycles under the same 25–85% SOC protocol, 3.0 Ah charge/discharge throughput per cycle, and 120 EFC endpoint used by all policies. CDI thresholds, compliance windows, current-action bounds, charging-duration targets, SOC tolerances, goal-compliance percentages, and stability criteria were selected by the authors for this study, and are set purely for evaluation and direct comparisons.

The model's ability to reduce degradation-associated met-

rics relative to standard Constant Current-Constant Voltage (CCCV) and other tested policies is significant to its evaluation. The fraction of cycles in which particular operating ranges were maintained to do so was tracked for a few metrics, with author-selected percentages used. Battery temperature is predicted to be maintained within 25–35°C for at least 85% of the time. It is expected that the charging rate will remain between approximately 0.5C to 1.5C for ∼80% of the entire simulation, balancing both charging speed and health. Charging duration will be monitored and it is expected to remain close to ∼60 minutes for about 65% of the time, with reductions of approximately 90 minutes from the average baseline charging duration without compromising battery health.

The model should produce a CDI of 0.035 or below over a period of 200 cycles in the simulation, utilizing these goals to contribute to long-term performance and health. Instantaneous CDI rate will additionally be monitored and this model will maintain a rate below 1.5×10^{-6} min^{-1} to avoid risks such as the destabilization of SEI layers due the acceleration of side reactions. Charging protocols that exit these author-selected bounds are considered unsafe because they may contribute to increases in degradation.

CCCV simulation results were used as a baseline for the evaluation of the controller's physical degradation outputs. The DQN was expected to remain below the CCCV values for LLI, LAM, capacity loss to SEI, and capacity fade percentage, while maintaining higher discharge capacity. Achieving this would result in DQN performance being considered favorable. However, the baseline for charging duration was not set at below 72.0 minutes because the DQN is expected to attain a stricter goal of a charge time below 60 minutes to provide a practical reduction. It could, therefore, claim success in minimizing charge time from the standard.

The DQN controller was considered stable across multiple seeds if the standard deviation for the battery metrics mentioned above remained within 5% of the five-seed mean. The 95% confidence interval was also expected to remain with a half interval width of at most 5% of the five-seed mean for all metrics. All five seeds were expected to meet or exceed thermal and charging compliance goals, as well as the degradation and charging performance targets.

For evaluation of the penalty-reward weight tuning, each variation was compared to the default penalty-reward function using the same health and performance metrics. Variations were considered to significantly impact controller behavior if key metrics were altered by more than 5% from the baseline DQN metrics. The 5% deviation threshold was applied to all aforementioned degradation metrics, mean charging duration, mean charging rate, mean and maximum temperature, and EFC. In addition, all compliance windows and targets for battery metrics are expected to be met for manual selection of weight variations to not be considered significantly impactful to the study.

Policy rankings were additionally used to evaluate the DQN's overall success in comparison to the other charging strategies. The DQN was expected to rank first for final CDI, CDI per cycle, CDI per EFC, LLI, capacity loss to SEI, discharge capacity, and capacity fade percentage, as these were the primary indicators of degradation reduction and capacity retention. For negative LAM, positive LAM, and resistance growth, the DQN was expected to rank 2nd or 3rd, since these metrics are likely more favorable to the lowest-current policies. In terms of charging duration, the DQN was expected to be faster than the CCCV policy, the fixed 0.72C policy, and the heuristic. However, it was expected to be slower than fixed 1.0C, fixed 1.5C, and the greedy controller as they would reduce charging time with faster rates, but increase degradation substantially. The DQN was evaluated in terms of its ability to provide a strong balance between degradation prevention and charge-time minimization.

Results

CDI Weight Sensitivity Analysis

Across all six methods, the Deep Q-Network (DQN) remained the policy with the lowest Composite Degradation Index (CDI). The DQN had CDI values of 0.03212 under the default case, 0.028806 under equal stress weighting, 0.035947 under high-current weighting, 0.036845 under high-voltage weighting, 0.038504 under high-thermal weighting, and 0.021714 under low-stress weighting. The Constant Current-Constant Voltage (CCCV) policy ranked second in all cases. Fixed 1.5C consistently produced the highest CDI value, the greedy policy remained with the second highest CDI, and the fixed 1.0C policy consistently held the third highest value. The fixed 0.72C tapering and the heuristic interchanged between the third and fourth lowest positions in terms of CDI across all policies. These policy rankings indicate that the manually selected weights to compose CDI didn't provide an advantage unique to the DQN policy.

Regression Models

Standardized Ridge coefficients showed that Effective Full Cycles (EFC) and mean cycle temperature were generally the strongest associated predictors for degradation metrics. Loss of Lithium Inventory (LLI) had coefficients of 1.1335 for EFC, 1.0038 for temperature, 0.2051 for charging rate, 0.0250 for State of Charge (SOC), and 0.0178 for voltage. Capacity loss to Solid Electrolyte Interphase (SEI) followed this pattern, with 0.0845 for EFC, 0.0773 for temperature, 0.0163 for charging rate, 0.00193 for SOC, and 0.00130 for voltage. Capacity fade percentage and discharge capacity showed simi-

© The National High School Journal of Science 2026

lar high aging-related coefficients. CDI slightly differed, with 0.0287 for temperature, 0.0225 for EFC, 0.00421 for charging rate, 0.000718 for voltage, and 0.000687 for SOC. Loss of Active Material (LAM) and resistance-growth coefficients showed smaller coefficients and inconsistent predictor patterns.

Results from the held-out 3.0C condition indicated that the Ridge regression was not able to translate and predict battery health under an aggressive charging condition not included in training. It performed best for negative and positive LAM, with the strongest R^2 values, lowest NRMSE, and small residual error. For all other metrics, including CDI, discharge capacity, and LLI, a substantially weaker performance was conducted. Some R^2 values dipped into the negatives, while certain NRMSE values eclipsed 45%. Residual spreads were also much larger. These highlight how the model was able to track LAM behavior, but was unreliable overall in terms of broader degradation outputs under aggressive charging.

Held-out 45°C validation indicated that the Ridge regression model again showed limited reliability. The model particularly held the highest ranking metrics for LLI and capacity fade. These targets maintained moderate R^2 and NRMSE values, with a more controlled residual error. Negative and positive LAM performed poorly, with substantially negative values for R^2 and near-100% NRMSE values. The model's ability to capture a minimal amount of temperature-influenced degradation trends was demonstrated here, as it slightly tracked the shift in some metrics, including lithium inventory loss, capacity retention, and SEI-related degradation. Its reliability declined when faced with active material loss and resistance growth, marking a complete reversal in the model performance metrics from the held-out 3.0C condition.

Penalty-Reward Weight Tuning

Across three seeds, the default reward setup produced a mean CDI of 0.03180, mean charging duration of 44.78 minutes, LLI of 1.49188%, SEI capacity loss of 0.10870 Ah, negative/positive LAM of 0.06939%/0.06374%, capacity fade of 2.27856%, discharge capacity of 3.92217 Ah, resistance growth of 0.00572 Ohm, mean applied charging rate of 0.804C, mean battery temperature of 28.65°C, maximum battery temperature of 34.21°C, and mean CDI rate of 1.36×10^{-6} min^{-1}.

Across the three seeds, for the non-default reward settings, final CDI remained within [0.03163, 0.03179], mean charging duration within [44.34, 45.27] minutes, LLI within [1.49063, 1.49419]%, SEI capacity loss within [0.10860, 0.10890] Ah, negative LAM within [0.06934, 0.06947]%, positive LAM within [0.06326, 0.06396]%, capacity fade within [2.27671, 2.28198]%, discharge capacity within [3.92199, 3.92227] Ah, resistance growth within [0.00572, 0.00573] Ohm, mean applied charging rate within [0.796, 0.812]C, mean battery temperature within [28.62, 28.67]°C, and mean CDI rate within $[1.35 \times 10^{-6}, 1.37 \times 10^{-6}]$ min^{-1}. Charge-time compliance ranged between [71.5, 74.5]%, the charging-rate compliance window was [98.5, 100]%, and temperature and charging rate both had a window compliance of 100%.

No variation changed any measured output by more than the established limit of 5%. The largest same-seed change was 3.94% for temperature-prioritized charging duration, where for seed 100 the value increased from 44.78 to 46.55 minutes. A change of 3.38% was caused in the mean charging rate by the balanced alternate variation, where for seed 1 the value increased from 0.801C to 0.828C. Other shifts that took place additionally fell below the predefined threshold, deeming the DQN outputs not sensitive to the tested penalty-reward variations.

DQN Evaluation

Across the five DQN seeds, the controller showed stable outputs for all of its operating behavior and degradation metrics. Final CDI, mean CDI rate, LLI, SEI capacity loss, negative and positive LAM, capacity fade, discharge capacity, and resistance growth all held narrow min-max ranges. Their standard deviations and 95% CI half-widths all remained below the stability target of 5% from the mean. For degradation metrics, discharge capacity held the narrowest variations for SD and CI half-width, while the mean CDI rate held the highest. All metrics except positive LAM and resistance growth were able to meet the predefined CCCV-based threshold. These results indicate that DQN was able to maintain favorable, but not entirely superior, degradation and charging performance relative to CCCV.

Temperature-window compliance, charging-rate-window compliance, and CDI-rate upper-limit compliance were each maintained at 100% for all seeds, leading their SDs and CI-half-widths to remain at 0%. The percentage of cycles completed within 60 minutes remained above the expected target. Consistency in the results demonstrates the low variability of DQN, ensuring one seed didn't substantially differ to skew degradation and performance outcomes.

Low variation in results likely occurred due to all seeds following the same settings, including PyBaMM configuration, action space, SOC window, EFC endpoint, etc. These fixed settings lessen the range of possible values. Discharge capacity likely varied the least due to the 120 EFC endpoint for all seeds, which limited the change to capacity retention. Charge-duration compliance varied the greatest amount because of charging decisions the DQN could've taken within its relatively wide-ranging action space. The consistent 100% compliance for temperature window, charging-rate window, and CDI-rate limit indicate that the learned policy actively avoided

Table 1 Ridge regression performance under the held-out 3.0C condition. Results are reported for each target using R^2, normalized root mean squared error, 95 % confidence intervals, mean residual, and residual standard deviation.

Target	R^2	NRMSE (%)	R^2 95 % CI	NRMSE 95 % CI (%)	Mean Residual	Residual SD
CDI	-1.614	46.85	[-1.841, -1.443]	[44.02, 50.22]	-0.002	0.017
Discharge capacity	-0.121	30.68	[-0.194, -0.067]	[28.94, 32.72]	0.002	0.053
LLI	-0.077	30.07	[-0.144, -0.028]	[28.42, 32.06]	0.003	0.690
SEI capacity loss	-0.274	32.71	[-0.353, -0.220]	[30.93, 34.86]	0.000747	0.054
Capacity fade	-0.176	31.43	[-0.248, -0.127]	[29.73, 33.49]	0.014	1.086
Negative LAM	0.870	10.48	[0.857, 0.878]	[9.84, 11.17]	-0.003	0.012
Positive LAM	0.435	21.78	[0.347, 0.494]	[20.19, 23.51]	-0.014	0.026
Resistance growth	-3.851	58.13	[-5.031, -3.032]	[55.27, 63.40]	5.629×10^{-9}	3.101×10^{-9}

Table 2 Ridge regression performance under the held-out 45°C condition. Results are reported for each target using R^2, normalized root mean squared error, 95 % confidence intervals, mean residual, and residual standard deviation.

Target	R^2	NRMSE (%)	R^2 95 % CI	NRMSE 95 % CI (%)	Mean Residual	Residual SD
CDI	0.100	27.53	[-0.054, 0.209]	[25.69, 29.55]	0.031	0.040
Discharge capacity	0.442	21.70	[0.356, 0.501]	[20.31, 23.20]	-0.067	0.109
LLI	0.475	20.95	[0.399, 0.526]	[19.62, 22.43]	0.767	1.423
SEI capacity loss	0.464	21.16	[0.388, 0.516]	[19.82, 22.67]	0.058	0.110
Capacity fade	0.478	20.98	[0.404, 0.528]	[19.66, 22.42]	1.147	2.195
Negative LAM	-9.579	94.26	[-11.454, -8.292]	[87.34, 101.72]	0.017	0.021
Positive LAM	-7.920	86.64	[-9.190, -7.062]	[80.63, 93.08]	0.011	0.021
Resistance growth	-3.946	48.70	[-4.952, -3.120]	[46.20, 54.23]	-3.158×10^{-8}	1.571×10^{-8}

behavior that would violate the constraints.

SPM Policy Comparison

For comparison purposes, the five-seed mean for every Deep Q-Network (DQN) output was used. For a majority of the metrics measured, DQN was able to rank highest. It produced the lowest CDI, LLI, SEI capacity loss, capacity fade, and negative LAM, while holding the highest retained discharge capacity. The fixed 1.0C was the closest competitor to the DQN, maintaining only slightly larger values for a majority of metrics. Although this was the case, it produced a higher CDI relative to DQN. This highlights how physical degradation outputs did not always translate into the same overall CDI ranking, since the constructed metric also incorporated charging-rate, thermal, and time-related stress behavior. CCCV and the 0.72C policy performed better in terms of mean battery temperature, resistance growth, and positive LAM, but their low-stress behavior contributed to a significantly longer charging time while still producing higher values for other degradation metrics like LLI and CDI. Charging duration trends indicated the expected tradeoff between speed and health prioritization for other policies. DQN was not able to rank first for charge time, which was expected, as fixed-rate policies with faster charging rates and more aggressive tendencies were included in the comparison. The greedy controller, fixed 1.0C, and fixed 1.5C all followed this trend, while contributing to higher degradation values for CDI, LLI, SEI loss, etc., as well as higher thermal stress. The DQN therefore held an intermediate charging-speed range, charging substantially faster than CCCV and fixed 0.72C while avoiding the stronger degradation penalties seen in the fastest policies. The heuristic fell in between these groups, charging much faster than CCCV and the fixed 0.72C policy, but not matching the DQN for degradation metrics. DQN was able to exceed all predefined operating window compliance thresholds. The DQN maintained 100% compliance for the temperature window, charging rate window, and CDI-rate limit. It also meets the charge-duration target for three-quarters of all cycles. CCCV and the 0.72C policy were slightly safer thermally, but due to the low-stress and slow charging they brought, they didn't meet the charging-duration compliance threshold. The faster fixed-rate policies and greedy controller were able to exceed the standard DQN set for charge-duration compliance, but fell far below in terms of degradation metrics. These results highlight how DQN did not need to utilize aggressive behaviors or violate compliance thresholds to attain lower degradation and charging-performance outputs. With these metrics, it is apparent that the primary advantage of DQN was not ranking highest in each isolated metric, but providing the strongest balance between degradation reduction and practical charging.

SPMe and DFN Validation

Under SPMe and DFN, DQN remained one of the strongest policies, ranking first for SEI capacity loss, capacity fade, discharge capacity, negative LAM, and resistance growth. However, CCCV ranked highest for CDI in both models and LLI in SPMe, attaining the lowest values in those metrics while DQN held the second-lowest. Positive LAM continued to be a weak point for the learned policy in terms of degradation, with CCCV and the 0.72C policy ranking higher. A major-

© The National High School Journal of Science 2026

Table 3 Five-seed stability results for the Hybrid DQN controller. Metrics are reported using the five-seed mean ± standard deviation, min-max range, 95 % confidence interval, standard deviation as a percentage of the mean, and CI half-width as a percentage of the mean.

Metric	Mean ± SD	Min-Max Range	95 % CI	SD (% of mean)	CI Half-Width (% of mean)
Final CDI	0.03238 ± 0.00020	[0.03213, 0.03268]	[0.03212, 0.03263]	0.63%	0.78%
Mean CDI rate	$1.42 \times 10^{-6} \pm 1.08 \times 10^{-8}$ min^{-1}	[1.41×10^{-6}, 1.44×10^{-6}] min^{-1}	[1.41×10^{-6}, 1.43×10^{-6}] min^{-1}	0.76%	0.95%
LLI	1.48126 ± 0.00104%	[1.48032, 1.48282]%	[1.47997, 1.48255]%	0.07%	0.09%
SEI capacity loss	0.10781 ± 0.00009 Ah	[0.10774, 0.10794] Ah	[0.10770, 0.10792] Ah	0.08%	0.10%
Negative LAM	0.06916 ± 0.00003%	[0.06912, 0.06919]%	[0.06913, 0.06919]%	0.04%	0.05%
Positive LAM	0.06556 ± 0.00020%	[0.06527, 0.06576]%	[0.06531, 0.06580]%	0.30%	0.37%
Capacity fade	2.26294 ± 0.00167%	[2.26155, 2.26548]%	[2.26087, 2.26502]%	0.07%	0.09%
Discharge capacity	3.92301 ± 0.00008 Ah	[3.92288, 3.92308] Ah	[3.92291, 3.92311] Ah	0.002%	0.003%
Resistance growth	0.00569 ± 0.00001 Ohm	[0.00568, 0.00570] Ohm	[0.00568, 0.00569] Ohm	0.09%	0.11%
Mean battery temperature	28.81 ± 0.01°C	[28.79, 28.83]°C	[28.79, 28.83]°C	0.05%	0.06%
Mean charging duration	42.00 ± 0.25 min	[41.75, 42.36] min	[41.69, 42.31] min	0.60%	0.74%
% cycles less than or at 60 minutes	76.3 ± 0.67%	[75.5, 77.0]%	[75.47, 77.13]%	0.88%	1.09%
Mean charging rate	0.857 ± 0.005C	[0.850, 0.862]C	[0.851, 0.863]C	0.60%	0.74%

Table 4 Comparison of key metrics under SPM, including the policies: the Hybrid DQN; baseline CCCV; fixed 0.72C taper, 1.0C, and 1.5C charging policies; heuristic; and greedy controller. The highlighted value corresponds to the top ranking policy for the respective metric.

Policy	CDI	Avg. Charge Time (min)	LLI (%)	SEI Loss (Ah)	Neg/Pos LAM (%)	Capacity Fade (%)	Discharge Capacity (Ah)	Resistance Growth (Ohm)	Mean Temp. (°C)
DQN (5-Seed mean)	**0.03238**	42.00	**1.48126**	**0.10781**	**0.06916** / 0.06556	**2.26294**	**3.92301**	0.00569	28.81
CCCV	0.03372	72.00	1.58906	0.11700	0.07117 / **0.04426**	2.42746	3.91470	0.00562	**27.40**
Fixed 0.72C Taper	0.03862	65.00	1.64989	0.12118	0.07242 / 0.05196	2.51655	3.91022	**0.00527**	27.79
Fixed 1C	0.03951	40.15	1.48455	0.10800	0.06919 / 0.06900	2.26650	3.92281	0.00569	28.87
Fixed 1.5C	0.04959	35.57	1.51335	0.10955	0.06947 / 0.08106	2.29537	3.92136	0.00576	29.49
Heuristic	0.03887	48.48	1.51323	0.11051	0.07029 / 0.05962	2.31210	3.92051	0.00578	28.39
Greedy Controller	0.04870	**35.13**	1.50265	0.10878	0.06926 / 0.08044	2.28133	3.92207	0.00573	29.48

ity of the degradation rankings were retained from SPM, but some were altered under the higher-fidelity models, confirming that the DQN was at least strongest in capacity-retention and preventing SEI-related degradation across all models.

Charging duration and temperature also held similar rankings under both of these models. The same tradeoff was experienced by faster policies, where they sacrificed battery-health for charging speed. Fixed 1.0C and 1.5C policies remained the fastest policies in SPMe and DFN, while CCCV and the 0.72C policy were still the slowest in both. Regarding mean battery temperature, CCCV retained its position as the lowest-temperature policy, while DQN continued to hold an intermediate rank in both models, as done in SPM.

The higher-fidelity models allowed for local anode-potential behavior to be monitored. In SPMe, the DQN reached a minimum local anode potential of −0.06235 V and an anode-below-0-V fraction of 0.18144. In DFN, the DQN reached a minimum local anode potential of −0.05024 V and an anode-below-0-V fraction of 0.15522, performing with similar results as the DQN trial in SPMe. In both models, DQN performed better than the most aggressive policies, fixed 1.5C and the greedy controller, but worse than CCCV and the fixed 0.72C tapering policy. Since local anode potentials below 0 V vs. Li/Li^+ indicate increased lithium-plating risk, these results show that the DQN did not entirely eliminate fast-charging-related plating-risk factors.

The SPMe and DFN outputs partly support the primary SPM trends, but didn't entirely reproduce the same rankings. These results overall support the primary conclusion that DQN was able to balance degradation and charging-performance well, even if it still lacked the highest rank in a few categories under higher-fidelity models.

Discussion

The primary goal of this study was to determine whether an adaptive charging protocol could reduce degradation relative to Constant Current-Constant Voltage (CCCV) and additional charging protocols while maintaining charging efficiency. The Deep Q-Network (DQN) met most of its objectives. Relative to CCCV, the primary policy for comparison, the DQN policy was able to attain a CDI reduction of 0.00134 and a charging-duration reduction of 30.0 minutes. In the SPM policy comparison, DQN did not outperform every policy for every individual output, but ranked first for CDI, LLI, SEI capacity loss, capacity fade, negative LAM, and discharge capacity. CCCV and fixed 0.72C tapering held stronger results for mean temperature, positive LAM, and resistance growth. Fixed 1.0C, fixed 1.5C, and greedy control charged faster than DQN but produced worse degradation and thermal behavior. These results highlight the learned policy's ability to reduce early-cycle degradation accumulation.

The stability and sensitivity analyses supported this interpretation. Across five DQN seeds, all observed metrics showed narrow standard deviations and 95% confidence intervals beneath the 5% target. Penalty-reward weight tuning also

Table 5 Policy-ranking comparison under SPMe and DFN validation. Rankings compare the charging policies across degradation, charging-duration, and mean-temperature metrics to determine whether the main SPM policy trends were maintained under higher-fidelity models.

Metric	SPMe ranking	DFN ranking
CDI (lowest to highest)	CCCV, DQN, Heuristic, 0.72C, 1.0C, Greedy, 1.5C	CCCV, DQN, 0.72C, Heuristic, 1.0C, Greedy, 1.5C
Charging duration (fastest to slowest)	1.5C, 1.0C, DQN, Greedy, Heuristic, 0.72C, CCCV	1.5C, 1.0C, DQN, Greedy, Heuristic, 0.72C, CCCV
LLI (lowest to highest)	CCCV, DQN, Heuristic, 1.0C, Greedy, 0.72C, 1.5C	DQN, 1.0C, Heuristic, CCCV, Greedy, 1.5C, 0.72C
SEI capacity loss (lowest to highest)	DQN, Heuristic, CCCV, 1.0C, Greedy, 0.72C, 1.5C	DQN, 1.0C, Heuristic, CCCV, Greedy, 1.5C, 0.72C
Capacity fade (lowest to highest)	DQN, Heuristic, CCCV, 1.0C, Greedy, 0.72C, 1.5C	DQN, 1.0C, Heuristic, CCCV, Greedy, 1.5C, 0.72C
Discharge capacity (highest to lowest)	DQN, Heuristic, CCCV, 1.0C, Greedy, 0.72C, 1.5C	DQN, 1.0C, Heuristic, CCCV, Greedy, 1.5C, 0.72C
Negative LAM (best to worst)	DQN, 1.0C, 1.5C, Greedy, Heuristic, CCCV, 0.72C	DQN, 1.0C, 1.5C, Heuristic, Greedy, CCCV, 0.72C
Positive LAM (lowest to highest)	CCCV, 0.72C, Heuristic, DQN, 1.0C, 1.5C, Greedy	CCCV, 0.72C, Heuristic, DQN, 1.0C, Greedy, 1.5C
Resistance growth (lowest to highest)	DQN, Heuristic, CCCV, 1.0C, Greedy, 0.72C, 1.5C	DQN, Heuristic, 1.0C, CCCV, Greedy, 0.72C, 1.5C
Mean cycle temperature (lowest to highest)	CCCV, Heuristic, 0.72C, Greedy, DQN, 1.0C, 1.5C	CCCV, Heuristic, 0.72C, Greedy, DQN, 1.0C, 1.5C

only produced slight variations in outputs, with none exceeding the 5% threshold. SPMe and DFN validation confirmed some conclusions, but shifted DQN's rankings for CDI and LLI. Negative local anode potentials also remained, indicating that DQN didn't remove all lithium-plating risk indicators under higher-fidelity models.

Implications and Relevance to Existing Literature

To ensure a fair comparison, all policies used the same 25–85% SOC window, 3.0 Ah charge/discharge throughput, and 120-EFC endpoint. Matching throughput ensured that each policy transferred the same amount of capacity through the cell every cycle, preventing a protocol from appearing more degradative if it transferred higher amounts. Experimental aging studies show how cycling is associated with capacity loss and resistance growth[23].

SOC windows also influence long-term degradation, so maintaining a consistent range allowed comparison to focus directly on each policy's health-prioritizing behavior[1].

Frequent cycling exposure was associated with higher degradation outputs, but EFC was not treated as the sole cause. Over substantial cycling, repeated ion insertion and extraction events can generate structural stress, contributing to particle cracking and SEI growth over newly exposed surface area[3,24]. As SEI growth progresses, cyclable lithium is consumed, resulting in Loss of Lithium Inventory (LLI) and increasing degradation[3]. Because these processes are mainly irreversible, cumulative degradation metrics become closely aligned with cycle accumulation[25].

The mechanism is part of cycle-induced aging, in which frequent cycling and storage periods reduce capacity and affect State of Health (SOH) in EV applications[26]. Therefore, using the same SOC window, throughput per cycle, and EFC endpoint helped minimize the role of battery-aging in the comparison.

Mean battery temperature rankings followed the expected relationship between charging rate and thermal behavior, reinforcing the role of thermal conditions on battery health. CCCV, fixed 0.72C, and the heuristic produced the lowest battery temperatures due to lower-current charging reducing heat-generation demand, while higher-current policies brought higher temperatures[27]. This occurrence indicates why DQN, with an average charging rate of 0.857C, did not rank first or last for mean temperature. Prior charging studies have similarly shown that current, voltage, and SOC operating range can affect cycle life and degradation accumulation[1,28].

The thermal trends mirror Arrhenius behavior, where elevated temperature can accelerate aging and SEI-related degradation[29]. In the model used in this study, SEI formation was represented by electrolyte decomposition occurring on the lithiated graphite surface[3].

Reaction byproducts then accumulate on the negative electrode, with continued growth consuming cyclable lithium and resulting in LLI[3]. Fixed 1.0C, fixed 1.5C, and the greedy controller reflect this pattern, producing higher degradation outputs than lower-temperature policies.

Charging duration trends reflected the tradeoff between speed and degradation control. Higher-current charging can increase heat generation, lithium-plating risk, SEI growth, current-induced stress, and mechanical electrode damage, helping explain why the fastest policies had weaker degradation results and higher battery temperatures[2,7,27]. Even without high-current charging, DQN significantly reduced charge time from CCCV while remaining within 10 minutes of the fastest charging policy. This matters over repeated cycling because battery aging increases resistance-related losses and reduces usable capacity[23,26].

DQN's lower charging duration suggests that its adaptiveness for minimizing degradation-raising currents allowed it to maintain practical charging.

DQN maintained strong results across most degradation metrics, including LLI, SEI capacity loss, capacity fade, and discharge capacity. This was likely caused by its dynamic current selection limiting the effects of high-current stress, temperature rise, and prolonged exposure to side reactions. Lower SEI loss and LLI aligned with its lower capacity fade and higher discharge capacity because SEI growth consumes us-

© The National High School Journal of Science 2026

able lithium, reducing the total retained battery capacity[3].

Fixed 1.0C was closest to the DQN policy in several metrics because it held a similar less-aggressive current. However, it still accumulated a higher CDI, indicating its inability to mitigate stress as effectively as DQN. Fixed 1.5C and greedy control ranked worse because their faster charging increased current-related and thermal stress, which accelerates degradation[2,27].

CCCV and the fixed 0.72C tapering policy ranked lower than DQN in several metrics despite lower charging rates, revealing how longer charging duration and cycling exposure still contributed significantly to degradation accumulation[23,26].

Under aggressive charging, repeated lithiation and delithiation can raise mechanical stress in graphite, causing negative-electrode cracking and active-material loss to be more likely[24].

This may explain why higher-current policies in this study produced higher negative LAM, while lower-current policies like DQN produced the lowest. For positive LAM and resistance growth, however, CCCV and fixed 0.72C tapering ranked better, suggesting that these outputs possibly favored primarily the lowest charging currents in the tested model.

The controller exceeded its charging-duration, CDI-rate limit, and temperature-window compliance goals, suggesting that the decline in the CDI and other degradation metrics was not simply a result of harsh constraints or incomplete charging, rather its moderate and adaptive behavior. However, CCCV and other policies remained favorable for some outputs, while CDI and LLI rankings shifted during SPMe and DFN validation. These results should therefore be interpreted as a model-specific validation of DQN performance, not proof that it would produce the same outputs under varied simulation settings. The outputs, however, are consistent with health-prioritized and controlled charging studies that balance charging duration with incorporated electrochemical, thermal, degradation, or plating-risk constraints[6,7,15].

Limitations

Limitations still exist, so the study should be interpreted as a simulation-based proof of concept. The simulated outputs are resolved by PyBaMM and remain susceptible to the assumptions of SPM and the OKane2022 parameter set. Even though these incorporate battery-physics-aligned equations, they remain approximations of real battery behavior. CDI is also a manually constructed degradation-stress index rather than direct simulated capacity fade, so other PyBaMM-resolved metrics should receive greater emphasis. The regression analysis is limited by its small simulation space, so Ridge coefficients can only be interpreted as feature associations under the tested conditions. Although penalty-reward weight tuning was included, only a limited amount of variations were tested and they do not represent all possible structures. Higher-fidelity validation under SPMe and DFN also showed remaining plating-risk indicators.

Another limitation is that the DQN received internal PyBaMM degradation states, including LLI, CDI, and CDI rate, as part of its observation vector. A real Battery Management System (BMS) would be required to estimate them from measurable signals such as terminal voltage, current, and temperature[30]. Due to this, the controller should be perceived as a simulation-based demonstration rather than one ready for BMS deployment. The study also did not include sensor noise in voltage, current, temperature, or SOC measurements. Since a BMS relies on limited and noisy signals in real-world operation, DQN decisions may be less reliable near voltage, SOC, or thermal boundaries[30].

The discharge process in the study was controlled and idealistic, but is erratic in realistic scenarios. Real EV discharge cycles include acceleration, braking, regenerative braking, traffic conditions, different trip lengths, and climate-control use. Each of these actions can shift SOC, temperature, current demand, and degradation behavior in ways that were unaccounted for by the fixed discharge phase. All simulations also used the same battery parameter set, so battery-to-battery variation was not tested. In real batteries, cells may differ in capacity, aging behavior, and impedance, so the same charging policy may behave differently across parameter sets[31].

In practical applications, applying the 25–85% SOC bound is not always feasible because long-distance drivers may require the full available driving range, even if that contributes to long-term degradation. The study additionally remained at the cell level, excluding pack-level balancing and vehicle-level implementation. Future experimental validation on physical battery cells would be significant in revealing the model's functionality in real-world settings, rather than just theoretical simulations. Before vehicle-level implementation, a state-estimation layer would need to be added before applying this controller to infer internal degradation and battery states from measurable signals like voltage, current, temperature, and time-series information. Studies have demonstrated RL charging optimization using real experimental battery data and ML estimation of anode plating potential from measurements from three-electrode experiments, illustrating possible paths for future implementation[11,12]. Realistic driving variability could be incorporated to no longer assume a fixed discharging rate and allow external variables to influence the battery state. Future versions of the controller should include stronger anode-potential or lithium-plating-risk constraints, since SPMe and DFN validation showed that these indicators were not fully removed by the learned DQN policy. More cycles could be run as well to verify whether these trends in the outputs persist or possibly extend in impact on total battery lifespan. The results

should be interpreted as an early-cycle reduction in degradation accumulation rather than proof of an extension to the lifespan of the battery. The controller should be tested under realistic drive cycles, noisy sensor measurements, and multiple battery parameter sets to validate whether DQN results can persist under more practical, varied conditions.

Conclusion

The results support the overall proposition of this paper, that degradation-aware adaptive charging can produce positive battery health outcomes while maintaining practical charging. Under matched settings, the DQN produced the lowest CDI, as well as several degradation metrics, among the SPM policies while maintaining a lower charging duration. Although DQN did not outperform every policy in all metrics, these findings indicate that within the existing simulation framework, it is able to provide the strongest overall balance between battery-health preservation, practical charging, and operating-window compliance. Even though the charging protocol may not be able to perform similarly universally, the study provides a useful simulation-based foundation for future research aimed at improving EV battery lifespan.

Acknowledgments

Much thanks to Dr. Adam Li, Dr. Xiao Dong, and Polygence for mentorship and guidance throughout the development of this study as well as the provision of the framework to facilitate the experimentation. I would like to thank Savar Shandilya, a fellow student at Washington High School, for offering feedback to organize the paper as well.

References

1 N. Roy Chowdhury, A. J. Smith, K. Frenander, A. Mikheenkova, R. W. Lindström, T. Thiringer, Influence of state of charge window on the degradation of Tesla lithium-ion battery cells. Journal of Energy Storage. Vol. 76, pg. 110001, 2024., https://doi.org/10.1016/j.est.2023.110001.

2 J. G. Qu, Z. Y. Jiang, J. F. Zhang, Investigation on lithium-ion battery degradation induced by combined effect of current rate and operating temperature during fast charging. Journal of Energy Storage. Vol. 52, pg. 104811, 2022., https://doi.org/10.1016/j.est.2022.104811.

3 S. E. J. O'Kane, W. Ai, G. Madabattula, D. Alonso-Alvarez, R. Timms, V. Sulzer, J. S. Edge, B. Wu, G. J. Offer, M. Marinescu, Lithium-ion battery degradation: how to model it. Physical Chemistry Chemical Physics. Vol. 24, pg. 7909–7922, 2022., https://doi.org/10.1039/D2CP00417H.

4 Q. Wang, B. Mao, S. I. Stoliarov, J. Sun, A review of lithium ion battery failure mechanisms and fire prevention strategies. Progress in Energy and Combustion Science. Vol. 73, pg. 95–131, 2019., https://doi.org/10.1016/j.pecs.2019.03.002.

5 H. Liu, I. H. Naqvi, F. Li, C. Liu, N. Shafiei, Y. Li, M. Pecht, An analytical model for the CC-CV charge of Li-ion batteries with application to degradation analysis. Journal of Energy Storage. Vol. 29, pg. 101342, 2020., https://doi.org/10.1016/j.est.2020.101342.

6 Z. Wei, Z. Quan, J. Wu, Y. Li, J. Pou, H. Zhong, Deep deterministic policy gradient-DRL enabled multiphysics-constrained fast charging of lithium-ion battery. IEEE Transactions on Industrial Electronics. Vol. 69, pg. 2588–2598, 2022., https://doi.org/10.1109/TIE.2021.3070514.

7 N. Wassiliadis, J. Kriegler, K. A. Gamra, M. Lienkamp, 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. Journal of Power Sources. Vol. 561, pg. 232586, 2023., https://doi.org/10.1016/j.jpowsour.2022.232586.

8 Y. Lu, X. Han, Y. Li, X. Li, M. Ouyang, Health-aware fast charging for lithium-ion batteries: model predictive control, lithium plating detection, and lifelong parameter updates. IEEE Transactions on Industry Applications. Vol. 60, pg. 7389–7398, 2024., https://doi.org/10.1109/TIA.2024.3427049.

9 H. El Ouazzani, I. El Hassani, N. Barka, T. Masrour, MSCC-DRL: multistage constant current based on deep reinforcement learning for fast charging of lithium ion battery. Journal of Energy Storage. Vol. 75, pg. 109695, 2024., https://doi.org/10.1016/j.est.2023.109695.

10 M. Yuan, C. Zou, Lifelong reinforcement learning for health-aware fast charging of lithium-ion batteries. IEEE Transactions on Transportation Electrification. Vol. 12, pg. 1129–1140, 2026., https://doi.org/10.1109/TTE.2025.3625421.

11 J. He, T. Yang, L. Xie, Y. Yang, C. Chen, J. Wei, A data-driven reinforcement learning enabled battery fast charging optimization using real-world experimental data. IEEE Transactions on Industrial Electronics. Vol. 72, pg. 430–438, 2025., https://doi.org/10.1109/TIE.2024.3398687.

12 Y. Zhang, T. Wik, J. Bergström, C. Zou, Machine learning-based lifelong estimation of lithium plating potential: a path to health-aware fastest battery charging. Energy Storage Materials. Vol. 74, pg. 103877, 2025., https://doi.org/10.1016/j.ensm.2024.103877.

13 M. A. Chowdhury, S. S. S. Al-Wahaibi, Q. Lu, Adaptive safe reinforcement learning-enabled optimization of battery fast-charging protocols. AIChE Journal. Vol. 71, pg. e18605, 2025., https://doi.org/10.1002/aic.18605.

14 Z. Zhang, T. Guo, Y. Liu, X. Pang, Z. Zheng, Fast-charging optimization method for lithium-ion battery packs based on deep deterministic policy gradient algorithm. Batteries. Vol. 11, pg. 199, 2025., https://doi.org/10.3390/batteries11050199.

15 S. Park, A. Pozzi, M. Whitmeyer, H. Perez, A. Kandel, G. Kim, Y. Choi, W. T. Joe, D. M. Raimondo, S. Moura, A deep reinforcement learning framework for fast charging of Li-ion batteries. IEEE Transactions on Transportation Electrification. Vol. 8, pg. 2770–2784, 2022., https://doi.org/10.1109/TTE.2022.3140316.

16 V. Sulzer, S. G. Marquis, R. Timms, M. Robinson, S. J. Chapman, Python battery mathematical modelling (PyBaMM). Journal of Open Research Software. Vol. 9, pg. 14, 2021., https://doi.org/10.5334/jors.309.

17 S. G. Marquis, V. Sulzer, R. Timms, C. P. Please, S. J. Chapman, An asymptotic derivation of a single particle model with electrolyte. Journal of The Electrochemical Society. Vol. 166, pg. A3693–A3706, 2019., https://doi.org/10.1149/2.0341915jes.

18 D.-I. Stroe, M. Swierczynski, A.-I. Stroe, S. K. Kær, Generalized characterization methodology for performance modelling of lithium-ion batteries. Batteries. Vol. 2, pg. 37, 2016., https://doi.org/10.3390/batteries2040037.

19 Scikit-learn developers, Ridge. Scikit-learn API documentation, accessed July 24, 2026., https://scikit-learn.org/stable/module

© The National High School Journal of Science 2026

s/generated/sklearn.linear_model.Ridge.html.

20 Scikit-learn developers, StandardScaler. Scikit-learn API documentation, accessed July 24, 2026., https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html.

21 V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning. Nature. Vol. 518, pg. 529–533, 2015., https://doi.org/10.1038/nature14236.

22 V. Nair, G. E. Hinton, Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th International Conference on Machine Learning. pg. 807–814, 2010., https://icml.cc/2010/papers/432.pdf.

23 S. Ohneseit, M. C. Holocher, A. Kalk, N. Uhlmann, H. J. Seifert, C. Ziebert, Aging behavior beyond SOH 80: an experimental aging study on commercial lithium-ion batteries with different cathode materials: capacity loss, resistance change and impedance modeling. Batteries & Supercaps. Vol. 8, pg. e202400713, 2025., https://doi.org/10.1002/batt.202400713.

24 N. Lin, Z. Jia, Z. Wang, H. Zhao, G. Ai, X. Song, Y. Bai, V. S. Battaglia, C. Sun, J. Qiao, K. Wu, G. Liu, Understanding the crack formation of graphite particles in cycled commercial lithium-ion batteries by focused ion beam-scanning electron microscopy. Journal of Power Sources. Vol. 365, pg. 235–239, 2017., https://doi.org/10.1016/j.jpowsour.2017.08.045.

25 J. Vetter, P. Novak, M. R. Wagner, C. Veit, K.-C. Möller, J. O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, A. Hammouche, Ageing mechanisms in lithium-ion batteries. Journal of Power Sources. Vol. 147, pg. 269–281, 2005., https://doi.org/10.1016/j.jpowsour.2005.01.006.

26 E. Redondo-Iglesias, P. Venet, S. Pelissier, Modelling lithium-ion battery ageing in electric vehicle applications: calendar and cycling ageing combination effects. Batteries. Vol. 6, pg. 14, 2020., https://doi.org/10.3390/batteries6010014.

27 Y. Xie, S. Shi, J. Tang, H. Wu, J. Yu, Experimental and analytical study on heat generation characteristics of a lithium-ion power battery. International Journal of Heat and Mass Transfer. Vol. 122, pg. 884–894, 2018., https://doi.org/10.1016/j.ijheatmasstransfer.2018.02.038.

28 P. Keil, A. Jossen, Charging protocols for lithium-ion batteries and their impact on cycle life: an experimental study with different 18650 high-power cells. Journal of Energy Storage. Vol. 6, pg. 125–141, 2016., https://doi.org/10.1016/j.est.2016.02.005.

29 G. Kučinskis, M. Bozorgchenani, M. Feinauer, M. Kasper, M. Wohlfahrt-Mehrens, T. Waldmann, Arrhenius plots for Li-ion battery ageing as a function of temperature, C-rate, and ageing state: an experimental study. Journal of Power Sources. Vol. 549, pg. 232129, 2022., https://doi.org/10.1016/j.jpowsour.2022.232129.

30 X. Lin, Y. Kim, S. Mohan, J. B. Siegel, A. G. Stefanopoulou, Modeling and estimation for advanced battery management. Annual Review of Control, Robotics, and Autonomous Systems. Vol. 2, pg. 393–426, 2019., https://doi.org/10.1146/annurev-control-053018-023643.

31 M. Baumann, L. Wildfeuer, S. Rohr, M. Lienkamp, Parameter variations within Li-ion battery packs: theoretical investigations and experimental quantification. Journal of Energy Storage. Vol. 18, pg. 295–307, 2018., https://doi.org/10.1016/j.est.2018.04.031.