Thermal Management of Liquid-Cooled 21700 Lithium-Ion Batteries: A Comparative Evaluation of LQR and Reinforcement Learning Control

Kelfin Munene Njagi, Benjamin Aidoo, Robert Sowah

Department of Engineering, Ashesi University, Berekuso, Eastern Region, Ghana

ABSTRACT

Effective thermal management is essential for lithium-ion battery safety and performance. This study develops a control-oriented five-node lumped thermal model of a liquid-cooled cylindrical 21700 cell incorporating Bernardi heat generation, state-of-charge dynamics, and temperature-dependent internal resistance. The model compares Linear Quadratic Regulator (LQR) and Reinforcement Learning (RL) controllers under identical conditions. Simulation results show that LQR maintains the cathode temperature closer to the operating point and reduces cooling energy by 39.13% compared with continuous pump operation. At an initial temperature of 45 °C, LQR requires less cooling energy than RL.

KEYWORDS

Battery thermal management, 21700 lithium-ion battery, liquid cooling, Linear Quadratic Regulator, reinforcement learning, optimal control

I. INTRODUCTION

Lithium-ion batteries are central to electric vehicles because of their high energy density and reliable cycling performance. During operation, heat is generated through resistive losses and electrochemical reactions, with heat generation increasing under higher discharge rates. If not adequately removed, this heat can raise cell temperature beyond the recommended operating range, accelerating ageing, reducing power capability, and increasing the risk of thermal runaway(Feng et al. 2018). Maintaining battery temperature within an appropriate range is therefore important for safe and reliable operation.

Several approaches have been developed for battery thermal management, including air cooling, phase-change materials, heat pipes, and liquid cooling. Air cooling is simple but becomes less effective at high discharge currents and in compact battery configurations where airflow is restricted (Kim et al. 2019). Phase-change materials and heat pipes can improve heat spreading but may introduce additional weight or cost. Liquid cooling provides higher heat transfer capability and more uniform temperature control, making it attractive for electric vehicle battery systems (Kim et al. 2019). Its effectiveness, however, depends on how coolant flow and pump operation are controlled.

Thermal modelling is important for predicting battery temperature evolution and designing thermal management systems. Lumped thermal models are computationally efficient but generally represent the battery using a single temperature state and therefore cannot capture internal temperature gradients. Multi-node models divide the cell into several thermal regions connected by equivalent thermal resistances and capacitances, providing a compromise between model complexity and computational efficiency. Such models are particularly suitable for control-oriented analysis because they capture the dominant thermal dynamics without requiring a highly detailed spatial model.

Conventional battery thermal management systems often use fixed-speed pumps, temperature-threshold control, or proportional–integral–derivative controllers. Although these approaches are relatively simple, they may provide excessive cooling or become less effective under changing operating conditions and dynamic heat generation (Wang et al. 2024). Model-based control methods incorporate the thermal dynamics directly into the control design and can therefore regulate cooling effort more systematically (Zhang et al. 2022). In particular, the Linear Quadratic Regulator (LQR) provides a means of balancing temperature regulation against cooling energy consumption, while Reinforcement Learning (RL) offers a data-driven approach for optimizing battery thermal-management decisions.

Despite advances in battery thermal modelling and control, limited work has directly compared optimal state-space control and reinforcement learning within the same liquid-cooled battery model and operating conditions. This study therefore develops a five-node control-oriented thermal model of a cylindrical 21700 lithium-ion battery and evaluates an LQR controller and an RL controller using the same physical model and operating conditions. Their performance is compared in terms of temperature regulation, pump behaviour, and cooling energy consumption under normal and elevated initial-temperature conditions. This common framework allows differences in performance to be assessed primarily in terms of the control strategy.

II. METHODS

A. Thermal Model Development

The thermal behaviour of the lithium-ion battery cell is represented by a lumped-parameter thermal network that captures the dominant heat transfer processes within the cylindrical cell. In this approach, the battery is divided into several thermal regions connected by equivalent thermal resistances and capacitances. This representation provides a control-oriented model that describes the battery's dynamic temperature behaviour while maintaining computational simplicity.

In this study, the cylindrical 21700 lithium-ion cell is modelled using a five-node thermal network. The model represents the cathode-side surface, cathode region, electrolyte layer, anode region, and anode-side surface. Heat transfer between adjacent regions is modelled through equivalent conductive thermal resistances, while each region is assigned a thermal capacitance representing its heat storage capability. Heat generated within the cell propagates through the thermal network before being removed by the liquid cooling system.

This multi-node thermal representation captures internal temperature gradients within the battery while maintaining a relatively simple structure suitable for control design and simulation analysis.

B. Heat Generation Model

The internal heat generated within the lithium-ion cell is described using the Bernardi formulation, which combines irreversible ohmic losses and reversible electrochemical heat effects. This expression accounts for the dominant mechanisms responsible for temperature rise during battery operation. The total heat generation rate inside the cell is expressed as

$$S(t) = I^2 R_{tot}(T, SOC) + I\, T_c \frac{dE}{dT}$$

where I is the applied current, $R_{tot}(T, SOC)$ is the internal resistance as a function of temperature and state of charge, T_c is the cathode temperature, and $\frac{dE}{dT}$ is the temperature coefficient of the open-circuit voltage.

In this work, the entropy coefficient is assumed constant and follows the Bernardi formulation

$$\frac{dE}{dT} = -10^{-4} \text{ V/K}$$

Which indicates that the open-circuit voltage decreases slightly as the cell temperature increases. The heat-generation term obtained from this expression is applied to the thermal model as an internal heat source driving temperature evolution in the five-node thermal network.

C. State of Charge Dynamics

The state of charge (SOC) is included in the model to represent the remaining usable capacity of the lithium-ion cell during discharge. The SOC state evolves according to a Coulomb counting relationship that links the applied current to the change in stored charge.

The SOC dynamics are expressed as

$$S\dot{O}C = -\frac{I}{C_{nom}}$$

Where I is the applied current and C_{nom} is the nominal capacity of the battery cell. The negative sign indicates that the state of charge decreases during discharge.

Although SOC does not directly affect the temperature states, it influences the battery's internal resistance through the resistance model used in the heat generation calculation. As a result, SOC indirectly influences thermal behaviour by altering the amount of heat generated within the cell.

D. Dynamic Resistance Model

The internal resistance of the lithium-ion cell varies with both temperature and state of charge. To capture this behaviour, a temperature-dependent resistance model is introduced. The formulation combines a state-of-charge correction factor with an Arrhenius-type temperature dependence.

The resistance is expressed as

$$R(T,SOC) = R_0 \, (1 + 0.4(1 - SOC))\exp[\frac{E_a}{R_g}(\frac{1}{T_k} - \frac{1}{T_{ref}})]$$

Where $R_0$is the reference resistance, $SOC$is the state of charge, $E_a$is the activation energy, and $R_g$is the universal gas constant. The temperatures $T_K$and T_{ref}represent the operating temperature and reference temperature in Kelvin, respectively. The term $1 + 0.4(1 - SOC)$accounts for the increase in resistance at lower states of charge, which is commonly observed in lithium-ion cells. The exponential term describes the dependence of resistance on temperature using an Arrhenius relationship, reflecting the decrease in internal resistance as temperature increases.

E. Five-Node Thermal Model

The five-node thermal network adopted in this study is based on the multi-node thermal modelling approach of Ferreira and Tang (2024), with the formulation adapted to the geometry, materials, and liquid-cooling configuration of the cylindrical 21700 lithium-ion cell considered here.

The cylindrical lithium-ion cell is represented by a lumped multi-node thermal network that captures radial heat transfer through the battery's internal layers. The model divides the cell into five thermal regions: the cathode-side surface, the cathode block, the electrolyte layer, the anode block, and the anode-side surface. Each region is represented by a temperature state and is connected to neighbouring regions through equivalent thermal resistances.

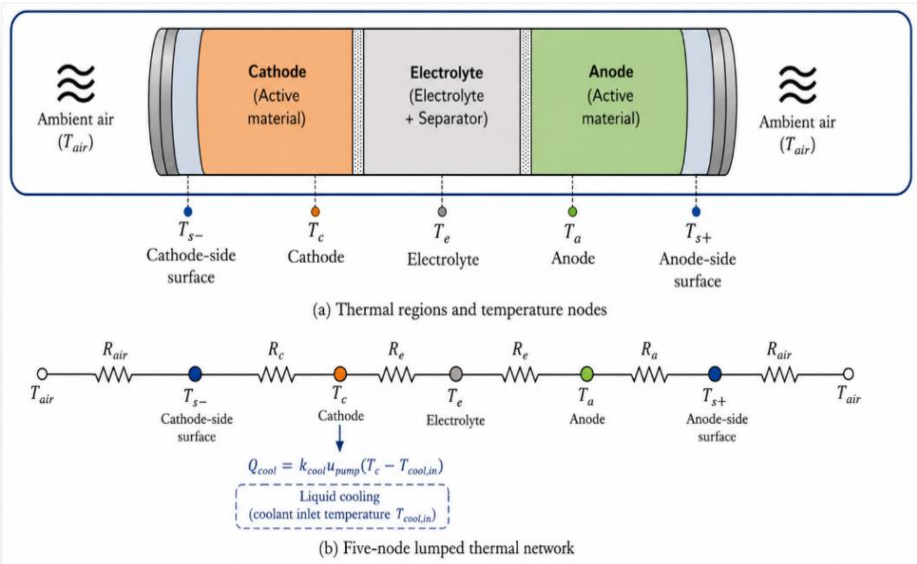

Figure 1: Five-node thermal network of the cylindrical 21700 lithium-ion battery, adapted from (Ferreira and Tang 2024).

Heat generated inside the cell propagates through the thermal network by conduction between adjacent regions and is partially removed through convection at the outer surfaces. The resulting temperature dynamics are obtained by applying an energy balance to each thermal node, accounting for conductive heat transfer, internal heat generation, and cooling effects.

The temperature state vector is defined as

$$\mathbf{x} = [T_{s^-} \quad T_c \quad T_e \quad T_a \quad T_{s^+}]^T,$$

Where T_{s^-}, T_c, T_e, T_a, and T_{s^+} denote the temperatures of the cathode-side surface, cathode region, electrolyte layer, anode region, and anode-side surface, respectively.

Applying energy conservation to each thermal node yields the governing equations of the five-node thermal model.

The cathode-side surface temperature is described by

$$\dot{T}_{s^-} = \frac{T_{air} - T_{s^-}}{\lambda_{air} R_{air}} - \frac{T_{s^-} - T_c}{\lambda_{air} R_c},$$

where T_{air} is the ambient temperature, λ_{air} is the effective thermal capacitance of the surface node, R_{air} is the convective thermal resistance between the cell surface and the surrounding air, and R_c is the conduction resistance between the cathode-side surface and the cathode region.

The cathode temperature evolves according to

$$\dot{T}_c = \frac{T_{s^-} - T_c}{\lambda_c R_c} + \frac{S(t)}{\lambda_c} - \frac{Q_{cool}}{\lambda_c},$$

Where λ_c the thermal capacitance of the cathode region is, $S(t)$ is the internal heat generation rate, and Q_{cool} represents the heat removed by the liquid-cooling system,

$$Q_{cool} = k_{cool} u_{pump} (T_c - T_{cool,in}),$$

With k_{cool} denoting the cooling coefficient, u_{pump} the normalized pump command, and $T_{cool,in}$ the coolant inlet temperature.

The electrolyte temperature is expressed as

$$\dot{T}_e = \frac{T_c - T_e}{\lambda_e R_e} + \frac{S(t)}{\lambda_e},$$

Where λ_e is the thermal capacitance of the electrolyte region and R_e is the equivalent thermal resistance between adjacent internal layers.

Similarly, the anode temperature is given by

$$\dot{T}_a = \frac{T_e - T_a}{\lambda_a R_e} + \frac{S(t)}{\lambda_a},$$

Where λ_a denotes the thermal capacitance of the anode region.

Finally, the temperature of the anode-side surface is governed by

$$\dot{T}_{s^+} = \frac{T_a - T_{s^+}}{\lambda_{air} R_a} - \frac{T_{s^+} - T_{air}}{\lambda_{air} R_{air}},$$

Where $R_a$is the thermal resistance between the anode region and the outer surface?

Together, these coupled differential equations describe radial heat conduction, internal heat generation, convective heat exchange with the surrounding environment, and heat removal through liquid cooling. The resulting nonlinear thermal model captures the dominant temperature dynamics of the 21700 lithium-ion cell while remaining computationally efficient for controller design and simulation.

F. State-Space Representation

The governing equations developed in the previous section can be combined into a nonlinear state-space model suitable for controller design. The temperature states evolve according to

$$\dot{\mathrm{x}} = f(\mathrm{x}, \mathrm{u}),$$

Where

$$\mathrm{x} = [T_{s^-} \quad T_c \quad T_e \quad T_a \quad T_{s^+}]^T$$

Is the thermal state vector and

$$\mathrm{u} = \begin{bmatrix} I \\ u_{pump} \end{bmatrix}$$

Is the input vector comprising the battery discharge current, I and the coolant pump command, u_{pump}.

The nonlinear function $f(\cdot)$incorporates conductive heat transfer between adjacent thermal regions, convective heat exchange with the surrounding environment, internal heat generation described by the Bernardi formulation, and the cooling effect produced by the liquid-cooling system. Since the heat generation and internal resistance depend on battery operating conditions, the resulting model exhibits nonlinear behaviour. This nonlinear state-space representation provides the basis for both control strategies investigated in this study. For the Linear Quadratic Regulator, the model is linearized about a representative operating point to obtain a linear state-space approximation. In contrast, the Reinforcement Learning controller interacts directly with the nonlinear model during training and evaluation, allowing both approaches to be assessed using the same physical system.

G. LQR Controller Design

The Linear Quadratic Regulator (LQR) was selected as the benchmark optimal control strategy because it provides systematic regulation of battery temperature while minimizing the control effort required from the cooling system. Unlike conventional rule-based or fixed-speed cooling methods, LQR determines the pump command by considering the complete thermal state of the battery and optimally balancing temperature regulation against auxiliary energy consumption.

To design the controller, the nonlinear thermal model was linearized about a representative operating point corresponding to an ambient temperature of 25°C, a 1C discharge rate, a state of charge of 0.7, and zero pump command. Around this operating condition, the nonlinear model can be approximated by the linear state-space system

$$\dot{x} = Ax + Bu,$$

Where $A$and $B$are the state and input matrices obtained through first-order linearization. The LQR controller determines the optimal control input by minimizing the quadratic performance index

$$J = \int_0^{\infty} (x^T Q x + u^T R u)\ dt,$$

Where $Q$is the state weighting matrix that penalizes temperature deviations from the desired operating condition, while Rpenalizes excessive pump actuation. Appropriate selection of these weighting matrices enables a trade-off between thermal regulation accuracy and cooling energy consumption.

The resulting optimal control law is expressed as

$$u = -Kx,$$

Where $K$is the optimal feedback gain obtained by solving the continuous-time algebraic Riccati equation? The controller continuously adjusts the coolant pump command according to the instantaneous thermal state of the battery, thereby reducing unnecessary cooling while maintaining safe operating temperatures.

H. Reinforcement Learning Controller

To investigate the performance of a learning-based control strategy, a Reinforcement Learning (RL) controller was developed using the nonlinear battery thermal model. Unlike the LQR controller, which relies on a linearized mathematical model, the RL agent learns an optimal cooling policy through repeated interaction with the battery system. At each time step, the agent observes the thermal state of the battery, selects a coolant pump command, and receives a reward based on the resulting temperature regulation and cooling energy consumption.

The learning process was formulated as a sequential decision-making problem in which the objective was to maintain the battery temperature within the desired operating range while minimizing auxiliary cooling effort. The state vector supplied to the agent consisted of the five thermal states of the battery,

$$x = [T_{s^-} \quad T_c \quad T_e \quad T_a \quad T_{s^+}]^T,$$

While the control action corresponded to the normalized pump command,

$0 \leq u_{pump} \leq 1.$

This formulation enables the controller to continuously regulate coolant flow according to the instantaneous thermal condition of the battery.

The reinforcement learning agent was trained using the nonlinear thermal model developed in Section 3. During training, the controller repeatedly interacted with the simulated battery environment by observing the current thermal state, selecting a pump command, and receiving feedback through a reward function. This iterative process enabled the agent to progressively improve its control policy without requiring an explicit mathematical solution to the optimization problem.

The reward function was designed to encourage accurate temperature regulation while discouraging excessive cooling effort. Temperature deviations from the desired operating region were penalized together with unnecessary pump actuation, allowing the controller to balance thermal safety against auxiliary energy consumption. Through repeated training episodes, the agent gradually learned a control policy that maintained battery temperature within safe limits while adapting to changing operating conditions.

Unlike the LQR controller, which is derived from a linear approximation of the system dynamics, the reinforcement learning controller operates directly on the nonlinear battery model. This enables the agent to respond to nonlinear thermal behaviour without requiring model linearization, thereby providing a useful benchmark for evaluating the effectiveness of data-driven control in battery thermal management.

I. Simulation Setup

The proposed control strategies were evaluated using MATLAB/Simulink under identical operating conditions to ensure a fair comparison between classical optimal control and reinforcement learning. Both controllers employed the same five-node thermal model, identical battery parameters, and the same heat-generation model based on the Bernardi formulation.

The simulations considered a cylindrical 21700 lithium-ion cell subjected to a constant 1C discharge under an ambient temperature of 25°C. Two operating scenarios were investigated. The first considered a battery initially at ambient temperature to evaluate the controllers under normal operating conditions. The second considered an elevated initial temperature of 45°C to assess controller performance during severe thermal loading.

For each simulation, the controllers were evaluated using the following performance criteria:

- Maximum battery temperature.
- Temperature regulation capability.
- Pump command behaviour.
- Cooling energy consumption.
- Overall thermal stability.

To establish a performance baseline, the proposed controllers were compared with continuous pump operation, where the coolant pump remained fully active throughout the simulation. This benchmark represents a commonly adopted cooling strategy and provides a reference for assessing improvements in thermal regulation and auxiliary energy efficiency. The cooling energy consumed during each simulation was obtained by integrating the instantaneous cooling power over the entire operating period. This metric provides a quantitative measure of the energy required by each control strategy to maintain acceptable battery temperatures.

III. RESULTS AND DISCUSSION

This section presents the simulation results obtained from the proposed battery thermal management framework. The performance of the developed controllers is evaluated using the nonlinear five-node thermal model under identical operating conditions. The analysis focuses on the ability of each control strategy to regulate battery temperature while minimizing cooling energy consumption. Particular attention is given to the temperature response of the battery, coolant pump behaviour, and the overall energy required for thermal management. The results are used to compare the performance of the Linear Quadratic Regulator (LQR) and Reinforcement Learning (RL) controllers under both ambient and elevated initial temperature conditions.

A. Performance of the LQR Controller

The effectiveness of the proposed Linear Quadratic Regulator (LQR) was first evaluated by comparing its thermal response with two conventional cooling strategies: operation without active cooling (Pump OFF) and continuous cooling (Pump ON). These operating conditions provide lower and upper performance bounds for assessing the controller's ability to regulate battery temperature while minimizing unnecessary cooling effort.

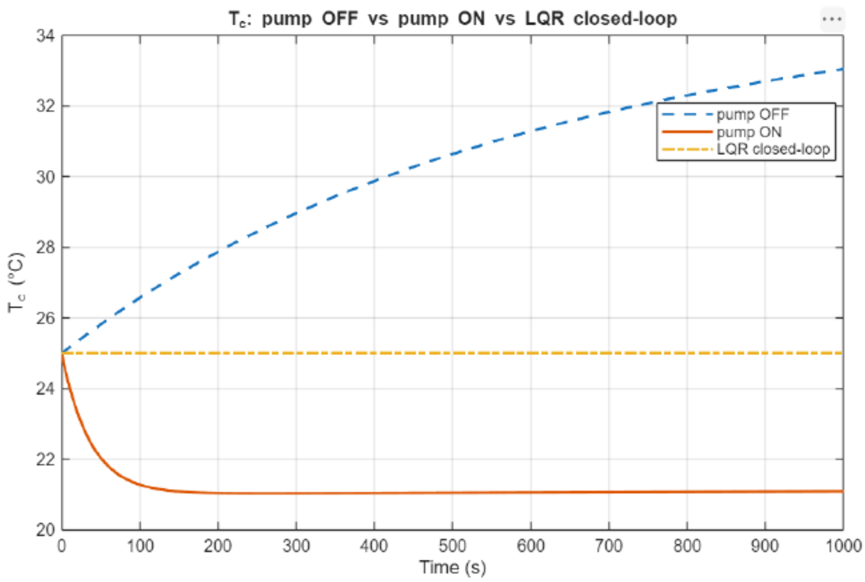

Figure 2: Comparison of cathode temperature under pump-off, pump-on, and LQR closed-loop control

Figure 2 compares the cathode temperature response obtained under the three cooling strategies throughout the discharge process. In the absence of active cooling (Pump OFF), the battery temperature increased continuously as internally generated heat accumulated within the cell. Although the temperature remained within the simulated operating limits, the absence of forced cooling resulted in the highest thermal excursion among the three operating conditions. Continuous pump operation (Pump ON) effectively suppressed the temperature rise by providing uninterrupted heat removal; however, this strategy maintained maximum cooling throughout the discharge period irrespective of the instantaneous thermal demand.

The proposed LQR controller achieved a temperature response that closely followed the continuous cooling case while requiring substantially less cooling effort. The difference between the Pump ON and LQR temperature trajectories remained small throughout the simulation, indicating that the controller successfully maintained the battery within the desired operating region without relying on continuous pump operation. This behaviour demonstrates that the LQR controller effectively balances thermal regulation with actuator utilisation by supplying cooling only when required.

Rather than reacting to temperature thresholds, the state-feedback controller continuously adjusted the coolant pump command according to the predicted thermal state of the battery. Consequently, unnecessary cooling during periods of low thermal demand was avoided while adequate cooling capacity remained available whenever the battery temperature began to increase. The smooth temperature profile obtained under LQR control further indicates that the controller maintained stable closed-loop operation without introducing oscillatory behaviour or abrupt changes in the thermal response.

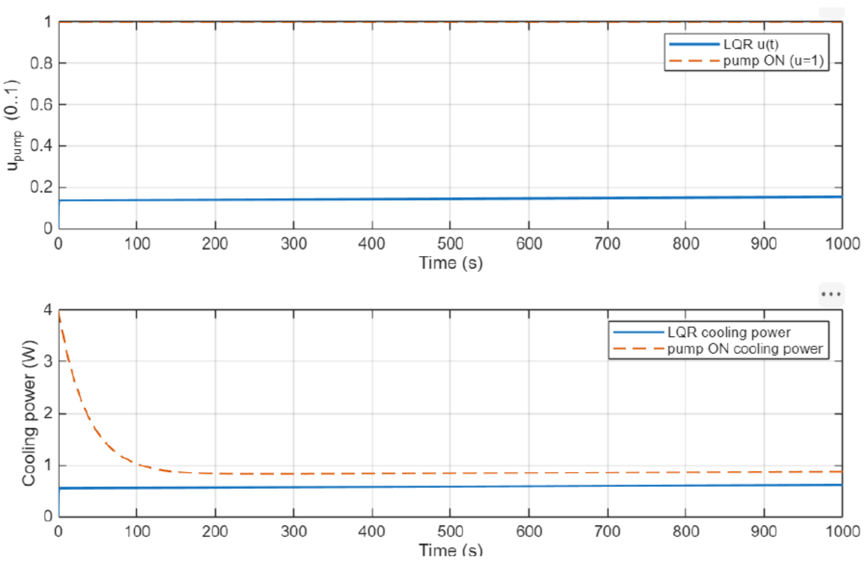

Figure 3: LQR pump command and corresponding cooling power compared with continuous pump operation

The control action generated by the LQR controller and the corresponding cooling power are presented in Figure 3. Unlike the continuous cooling strategy, which maintained maximum pump operation throughout the simulation, the LQR controller adjusted the pump command dynamically in response to the evolving thermal conditions. Higher pump activity was observed only during periods of increased heat generation, while reduced cooling effort was applied whenever the battery temperature approached the desired operating region.

This adaptive control behaviour translated directly into lower auxiliary energy consumption. The cumulative cooling energy required under continuous pump operation was 961.04 J, whereas the proposed LQR controller required only 584.97 J, representing a 39.13% reduction in cooling energy while maintaining comparable thermal regulation. These results demonstrate that optimal state-space control can significantly improve the energy efficiency of liquid-cooled battery thermal management without compromising temperature stability.

The reduction in cooling energy highlights one of the principal advantages of optimal feedback control for battery thermal management. By exploiting the thermal dynamics captured by the five-node model, the LQR controller minimised unnecessary actuator usage while maintaining effective heat removal. This improvement is particularly important in electric vehicle applications, where auxiliary power consumption directly influences the overall energy efficiency and driving range of the vehicle.

B. Performance of the Reinforcement Learning Controller

The Reinforcement Learning (RL) controller was evaluated using the same operating conditions adopted for the LQR controller to enable a direct comparison of the two control strategies. The controller interacted continuously with the nonlinear battery model, adjusting the coolant pump command according to the instantaneous thermal state of the battery. Unlike the LQR controller, which relies on a fixed state-feedback gain derived from a linearized model, the RL controller learned its control policy through repeated interaction with the nonlinear thermal environment.

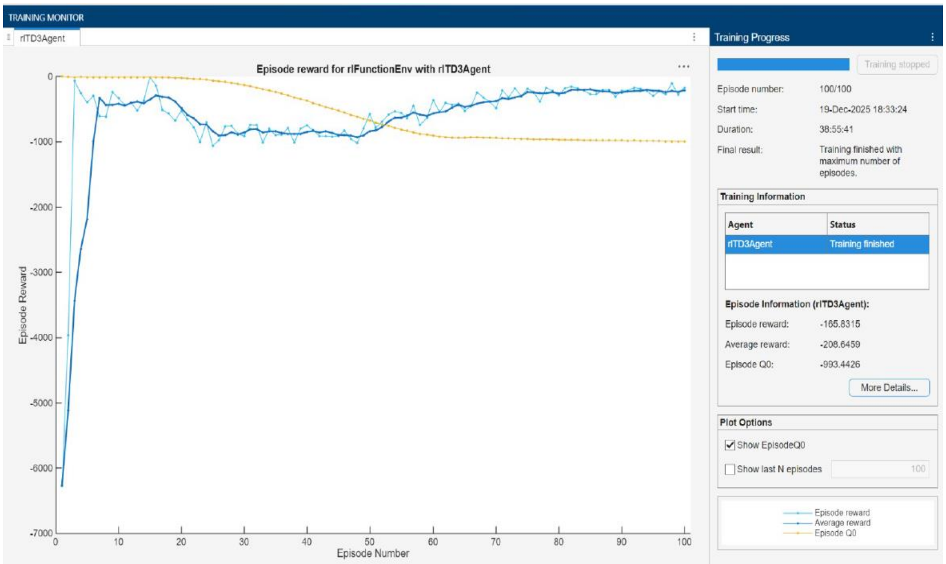

Figure 4: Training progress of the reinforcement learning agent

Figure 4 shows the evolution of the reinforcement learning process during training. At the beginning of training, the agent explored different control actions while gradually learning the relationship between battery temperature, pump operation, and the reward function. As training progressed, the cumulative reward increased and eventually approached a steady value, indicating that the controller had converted to a stable policy.

This convergence suggests that the agent consistently selected cooling actions that maintained acceptable battery temperatures while accounting for the energy required to operate the cooling system. The convergence behaviour demonstrates that the nonlinear battery model provided a stable learning environment for controller training. Once the reward stabilised, the learned policy was retained and used for all subsequent closed-loop simulations presented in this

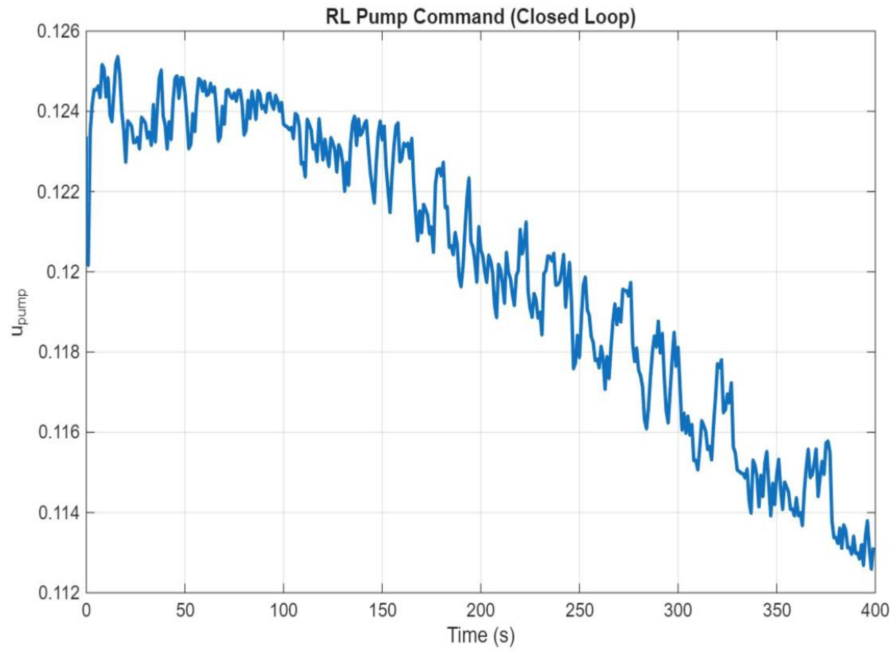

Figure 5: Closed-loop pump command generated by the reinforcement learning controller

Figure 5 presents the pump command generated by the trained reinforcement learning controller during battery discharge. Unlike a fixed-speed cooling strategy, the controller continuously adjusted the pump command according to the thermal condition of the battery. The pump activity increased during periods of higher thermal demand and decreased when less cooling was required, indicating that the controller adapted its actions to the changing operating conditions.

The variation in pump command reflects the policy learned during training. Rather than maintaining constant cooling, the controller determined the level of cooling required at each instant based on the observed thermal state. This adaptive behaviour allowed the controller to respond directly to the nonlinear temperature dynamics represented by the battery model.

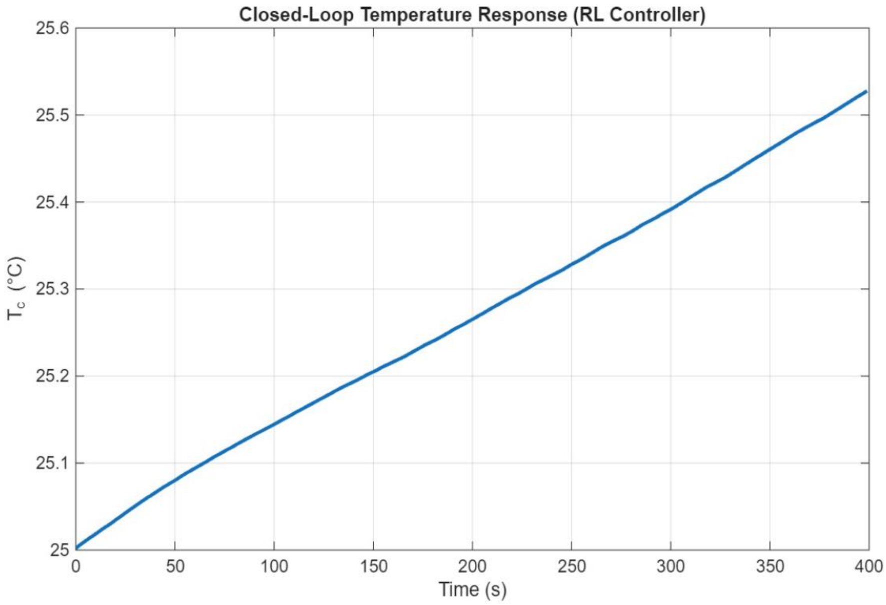

Figure 6: Closed-loop cathode temperature response under reinforcement learning control

Figure 6 shows the cathode temperature response obtained under reinforcement learning control during the discharge period. The temperature increased gradually from approximately 25.0 °C at the start of the simulation to about 25.5 °C after 400 s. This increase is consistent with continuous internal heat generation during battery discharge, where the rate of heat generation exceeds the rate of heat removed by the cooling system.

Despite the gradual increase in temperature, the response remained continuous throughout the simulation without abrupt changes. This behaviour indicates that the reinforcement learning controller maintained a consistent cooling action while the battery was subjected to a sustained thermal load. The relatively small increase in cathode temperature suggests that the controller was able to limit temperature rise over the simulation period.

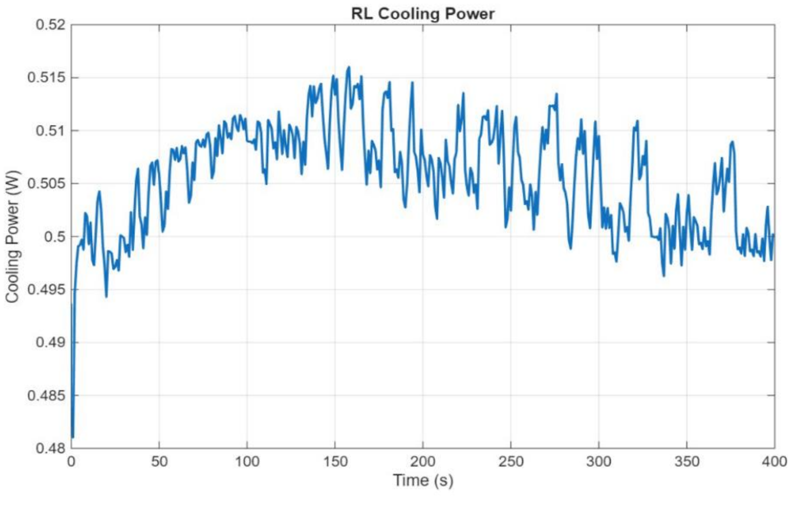

Figure 7: Cooling power under reinforcement learning control

Figure 7 shows the cooling power produced by the reinforcement learning controller during the discharge process. The cooling power increased rapidly at the beginning of the simulation before fluctuating within a relatively narrow range of approximately 0.50–0.515 W. Unlike a fixed cooling strategy, the controller continuously adjusted the cooling power in response to the changing thermal conditions of the battery. These variations indicate that the controller did not rely on a constant cooling level but instead adapted the cooling demand throughout the discharge period.

The fluctuations observed after the initial transient are characteristic of a reinforcement learning controller operating under continuous feedback. As the battery temperature evolved, the controller repeatedly updated its cooling action to satisfy the control objective while responding to the nonlinear dynamics of the thermal model. Despite these adjustments, the cooling power remained bounded within a relatively small operating range, suggesting that the controller avoided excessive changes in actuator demand.

The adaptive cooling behaviour observed in Figure 7 complements the temperature response presented in Figure 6. While the battery temperature increased gradually during discharge because of continuous heat generation, the controller continuously modified the cooling effort to limit the rate of temperature rise. This behaviour illustrates how the learned control policy balanced heat removal with actuator utilisation under the simulated operating conditions.

C. Comparative Analysis of the LQR and Reinforcement Learning Controllers

The performance of the LQR and reinforcement learning controllers was compared under identical operating conditions to assess their ability to regulate battery temperature and manage cooling demand. Both controllers were evaluated using the same battery model, discharge profile, and initial ambient temperature, ensuring that any observed differences resulted from the control strategy rather than the simulation environment.

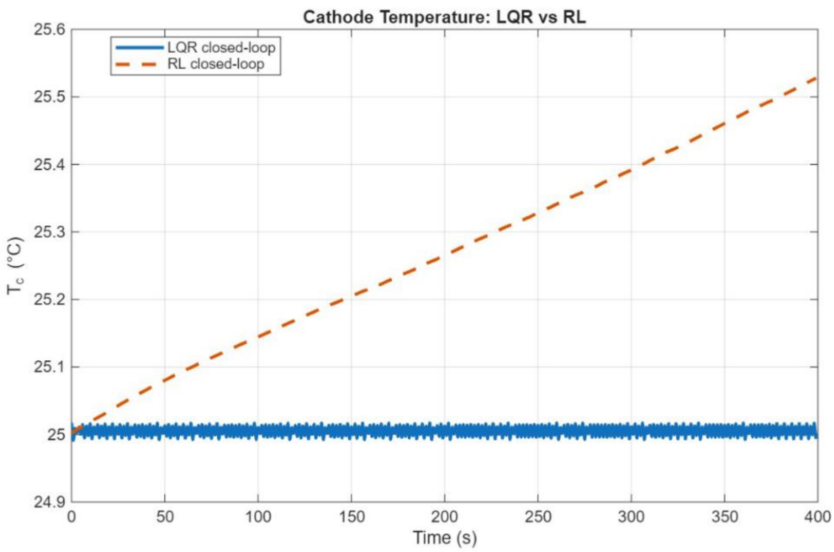

Figure 8: Comparison of cathode temperature response under LQR and reinforcement learning control.

Figure 8 compares the cathode temperature responses obtained using the LQR and reinforcement learning controllers under identical operating conditions. Both controllers maintained the battery temperature close to the initial operating point during the early stages of discharge. However, as the simulation progressed, the two temperature responses diverged. The LQR controller maintained the cathode temperature close to 25 °C throughout the simulation, whereas the reinforcement learning controller exhibited a gradual increase in temperature, reaching approximately 25.53 °C at the end of the discharge period.

The difference between the two responses became more evident with increasing discharge time. While the reinforcement learning controller continued to allow a gradual rise in temperature, the LQR controller maintained a nearly constant thermal response over the entire simulation interval. Under the simulated operating conditions, these results indicate that the LQR controller provided tighter temperature regulation than the reinforcement learning controller.

The observed difference reflects the distinct control strategies adopted by the two approaches. The LQR controller computes the cooling action from the linearized state-space model using an optimal state-feedback law, enabling continuous correction of temperature deviations from the operating point. In contrast, the reinforcement learning controller regulates the cooling system through a learned policy obtained from interaction with the nonlinear thermal model. Although both controllers maintained stable closed-loop operation, the LQR controller achieved closer tracking of the desired operating temperature under the conditions considered in this study.

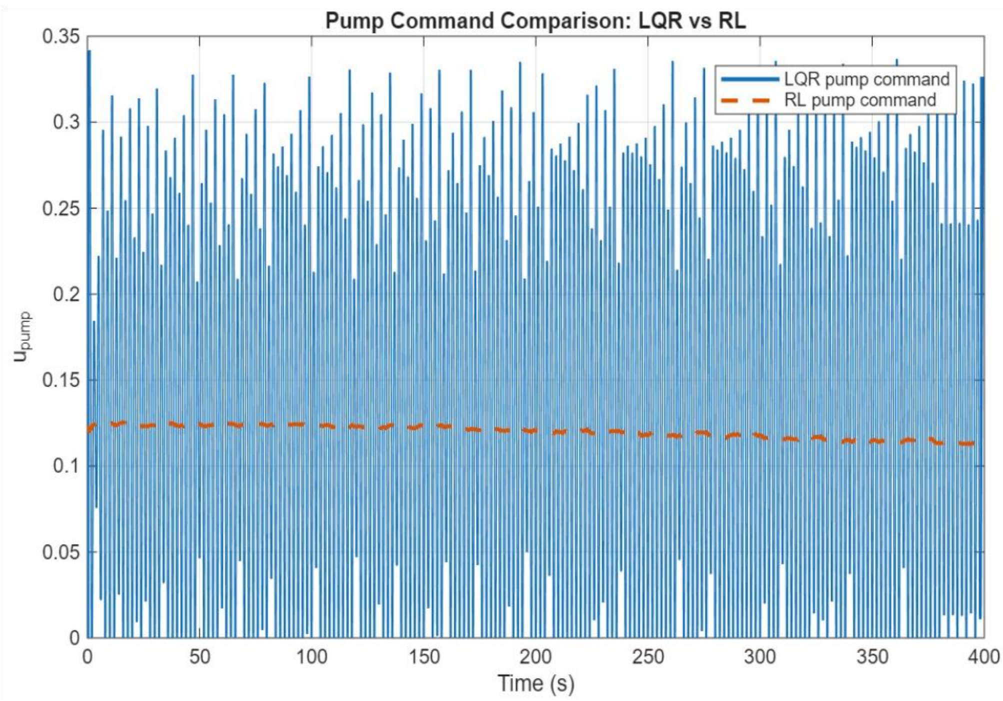

Figure 9: Pump command comparison under LQR and reinforcement learning control

Figure 9 compares the pump commands generated by the LQR and reinforcement learning controllers during battery discharge. The two controllers exhibit distinctly different control behaviours. The LQR controller continuously adjusted the pump command over a relatively wide operating range, producing frequent changes in the control signal throughout the simulation. In contrast, the reinforcement learning controller maintained a comparatively smoother pump command with only small variations around its operating level.

These differences reflect the manner in which the two controllers determine the cooling input. The LQR controller computes the pump command directly from the instantaneous state deviations, resulting in continuous corrections whenever the battery temperature changes. The reinforcement learning controller, on the other hand, follows the control policy acquired during training, leading to a more gradual adjustment of the cooling command during operation.

When considered together with the temperature responses shown in Figure 8, the results indicate that the two controllers adopt different approaches to thermal regulation. The LQR controller achieved tighter regulation of the cathode temperature through more active adjustment of the cooling input,

Whereas the reinforcement learning controller maintained a smoother control action but allowed a gradual increase in battery temperature over the discharge period. These observations highlight the trade-off between control aggressiveness and temperature regulation under the simulated operating conditions.

D. Performance at Elevated Initial Temperature (45 °C)

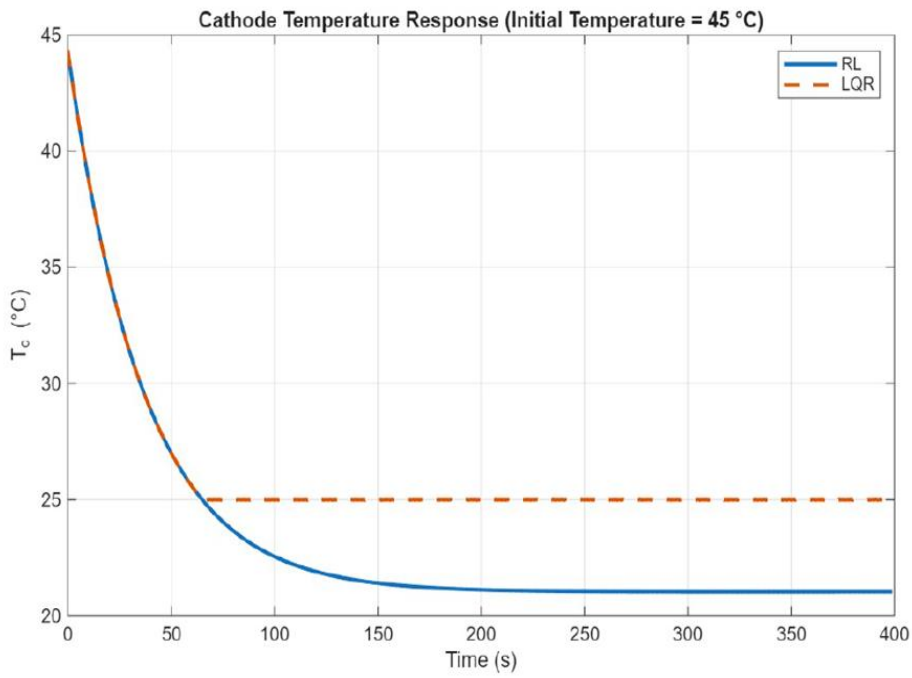

Figure 10: Cathode temperature response under reinforcement learning and LQR control at an initial temperature of 45 °C

Figure 10 compares the cathode temperature responses of the reinforcement learning and LQR controllers when the battery was initialized at 45 °C. Both controllers produced a rapid reduction in battery temperature during the initial stage of the simulation, reflecting the high cooling demand immediately after the start of operation. The two responses were nearly identical during the first 60–70 s, indicating that both controllers reacted effectively to the elevated thermal condition.

Beyond this period, the controller responses began to diverge. The reinforcement learning controller continued reducing the cathode temperature until it reached approximately 21 °C, whereas the LQR controller stabilized the temperature at approximately 25 °C. This difference indicates that the reinforcement learning controller maintained cooling for a longer duration, while the LQR controller regulated the battery temperature around the selected operating point.

These results demonstrate that both controllers were capable of reducing battery temperature from an elevated initial condition; however, they adopted different control strategies after the transient cooling phase. The reinforcement learning controller continued applying cooling beyond the nominal operating temperature, whereas the LQR controller maintained the temperature close to the reference value. This behaviour highlights the influence of the control policy on the resulting thermal response under high-temperature operating conditions.

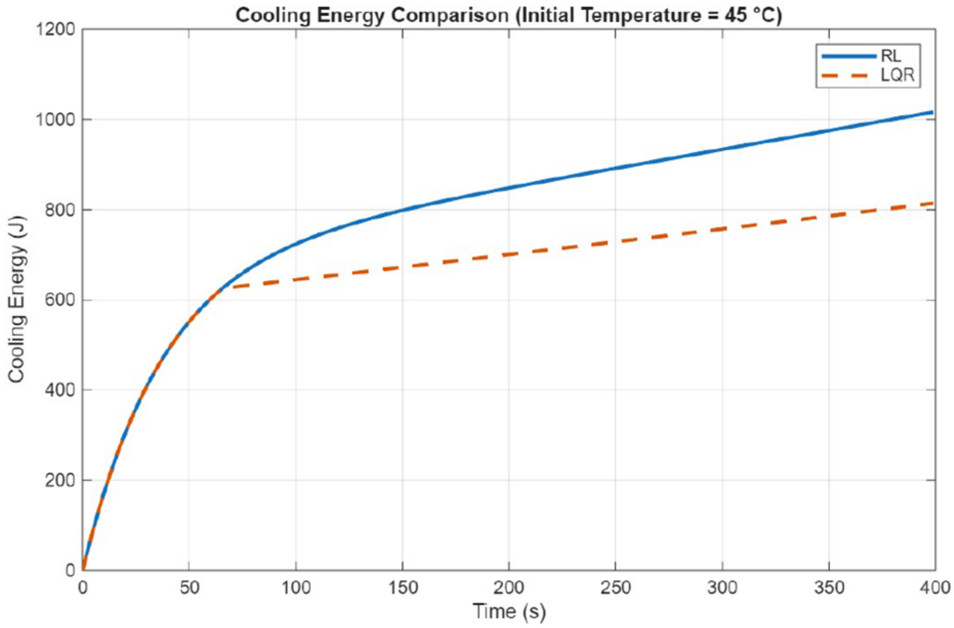

Figure 11: Comparison of cumulative cooling energy consumption under reinforcement learning and LQR control at an initial temperature of 45 °C.

Figure 11 compares the cumulative cooling energy consumed by the reinforcement learning and LQR controllers when the battery was initialized at 45 °C. During the initial cooling period, both controllers exhibited nearly identical energy consumption, reflecting the high cooling demand required to reduce the battery temperature from the elevated initial condition. As shown in Figure 10, both controllers rapidly decreased the cathode temperature during this stage, resulting in similar cooling energy requirements.

After approximately 70 s, the cumulative energy curves began to diverge. The reinforcement learning controller continued to accumulate cooling energy at a higher rate, reaching approximately 1015 J at the end of the simulation, whereas the LQR controller required approximately 810 J. This corresponds to a reduction in cooling energy of approximately 20% for the LQR controller under the elevated-temperature condition.

The difference in cumulative cooling energy is consistent with the temperature responses presented in Figure 10. After reducing the battery temperature from 45 °C, the reinforcement learning controller continued cooling beyond the nominal operating temperature, resulting in additional cooling energy consumption. In contrast, the LQR controller regulated the cathode temperature close to the desired operating point, avoiding unnecessary cooling while maintaining thermal regulation. These results demonstrate that the LQR controller achieved a more energy-efficient balance between temperature regulation and cooling effort under the simulated high-temperature operating condition.

The results obtained under both ambient and elevated initial temperatures indicate that the LQR controller consistently provided effective temperature regulation while requiring less cooling energy than the reinforcement learning controller. Although the reinforcement learning controller successfully reduced battery temperature under both operating conditions, its learned control policy applied additional cooling after the desired operating temperature had been reached, leading to higher cumulative energy consumption. Within the simulation framework adopted in this study, the LQR controller therefore demonstrated superior thermal regulation efficiency for the liquid-cooled 21700 lithium-ion cell.

This study was limited to a simulation-based evaluation of battery thermal management for a single liquid-cooled 21700 lithium-ion cell. Although the five-node lumped thermal model captured the dominant thermal behaviour required for controller development and performance assessment, it represents an idealized system and does not account for all physical phenomena encountered in practical battery applications.

The performance of the LQR and reinforcement learning controllers was evaluated under predefined operating conditions using fixed model parameters. Variations associated with battery ageing, manufacturing inconsistencies, parameter uncertainty, sensor measurement errors, actuator dynamics, and changing environmental conditions were not incorporated into the simulation framework. In addition, the study considered only a single-cell configuration and therefore did not investigate thermal interactions that occur in battery modules or full battery packs.

Furthermore, the proposed control strategies were validated exclusively through numerical simulations. Experimental implementation on a physical liquid-cooled battery system was beyond the scope of this study. Consequently, the reported controller performance should be interpreted within the assumptions and operating conditions adopted in the simulation model.

IV. CONCLUSION

This study investigated the performance of Linear Quadratic Regulator (LQR) and reinforcement learning (RL) controllers for battery thermal management of a liquid-cooled 21700 lithium-ion cell. Both controllers were implemented on the same five-node thermal model to enable a consistent evaluation of their temperature regulation and cooling energy requirements under identical operating conditions.

The simulation results demonstrated that both control strategies successfully regulated battery temperature during discharge. Under ambient initial conditions, the LQR controller maintained the cathode temperature close to the desired operating point, whereas the reinforcement learning controller allowed a gradual increase in temperature over the simulation period. When the battery was subjected to an elevated initial temperature of **45 °C**, both controllers rapidly reduced the battery temperature during the initial cooling stage. However, the reinforcement learning controller continued cooling beyond the desired operating temperature, while the LQR controller stabilized the battery temperature near the reference value. The differences in thermal regulation were reflected in the corresponding cooling energy requirements.

The reinforcement learning controller consumed more cumulative cooling energy under elevated temperature conditions because of its continued cooling action after the desired operating temperature had been reached. In contrast, the LQR controller achieved effective temperature regulation while requiring less cooling energy, demonstrating a better balance between thermal control performance and energy utilization within the simulation conditions considered in this study.

Overall, the results indicate that the LQR controller provided superior thermal regulation and cooling energy efficiency for the liquid-cooled 21700 lithium-ion cell evaluated in this work. The proposed modelling and control framework provides a useful basis for the design and assessment of advanced battery thermal management systems and establishes a benchmark for future investigations involving intelligent control strategies for electric vehicle battery applications. Future research should focus on the experimental implementation and validation of the proposed battery thermal management framework using a liquid-cooled battery test platform. Experimental studies would enable assessment of controller performance under practical operating conditions and facilitate validation of the simulation results obtained in this study.

The proposed five-node thermal model may also be extended to battery modules and full battery packs to investigate cell-to-cell thermal interactions, cooling uniformity, and the influence of different cooling configurations on overall thermal performance. Such studies would provide a more comprehensive evaluation of the proposed control strategies in electric vehicle battery systems.

Further improvements to the reinforcement learning controller may be achieved through the development of advanced reward functions that simultaneously optimize temperature regulation and cooling energy consumption. Future investigations may also explore deep reinforcement learning techniques capable of adapting to changing battery operating conditions, ageing effects, and parameter uncertainties while maintaining robust thermal control performance.

Finally, the proposed controllers may be evaluated under realistic electric vehicle operating scenarios by incorporating standard drive cycles, varying ambient temperatures, and dynamic load profiles. Such investigations would provide a broader understanding of controller performance under representative driving conditions and support the development of energy-efficient battery thermal management systems for next-generation electric vehicles.

V. DECLARATIONS

A. Availability of data and materials

The data generated and analysed during this study are available from the corresponding author on reasonable request. The simulation results reported in this article were generated using the modelling and simulation framework described in the manuscript.

B. Competing interests

The authors declare that they have no competing interests.

C. Funding

No external funding was received for the research, analysis, or preparation of this manuscript.

D. Authors' contributions

BA contributed to the conceptualisation of the study, provided technical guidance, and contributed to the interpretation and refinement of the research outcomes. RS provided research guidance throughout the study, contributed to the development and review of the methodology, and supported the interpretation of the results. KMN developed the battery thermal model and control strategies, implemented the LQR and reinforcement learning approaches, performed the simulations, analysed the results, and prepared the manuscript. All authors reviewed, revised, and approved the final manuscript.

E. Acknowledgements

The authors gratefully acknowledge Ashesi University for providing the academic environment and institutional support that facilitated the development of this research. Kelfin Munene Njagi sincerely appreciates Benjamin Aidoo and Robert Sowah for their valuable guidance, mentorship, and support throughout the development of this research and preparation of the manuscript.

REFERENCES

[1] Angani A, Kim H-W, Hwang M-H, et al (2023) A comparison between Zig-Zag plated hybrid parallel pipe and liquid cooling battery thermal management systems for Lithium-ion battery module. Appl Therm Eng 219:119599. https://doi.org/10.1016/j.applthermaleng.2022.119599.

[2] Bandhauer TM, Garimella S, Fuller TF (2011) A Critical Review of Thermal Issues in Lithium-Ion Batteries. J Electrochem Soc 158:R1. https://doi.org/10.1149/1.3515880.

[3] Bernardi D, Pawlikowski E, Newman J (1985) A General Energy Balance for Battery Systems. J Electrochem Soc 132:5–12.

[4] Cheng H, Jung S, Kim Y-B (2024) Battery thermal management system optimization using deep reinforced learning algorithm. Appl Therm Eng 236:121759. https://doi.org/10.1016/j.applthermaleng.2023.121759.

[5] Feng X, Ouyang M, Liu X, et al (2018) Thermal runaway mechanism of lithium ion battery for electric vehicles: A review. Energy Storage Mater 10:246–267. https://doi.org/10.1016/j.ensm.2017.05.013.

[6] Ferreira P, Tang S-X (2024) Quintuple Thermal Model for All-Solid-State Batteries and Temperature Estimation through a Cascaded Thermal-Electrochemical Model. In: 2024 IEEE Conference on Control Technology and Applications (CCTA). IEEE, pp 716–721.

[7] Kim J, Oh J, Lee H (2019) Review on battery thermal management system for electric vehicles. Appl Therm Eng 149:192–212. https://doi.org/10.1016/j.applthermaleng.2018.12.020.

[8] Kumar A, Dewangan AK, Yadav AK (2025) A state-of-the-art review on modelling and simulation of battery thermal management system using phase change material and liquid cooling: Enhancing performance, sustainability, and future research needs. J Energy Storage 133:117978. https://doi.org/10.1016/j.est.2025.117978.

[9] Lin C, Wen H, Liu L, et al (2021) Heat generation quantification of high-specific-energy 21700 battery cell using average and variable specific heat capacities. Appl Therm Eng 184:116215. https://doi.org/10.1016/j.applthermaleng.2020.116215

[10] Liu S, Zhang H, Xu X (2021) a study on the transient heat generation rate of lithium-ion battery based on full matrix orthogonal experimental design with mixed levels. J Energy Storage 36:102446. https://doi.org/10.1016/j.est.2021.102446.

[11] Marzook MW, Hales a, Patel Y, et al (2022) Thermal evaluation of lithium-ion batteries: Defining the cylindrical cell cooling coefficient. J Energy Storage 54:105217. https://doi.org/10.1016/j.est.2022.105217.

[12] Mosavi A (2023) Modelling Thermal Management of Battery Energy Storage System with Machine Learning.

[13] Pesaran AA (2001) Battery Thermal Management in EVs and HEVs: Issues and Solutions. In: Advanced Automotive Battery Conference. Las Vegas, Nevada.

[14] Sarchami A, Kiani M, Najafi M, Houshfar E (2023) Experimental investigation of the innovated indirect-cooling system for Li-ion battery packs under fast charging and discharging. J Energy Storage 61:106730. https://doi.org/10.1016/j.est.2023.106730.

[15] Sarmadian A, Widanage WD, Shollock B, Restuccia F (2023) Experimentally-verified thermal-electrochemical simulations of a cylindrical battery using physics-based, simplified and generalised lumped models. J Energy Storage 70:107910. https://doi.org/10.1016/j.est.2023.107910.

[16] Sheng L, Zhang H, Su L, et al (2021) Effect analysis on thermal profile management of a cylindrical lithium-ion battery utilizing a cellular liquid cooling jacket. Energy 220:119725. https://doi.org/10.1016/j.energy.2020.119725.

[17] Shin S, Jeong D, Kim Y (2026) Deep reinforcement learning-based thermal management of battery subpack in electric vehicle. Comput Chem Eng 204:109406. https://doi.org/10.1016/j.compchemeng.2025.109406.

[18] Smith K, Wang C-Y (2006) Power and thermal characterization of a lithium-ion battery pack for hybrid-electric vehicles. J Power Sources 160:662–673. https://doi.org/10.1016/j.jpowsour.2006.01.038.

[19] Sun Y, Bai R, Ma J (2022) Development and Analysis of a New Cylindrical Lithium-Ion Battery Thermal Management System. Chinese Journal of Mechanical Engineering 35:100. https://doi.org/10.1186/s10033-022-00771-8.

[20] Wang R, Zhang H, Chen J, et al (2024) Modeling and Model Predictive Control of a Battery Thermal Management System Based on Thermoelectric Cooling for Electric Vehicles. Energy Technology 12: https://doi.org/10.1002/ente.202301205.

[21] Wu Q, Huang R, Yu X (2023) Measurement of thermo physical parameters and thermal modelling of 21,700 cylindrical battery. J Energy Storage 65:107338. https://doi.org/10.1016/j.est.2023.107338.

[22] Yang T-F, Lin P-Y, Teng L-T, et al (2022) Numerical and experimental study on thermal management of NCM-21700 Li-ion battery. J Power Sources 548:232068. https://doi.org/10.1016/j.jpowsour.2022.232068.

[23] Zhang T, Yu Y, Yu H, et al (2026) intelligent predictive cooling strategy for liquid-cooled lithium-ion batteries under dynamic operating conditions. Int J Heat Mass Transf 259:128408. https://doi.org/10.1016/j.ijheatmasstransfer.2026.128408

[24] Zhang Z, Min H, Yu Y, et al (2022) an optimal thermal management system heating control strategy for electric vehicles under low-temperature fast charging conditions. Appl Therm Eng 207:118123. https://doi.org/10.1016/j.applthermaleng.2022.118123.