Comparative Analysis and Integration of MPC and RL-Control for Cabin Comfort in Heavy-Duty BEVs

Daniel Linse[1], Kai Franke[1], Prof. Dr.-Ing Stefan Pischinger[1], Dr.-Ing. David Hemkemeyer[2], Dr.-Ing. Tobias Voßhall[3], Dr.-Ing. Christoph Müller[2], Patrick Schutzeich[2]

[1] Chair of Thermodynamics of Mobile Energy Conversion Systems, RWTH Aachen University, Forckenbeckstraße 4, 52074 Aachen, Germany

[2] FEV Vehicle GmbH, Neuenhofstraße 181, 52078 Aachen, Germany

[3] FEV Europe GmbH, Neuenhofstraße 181, 52078 Aachen, Germany

linse@tme.rwth-aachen.de

Abstract: Cabin climate control in battery electric trucks is particularly challenging due to prolonged occupancy periods and the high energy demand of HVAC systems, which can significantly reduce driving range. This work investigates advanced control strategies for cabin thermal management using a model-in-the-loop simulation environment representative of long-haul operations. A model predictive controller (MPC) and a reinforcement learning (RL) framework were developed and benchmarked against a rule-based strategy over a 22-hour driving cycle for a cabin cooling scenario. The MPC improved thermal comfort by maintaining temperature, CO_2 concentration, and humidity within defined limits, while achieving energy usage comparable to the rule-based strategy, thereby demonstrating its capability for multi-objective control under realistic boundary conditions. The initial application of RL as a complementary data-driven approach indicates that comfort targets can be achieved through a direct trade-off between energy use and comfort. However, RL must be extended to true multi-target control to be properly evaluated against the RB approach. In conclusion, the integration effort of both control strategies was assessed, providing an understanding of their respective advantages and limitations for future thermal management applications.

Keywords: Model Predictive Control (MPC), Battery Electric Vehicle (BEV), Reinforcement Learning (RL), Cabin comfort

1 Introduction

The rapid expansion of electromobility is a cornerstone of Europe's climate strategy on the path to carbon neutrality. Policy packages such as the EU Green Deal and "Fit for 55" aim to accelerate the transition by tightening CO_2 limits and incentivizing zero-emission road transport [1]. In Germany, registrations of fully electric trucks increased from 24,380 in 2020 to 92,312 in 2025 [2], yet they still represent only a small fraction of the 3.83 million trucks registered overall in 2025 [3]. Despite growing adoption, there remains a significant gap. Enhancing the real-world performance and appeal of electric heavy-duty trucks continues to be crucial, as range limitations remain a key concern for fleet owners [4].

Contribution: 2025 FKFS Conference on Vehicle Aerodynamics and Thermal Management 15 – 16 October 2025 | Leinfelden-Echterdingen

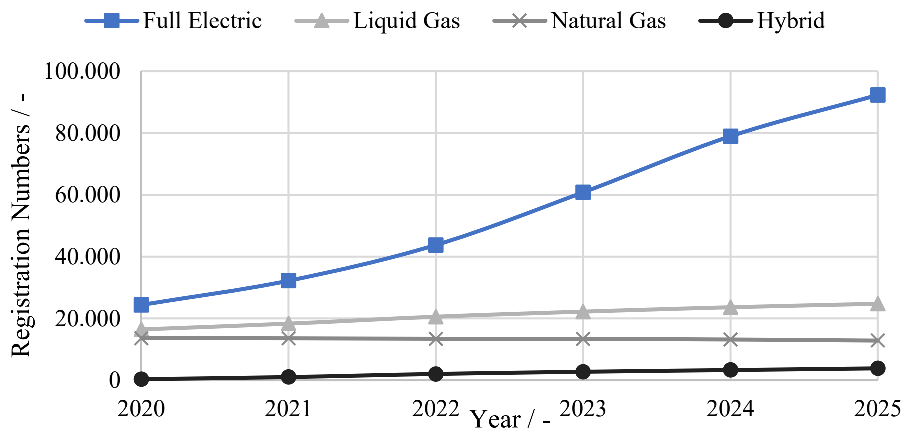

Figure 1: Number of trucks with alternative drivetrains, Germany (2020 - 2025) [2].

For battery electric vehicles (BEVs), cabin air conditioning is a major auxiliary load that can reduce usable driving range under real operating conditions. Especially in cold weather, the absence of powertrain waste heat means the cabin heating demand must be supplied electrically from the traction battery. This has a significant impact on the range [5]. In heavy-duty applications such as battery electric trucks, this challenge is intensified by prolonged cabin occupancy and overnight stays ("hotel function"), which increase HVAC energy demand [6]. Maintaining target temperature, humidity and CO_2 levels while minimizing energy consumption requires advanced, adaptable control strategies [7] [8].

This work investigates intelligent control strategies to improve the energy efficiency and comfort of cabin climate systems in heavy-duty BEVs, focusing on model predictive control (MPC). The MPC leverages accurate thermal models and external data sources, such as ambient forecasts, to manage the cabin climate dynamically under extended-occupancy conditions. It is implemented within a detailed MATLAB/Simulink simulation environment developed as part of the EU research project ESCALATE, which focuses on developing modular, cost-effective heavy-duty vehicles, leveraging the realistic digital twin presented in this study. Furthermore, a preliminary reinforcement learning (RL) framework is introduced as a complementary, data-driven control approach. The method is developed with a focus on general applicability to cabin climatization problems, emphasizing key design aspects such as state representation, reward shaping, and environment interaction. This framework forms the basis for future integration into thermal management systems and enables the development of hybrid control strategies that combine the strengths of RL and MPC.

2 Methodology

This section outlines the two control approaches of MPC and RL, which are integrated and investigated in this work. Both are implemented in a model-in-the-loop (MiL) simulation environment developed in MATLAB/Simulink.

These approaches both offer the ability to develop controllers which are able to do multi-objective control based on optimization methods. The simulation environment features a detailed full thermal system model and a reduced-order model (ROM) of the cabin and HVAC system. The latter enables efficient and realistic closed-loop control development and is described in chapter 2.3. The objective for both controllers in this work builds on the initial development of a cooling controller. It focuses on regulating air temperature, cabin CO_2 concentration, and relative humidity by controlling the compressor rate, blower rate, and recirculation rate.

2.1 Model Predictive Controller

Cabin air conditioning in electric vehicles presents a nonlinear, multivariable control problem (NLP), influenced by ambient conditions, driver demands, and internal system dynamics. Model Predictive Control (MPC) is well-suited for this task due to its ability to predict future states, enforce system constraints, and optimize control actions over a defined prediction horizon [9] [10]. In this work, MPC is implemented using the acados framework, which supports real-time optimization [11]. The user defines the system dynamics, control variables, and the cost function, while the framework handles formulation of the optimization problem and solver execution. The problem is automatically discretized in discrete timesteps k over the prediction horizon on which the cost function J is minimized via direct multiple shooting. Afterwards it is solved with a sequential quadratic programming method (SQP). For a comprehensive treatment of NLP solution methods, the reader is referred to in-depth literature [10] [11].

The main objectives of the controller are contained within a compact representation of the system boundaries (eq. 1) and the cost function (eq. 2) and include maintaining passenger comfort while minimizing energy consumption. The comfort terms are modelled as the system states x and are bound to a lower boundary x_{lb} and an upper boundary x_{ub}. The slack variable x_{slack} allows for constraint relaxation, thereby converting the hard boundary into a soft boundary. The cost function is defined as the sum of the state-related term J_{x} and the cost term J_{u}, which are both evaluated and accumulated over the prediction horizon.

$$x_{\mathrm{lb(k)}} \leq x(k) + x_{\mathrm{slack(k)}} \leq x_{\mathrm{ub(k)}} \tag{1}$$

$$J = \sum_{k=1}^{N} J_{\mathrm{x(k)}} + \sum_{k=0}^{N-1} J_{\mathrm{u(k)}} \tag{2}$$

Minimization of the cost function inherently enforces the state constraints. Equation 3 defines the state-related term as the weighted sum of all quadratic slack variables of the MPC. The weighting is determined by the tunable matrix Q. The central slack variables consist of the cabin air temperature T_{slack}, the CO_2 concentration $X_{\mathrm{CO2,\,slack}}$ and the relative humidity $X_{\mathrm{hum,\,slack}}$.

$$J_{\mathrm{x(k)}} = Q_{\mathrm{T_{slack}}} T_{\mathrm{slack}}(k)^2 + Q_{\mathrm{X_{CO2,\,slack}}} X_{\mathrm{CO2,\,slack}}(k)^2 + Q_{\mathrm{X_{hum,\,slack}}} X_{\mathrm{hum,\,slack}}(k)^2 \tag{3}$$

Finally, the cost term in eq. (4) contains the weighted quadratic energy consumption of the compressor E_{Cpr} and the cabin blower E_{Blower}. The power consumption of the recirculation flap is assumed negligible.

$$J_{\mathrm{u(k)}} = Q_{\mathrm{Cpr}} E_{\mathrm{Cpr}}(k)^2 + Q_{\mathrm{Blower}} E_{\mathrm{Blower}}(k)^2 \qquad (4)$$

With the problem formulation established, the MPC is then integrated into a model-in-the-loop (MiL) environment in MATLAB/Simulink, which consists of three core components: the MPC controller, the plant model, and the prediction module. The controller operates with a fixed sampling interval of 10 s, selected to ensure both robust control performance and real-time feasibility. This interval reflects the slow thermal dynamics of the cabin environment and the response characteristics of the actuators. At each step, the controller computes the optimal control and state trajectories over a prediction horizon of 10 minutes. This horizon length is sufficient to capture the dominant cabin thermal dynamics and anticipated disturbances, while allowing the system states to be adjusted within this time frame. After the optimization is completed, only the first element of the control input vector is applied to the plant model. The plant, representing the HVAC and cabin thermal system, simulates the system response at each step, and the updated states are fed back to the MPC, thereby closing the control loop. The prediction module provides forecasted boundary conditions, including relevant ambient factors and comfort constraints derived from route and trip data (e.g., planned stops). By solving the optimization problem over the moving horizon, the controller can anticipate future disturbances and proactively adjust its control actions. Ultimately, the MPC is applied to the reduced-order model, enabling efficient testing and iteration without compromising the physical relevance of the results.

2.2 Reinforcement Learning Controller

RL is investigated in this research as a data-driven control strategy for cabin climate management. A Deep Deterministic Policy Gradient (DDPG) agent is employed, following an actor–critic structure (Fig. 2). DDPG agents can operate in a continuous action space and are therefore suitable for fully variable control [12] [13]. An RL agent interacts with its environment in a closed loop: it observes the current state s_{t}, selects an action a_{t}, and receives the next state $s_{(\mathrm{t}+1)}$ together with a scalar reward r_{t}. By maximizing cumulative rewards through repeated interaction, the agent gradually learns strategies that optimize long-term performance. The DDPG architecture is based on the following three fundamental components:

- Actor (policy network θ^{μ}): maps states s_{t} to continuous actions $a_{\mathrm{t}} = \mu(s_{\mathrm{t}} \,|\theta^{\mu})$.
- Critic (Q-network θ^{Q}): evaluates these actions by estimating their long-term value $Q(s_{\mathrm{t}}, a_{\mathrm{t}})$.
- Training process: the Actor is updated via policy gradients, the Critic via minimization of a loss function. Soft target updates are applied to stabilize training.

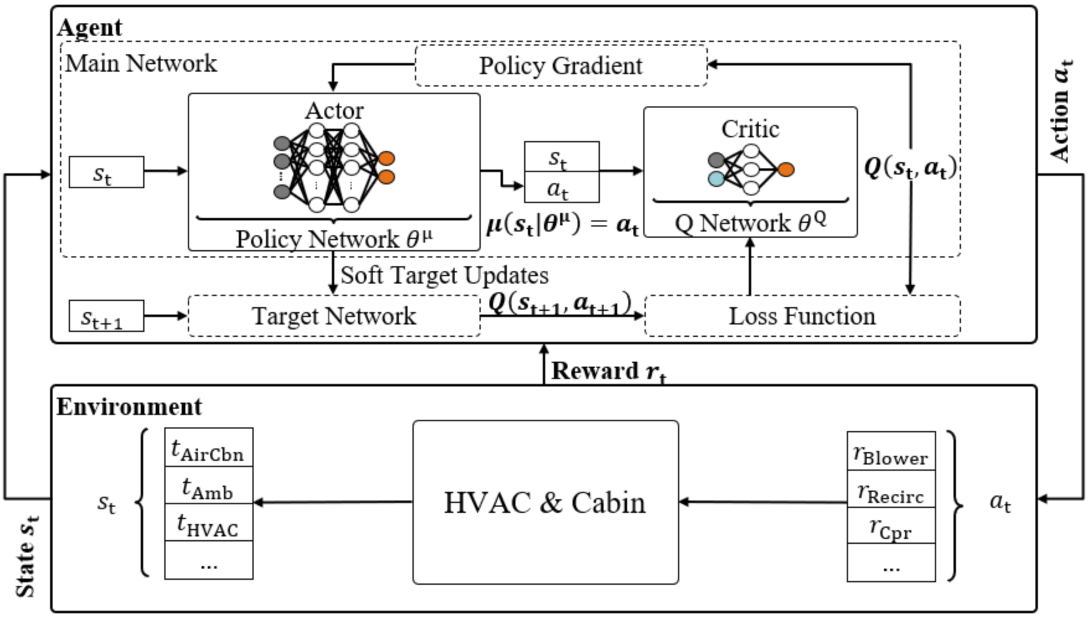

Figure 2: Integration of the RL agent framework inside the simulation environment.

Two complementary training approaches are pursued. In the first, the agent is trained directly on the high-fidelity digital twin, which avoids model reduction and ensures maximum physical fidelity, but at the cost of substantial computational resources. In the second, training is performed on the ROM, enabling faster iterations, direct benchmarking with MPC, and providing insight into potential hybrid MPC–RL strategies. The RL agent processes the system states cabin air temperature (t_{CbnAir}) and setpoint $t_{CbnAirSP}$, ambient temperature (t_{Amb}), HVAC outlet air temperature ($t_{HVACAirOut}$), and action signals (current $r_{Comp,t}$ and previous $r_{Comp,t-1}$). Based on these inputs, it determines the control action of the compressor (r_{Comp}). The agent is trained using a composite reward that balances comfort, energy efficiency, and smoothness of control (see Figure 3). The temperature reward targets the passenger comfort by penalizing deviations from the cabin air temperature setpoint. The energy reward promotes efficient operation by discouraging high activation levels of HVAC components. Finally, the smoothness reward penalizes abrupt control changes, targeting stable and hardware-friendly operation.

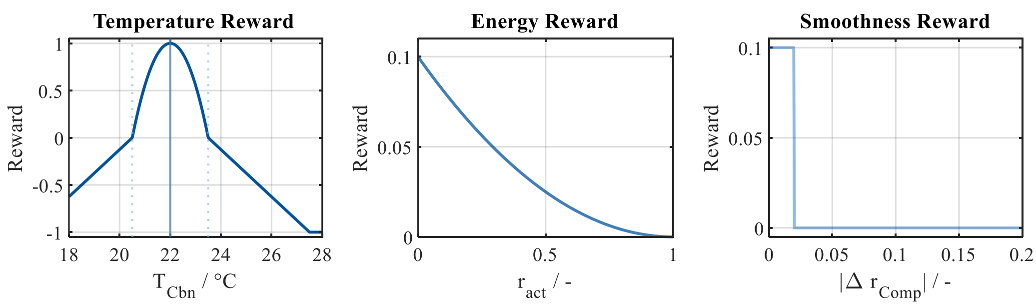

Figure 3: Sub-reward functions for temperature, energy, and smoothness.

2.3 Reduced-Order-Model of Cabin and HVAC

To enable efficient development of the controllers, a reduced-order-model of the vehicle cabin and HVAC system was created. The ROM captures the essential thermal dynamics while ensuring low computational load and enabling rapid simulation. Validation was conducted using the full thermal system model. This model was previously plausibilized to represent the thermal behavior of a generic truck, ensuring physical consistency.

The cabin is modeled as a single-zone thermal system, assuming spatially uniform air temperature. The core of the model is an energy balance over the enclosed air volume:

$$\frac{dE_{Cabin}}{dt} = \sum \dot{Q}_i + \sum \dot{m}_{in}\, h_{in} - \sum \dot{m}_{out}\, h_{out} \qquad (5)$$

This equation accounts for internal energy changes due to air temperature variation, enthalpy flows from ventilation, and additional heat sources or sinks $\dot{Q}_i$, such as solar radiation or internal gains. To represent the thermal inertia of the cabin interior, a lumped thermal mass is included in equation 6. This mass represents components like seats, dashboard and other internal masses and is thermally coupled to the cabin air via a resistance $R_{\text{th, Interior}}$.

$$\dot{Q}_{\text{Interior}} = \frac{T_{\text{Air,Cabin}} - T_{\text{Interior}}}{R_{\text{th, Interior}}} \qquad (6)$$

This formulation allows the model to replicate realistic heating and cooling dynamics, including the effect of delayed thermal response. The resistance $R_{\text{th, Interior}}$ and other parameters were validated within the Simulink full vehicle model. The HVAC system is modeled with key functionalities such as air mixing, heating, and cooling via heat exchangers, as well as flap positions to switch between fresh air and recirculation modes. The blower fan modulates the air mass flow supplied to the cabin.

In addition to thermal comfort, the model also considers air quality, specifically the accumulation of CO_2 in recirculation mode. The mass balance for CO_2 includes both external input and occupant respiration as source terms:

$$\dot{m}_{\text{Air}} X_{\text{Inlet}} + \dot{m}_{\text{CO}_2\text{, Passengers}} = M \frac{dx_{\text{Cabin}}}{dt} + \dot{m}_{\text{Air}}\, X_{\text{Cabin}} \qquad (7)$$

Here, x_{Inlet} and x_{Cabin} are the CO_2 concentrations of incoming air and cabin air, and M is the total cabin air mass. Leakages are neglected in this reduced order model. Under full recirculation, CO_2 concentration increases steadily, making the model suitable for studying the trade-off between energy efficiency and air quality. As a second metric for air quality the cabin air humidity is modelled by a control volume mass balance of the water vapor. The rate of change of the cabin humidity ratio $\dot{w}_{Air}$ is expressed as:

$$\dot{w}_{\text{Air}} = \frac{1}{m_{\text{Air}}} \left[\dot{m}_{\text{Air}}(w_{\text{mix}} - w_{\text{Cabin}}) + \dot{m}_{\text{gen}} + \dot{m}_{\text{deh}}\right] \qquad (8)$$

Here, m_{Air} denotes the total air mass inside the cabin. The first term accounts for the humidity change due to exchange with the supply airflow, with $\dot{m}_{Air}$ representing the air mass flow into the cabin and w_{mix} the humidity ratio of the mixed ambient and recirculated air. The second term $\dot{m}_{gen}$, describes the generation of water vapor by passengers (e.g., through breathing and perspiration). The third term $\dot{m}_{deh}$, represents the removal of moisture through condensation at the evaporator coil surface whenever the incoming air exceeds the saturation limit at coil temperature. Overall, this ROM provides the necessary balance between physical fidelity and computational efficiency, enabling its use in the computational expensive MPC and RL control development.

3 Results & Discussion

Using the MiL environment described previously, the performance of the MPC and the RL controller was evaluated under varying ambient conditions based on real-world measurement data. To assess robustness, stochastic disturbances were added to the ambient temperature, humidity, and solar radiation, introducing controlled misalignment between the plant model and the controller's internal prediction model.

The evaluation covered a 22 h 15 min cycle representative of long-haul truck operation. The cycle consisted of two driving phases of 3 h 45 min and 5 h, a 12 h overnight stay, and 45 min rest periods in between. Driver presence is assumed according to an expected occupancy schedule. During the overnight stay, the sleep mode is activated in which the cabin temperature setpoints are lowered to enhance comfort during rest. During the day, these values are set at 22 °C for the upper boundary and at 20 °C for the lower boundary. The relative humidity setpoints during the day are defined between 30 - 60% and are bound to a stricter window of 35 - 55% during sleep mode. For the CO_2 concentration an upper boundary of 1200 ppm is defined throughout occupancy of the cabin. This setup provides a comprehensive framework for evaluating the controllers under realistic, time-varying thermal boundary conditions, including heat soak, idle periods, and extended occupancy.

To establish a baseline, a rule-based strategy with fixed recirculation rates was applied across the cycle to control the cabin air temperature. Figure 3 illustrates the effect of increasing the recirculation rate on total HVAC energy consumption.

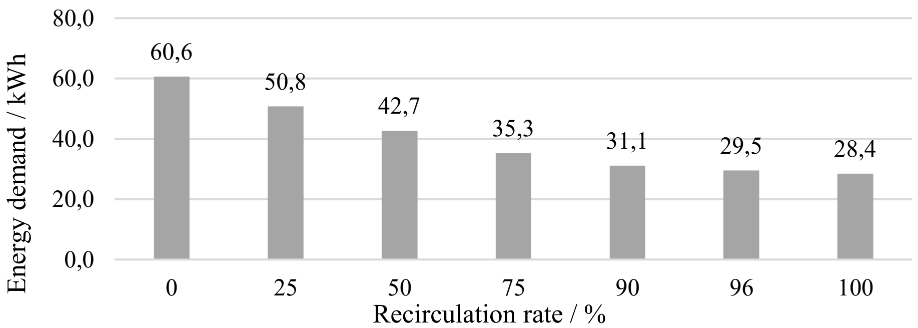

Figure 3: Energy consumption of the rule-based HVAC control at average ambient temperature of 32 °C with varying recirculation rates.

A clear trend is visible: as the recirculation rate increases from 0% to 100%, the energy demand steadily decreases from 60.6 kWh to 28.4 kWh, corresponding to a total reduction of approximately 53.1%. However, high recirculation rates also impact the performance in terms of cabin comfort such as temperature targets, relative humidity, and CO_2 concentration. Simulation results show that the CO_2 concentration threshold of 1200 ppm is only exceeded at recirculation rates above 96 percent, indicating that high recirculation levels can maintain acceptable air quality while still offering energy savings. These findings highlight the need for dynamic control strategies, which can continuously balance energy efficiency and air quality, rather than relying on static setpoints.

Consequently, the performance of the MPC, RL and RB strategy are compared in the following section. As initially stated, to complement the comparison a preliminary RL framework was developed and applied. Similar to the RB approach, only the cabin temperature was controlled using compressor actuation alone, while maintaining a fixed recirculation rate and blower speed. To ensure comparability, the recirculation rate of the RB approach was selected to align its energy usage with that of the MPC approach and ensure a sufficient rate of fresh air. To quantify the control quality of the strategies, the target deviation $|\Delta x(t)|$ is time-averaged over the periods in which the boundaries are exceeded, resulting in three comfort metrics for temperature, CO_2 and humidity. The first metric is the temperature comfort, which is displayed in figure 4.a. The air quality is evaluated for CO_2 levels in figure 4.b and for the relative humidity in figure 4.c.

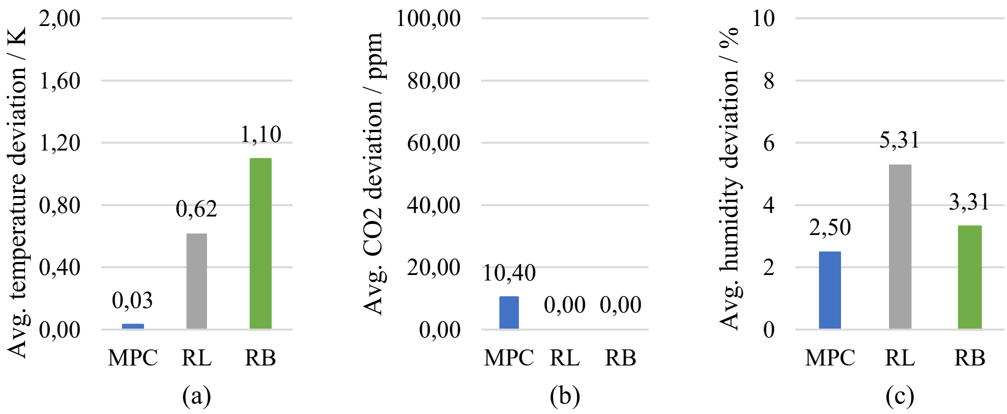

Figure 4: Comparison of comfort metrics for the MPC, RL, and RB strategy

The MPC strategy achieves the lowest average deviation of 0.03 K compared to 1.1 K with the RB strategy. This is also reflected in the resulting maximum deviations. The MPC maximum deviation from the control target of 6.21 K is present at the beginning of cabin conditioning, while the RB strategy reaches an even higher value of 16.11 K but at a later stage due to hot soak during the idle phase. During this idle phase the MPC achieves a much lower deviation of 2.15 K by pre-conditioning the cabin according to the predicted change in the temperature boundaries.

The RL strategy achieves an average deviation of 0.62 K but, lacking the predictive capability, does not pre-condition the cabin and therefore reaches maximum deviation of 16.45 K, similar to the RB strategy. These comfort gains of MPC and RL come at only a minor increase in energy consumption, with 30.18 kWh and 30.13 kWh respectively, compared to 29.48 kWh for the RB strategy.

Consequently, air quality discomfort is also evaluated. The MPC targets to maintain CO_2 levels below a soft-constrained upper bound. This bound is exceeded in some cases due to the balancing of the multi-target optimization and the improved convergence of the optimization. This results in a recorded average deviation of 10.4 ppm, which is negligible compared to typical indoor CO_2 concentration fluctuations and has no perceptible impact on passenger comfort. In contrast, the RB and RL strategies do not enforce any explicit CO_2 limit since CO_2 is not directly controlled. As the setpoint was defined to ensure sufficient fresh air, both strategies show no violation of the CO_2 discomfort metric

Lastly, humidity deviation is evaluated. The MPC strategy achieves the lowest average deviation at 2.5% compared to 3.31% for the RB approach, as the MPC explicitly enforces the humidity boundary. The RL strategy exhibits a considerably higher deviation of 5.31% since humidity is not actively controlled, leading to excessive dehumidification at the evaporator caused by increased compressor actuation. Given the average outdoor humidity of 31%, these deviations remain non-critical under the tested conditions due to overall system stability, though they may become more relevant in high-humidity heating scenarios.

Overall, MPC achieves the best balance between comfort and energy consumption by minimizing temperature and humidity deviations while maintaining CO_2 within acceptable limits. The preliminary RL controller shows intermediate performance by improving temperature target deviation, but lacks predictive capabilities and an extended actuator control. Nevertheless, it offers a promising foundation for further development toward a highly automatable controller design.

In the following the findings regarding the integration of both MPC and the initial RL approach are summarized in figure 5.

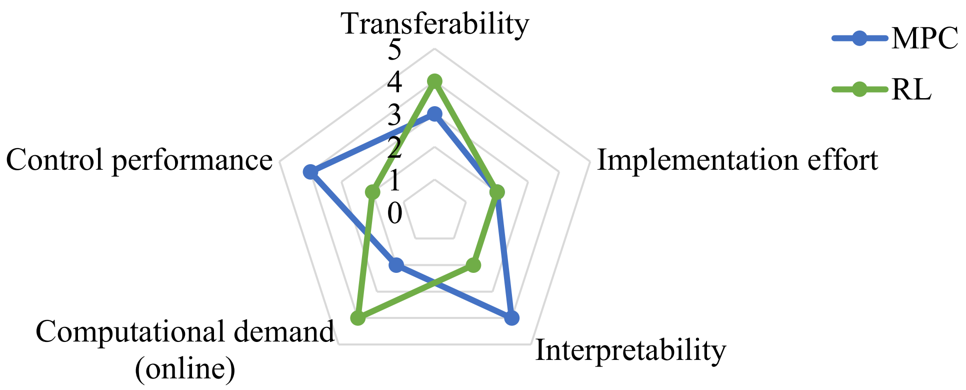

Figure 5: Comparison of evaluated performance metrics for MPC and RL approach for the MiL application in this work.

Here, various performance metrics that were evaluated during the development are rated according to the following scale:

1: Very poor (significant drawbacks), **2**: Poor (limited suitability),
3: Moderate (meets minimum requirements), **4**: Good (only minor limitations),
5: Excellent (highly suitable)

The transferability of the MPC is rated as moderate. A major challenge in this work was the application of the MPC framework to the full thermal system model. The higher complexity of the digital twin caused a divergence between the internal system dynamics of the MPC and the actual plant behavior, which prevented direct application and led to high deviations to the control targets in the cabin control case. This highlights a well-known limitation of MPC, namely the need for model reduction to embed an equation-based system model while ensuring real-time feasibility, which inevitably leads to a model-plant mismatch [14] [15]. Within digital twins such model reduction is often not straightforward.

One possible mitigation strategy is the use of data-driven surrogate models, although this requires additional efforts for system identification and validation [16]. In contrast, the RL agent could also be trained directly in the full model environment and achieved comparable control performance to the reduced-order model for the cabin conditioning task. This potentially eliminates the need for model reduction and represents a clear advantage in the digital twin environment. However, this advantage is specific to simulations, since the training of RL on real systems would require considerably more time and resources. For this reason, RL received a higher transferability rating of four.

The implementation effort of MPC is rated at two, as is the case for RL. For MPC, the main effort lies in the development and validation of an internal prediction model tailored to the cabin thermal dynamics. The model must be suitable for optimization and verified against various measurements or simulation data, which is an obstacle to fast integration. RL requires less manual modelling effort since the framework only needs to be provided with selected observations and actions. The training of the neural networks is then conducted automatically. Due to the inherent flexibility of the RL framework, it can be implemented in a wide range of environments and therefore offers broad applicability without requiring the detailed modelling knowledge that is essential for a grey-box MPC. However, in this work only a simplified control objective was investigated. The implementation effort and training requirements are expected to increase once additional control variables are incorporated.

Interpretability plays a decisive role in improving control performance. MPC scores higher in this category with a rating of four, whereas RL is rated at two. MPC benefits from the possibility of incorporating grey-box models, which supports verification of predictions and thereby improves understanding of controller behavior. Furthermore, the tuning of the cost function can be conducted in an intuitive manner.

RL, on the other hand, operates as a black box and lacks direct interpretability, which results in an iterative and often time-consuming process of reward shaping, network architecture selection, and hyperparameter tuning. Automating this process through optimization-based hyperparameter tuning can resolve this issue as shown in [12]. However, the DDPG algorithm used in this study is deterministic, which means that it provides consistent outputs when presented with the same observations. This property allows the user to draw limited conclusions from the observed controller behavior [13]. In line with these limitations, the RL strategy received a lower rating of two.

With respect to computational demand, RL performs better than MPC. RL is rated at four, while MPC is rated at two. In this work the MPC used an SQP solver with average solution times below 0.5 seconds for a 2.3 GHz CPU, which is sufficient for real-time operation in the cabin conditioning task. Nevertheless, execution times are expected to increase if the controller is implemented on an embedded microcontroller with limited processing power. RL execution times are negligible once the training is complete, which represents a clear advantage. The drawback lies in the training phase, which is computationally intensive. For the compressor control task, a total of 500 episodes were run in parallel with 4 CPU cores, resulting in more than 10 hours of training. This effort, however, occurs entirely offline and does not affect embedded system performance.

Overall, MPC is better suited to the problem under consideration, as its interpretability enables more targeted and reliable implementation, while RL should be regarded as a preliminary approach at this stage.

4 Conclusion

Model-predictive-control demonstrated its capability to enforce operational boundaries and balance energy efficiency for multi-objective control. Reinforcement learning was introduced as a framework for cabin climatization, and an initial controller was evaluated to determine the requirements for control design. Open challenges remain in managing multi-objective control, along with the need for further investigation into hyperparameter tuning and reward shaping. Both strategies provide distinct advantages, with MPC excelling in interpretability and offering a more transparent design process, while RL shows superior transferability to detailed thermal system models. Future work will extend the MPC with active heating and advanced humidity control, explore data-driven models as internal predictors, and expand RL to additional control variables. Furthermore, synergies between RL and MPC will be evaluated, with the aim of combining the interpretability and constraint-handling of MPC with the adaptability and transferability of RL to enable more efficient cabin climatization strategies.

Acknowledgements

This project has received funding from the European Union’s Horizon Europe research and innovation programme under the Grant Agreement No. 101096598

5 References

[1] Europäische Kommision, "Verordnung (EU) 2023/85 zur Änderung der Verordnung (EU) 2019/631 im Hinblick auf eine Verschärfung der CO2-Emissionsnormen für neue Personenkraftwagen und für neue leichte Nutzfahrzeuge im Einklang mit den ehrgeizigeren Klimazielen der Union," 2023.

[2] KBA, "Anzahl der Lastkraftwagen mit alternativen Antrieben in Deutschland in den Jahren 2020 bis 2025," [Online]. Available: https://de.statista.com/statistik/daten/studie/259803/umfrage/lkw-bestand-mit-alternativen-antrieben-in-deutschland/. [Accessed 08 09 2025].

[3] KBA, "Fahrzeugbestand von Straßenfahrzeugen in Deutschland 2025," [Online]. Available: https://www.bmv.de/SharedDocs/DE/Artikel/G/fahrzeugbestand.html. [Accessed 11 09 2025].

[4] C. Sugihara, S. Hardman and K. Kurani, "Social, technological, and economic barriers to heavy-duty truck electrification," *Research in Transportation Business & Management,* no. 51, p. 101064, 2023.

[5] Daimler Truck, "Cold, ice and snow successfully defied," 2023. [Online]. Available: https://www.daimlertruck.com/en/newsroom/pressrelease/cold-ice-and-snow-successfully-defied-mercedes-benz-trucks-tests-electric-trucks-in-finland-52179205. [Accessed 12 09 2025].

[6] K. Sheth, D. Patel and G. Swami, "Reducing Electrical Consumption in Stationary Long-Haul Trucks," *Open Journal of Energy Efficiency,* no. 13, pp. 88-99, 2024.

[7] S. H. Kim and H. J. Moon, "Case study of an advanced integrated comfort control algorithm with cooling, ventilation, and humidification systems based on occupancy status," *Building and Environment,* no. 133, pp. 246-264, 2018.

[8] Q. Wu, B. Ma and L. Guo, "Cabin Temperature and Humidity Comfort Control for Electric Vehicles in High Temperature and Humidity Environment," in *China Automation Congress (CAC)*, Qingdao, 2024.

[9] M. Auer, Ein Beitrag zur Erhöhung der Reichweite eines batterieelektrischen Fahrzeugs durch prädiktives Thermomanagement, Wiesbaden: Springer Vieweg, 2016.

[10] M. Papageorgiou, M. Leibold and M. Buss, "Optimierung: Statische, dynamische, stochastische Verfahren für die Anwendung," Springer Vieweg Berlin, Heidelberg, 2015.

[11] R. Verschueren, G. Frison, D. Kouzoupis, J. Frey, N. Van Duijkeren, A. Zanelli, B. Novoselnik, T. Albin, R. Quirynen and M. Diehl, "acados—a modular open-source framework for fast embedded optimal control," *Mathematical Programming Computation 14,* vol. Nr. 1, p. 147–183, 2022.

[12] J. Wang, C. Du, F. Yan, M. Hua, X. Gongye, Q. Yuan, H. Xu and Q. Zhou, "Bayesian optimization for hyper-parameter tuning of an improved twin delayed deep deterministic policy gradients based energy management strategy for plug-in hybrid electric vehicles," *Applied Energy,* no. 381, p. 125171, 2025,.

[13] E. H. Sumiea, S. J. Abdulkadir, H. S. Alhussian, S. M. Al-Selwi, A. Alqushaibi, M. G. Ragab and S. M. Fati, "Deep deterministic policy gradient algorithm: A systematic review," *Heliyon,* no. 10, p. 30697, 2024.

[14] A. S. Badwe, S. L. Shah, S. C. Patwardhan and R. S. Patwardhan, "Model-Plant Mismatch Detection in MPC Applications using Partial Correlation Analysis," in *Proceedings of the 17th World Congress*, Seoul, 2008.

[15] S. J. Kuntz and J. B. Rawlings, "Beyond inherent robustness: strong stability of MPC despite plant-model mismatch," *IEEE Transactions On Automatic Control,* pp. 1-13, 2025.

[16] K. Prag, M. Woolway and T. Celik, "Toward Data-Driven Optimal Control: A Systematic Review of the Landscape," *IEEE Access,* no. 10, pp. 32190-32212, 2022.