Penerbit ITDel

**Journal of Intelligent Computing, Systems, and Applications**
e-ISSN: 3123-1705

JICSA

Vol. 2 No. 1 (2026) 34-49
https://journal.del.ac.id/jicsa

# Experimental Comparison of Deployed DQN and Rule-Based Climate Controllers in Indoor Farming: Climate Performance and Electrical Energy Trade-offs

**Raj David Francisco Damanik[1*], Samuel Abednego Sormin[1], Albert Sagala[1]**

[1]*Electrical Engineering, Faculty of Informatics and Electrical Engineering, Institut Teknologi Del, Toba, North Sumatra 22381, Indonesia*

*Corresponding Author: els22018@students.del.ac.id

**Article Info**

Received: 29 July 2026
Revised: 15 September 2026
Accepted: 17 September 2026
Available online: 30 September 2026

**Keywords**

Deep Q-Network, reinforcement learning, indoor farming, climate control, electrical energy consumption

**Abstract**

Indoor farming requires coordinated environmental control while limiting electrical energy use. This experimental case study compares two deployed controller configurations on the same physical platform using separate 24-hour records: a fixed Deep Q-Network (DQN) policy and a hysteresis rule-based controller. The DQN used an 11-feature state, an 11-64-64-8 neural network, eight discrete actuator combinations, and operational action constraints. The deployed control intervals were 10 min for DQN and 5 min for rule-based control. Performance was evaluated using target-band compliance, midpoint mean absolute error (MAE), excursion magnitude, measured actuator electrical energy, and actuator-status telemetry. Temperature compliance was 79.86% for DQN and 77.37% for rule-based control, with midpoint MAE values of 1.10 and 1.28 °C. During lamp-on periods, CO2 compliance was 45.31% and 18.23%, with MAE values of 119.31 and 221.69 ppm, respectively. Relative-humidity regulation was limited by the absence of active dehumidification. Controller-dependent electrical energy was 4.508 kWh in the DQN window and 3.788 kWh in the rule-based window. CO2 gas consumption could not be quantified because no flow meter was installed. Because the deployments differed in date, outdoor conditions, update interval, and auxiliary operational constraints, the findings are interpreted as configuration-level field observations rather than isolated causal effects of the decision algorithms.

## 1. Introduction

Controlled-environment agriculture (CEA) and indoor farming use sensing, automation, and environmental-control equipment to maintain selected growing conditions while reducing dependence on

This is an open access article under the CC BY-NC-SA 4.0 license.

outdoor weather. Reviews of CEA describe vertical farms and greenhouses as highly controllable systems in which temperature, relative humidity, light, and carbon dioxide (CO2) are key environmental variables [1], [2]. Recent sensing reviews likewise emphasize continuous environmental measurements and connected monitoring as core components of indoor-farming management [3]. This controllability is accompanied by a substantial resource burden: comparative work on plant factories and greenhouses identifies purchased energy as a major consideration [4], while recent plant-factory reviews identify artificial lighting and heating, ventilation, and air-conditioning (HVAC) as dominant electrical loads [5]. Therefore, practical controller evaluation should consider both environmental regulation and the electrical and actuator behavior required to achieve it.

Conventional threshold, on/off, and hysteresis control remain attractive because their logic is transparent, computationally inexpensive, and straightforward to implement. At the same time, contemporary smart-greenhouse research increasingly combines sensing, data-driven models, optimization, and intelligent control to address nonlinear and coupled climate dynamics [6], [7]. Reviews of deep learning in CEA report applications in microclimate prediction and energy-aware control [8]; greenhouse studies have investigated synthetic-data support for intelligent climate models [9], neural-network-based decoupling control in plant factories [10], deep-learning-assisted predictive energy management [11], integrated greenhouse energy management [12], and machine-learning forecasting of greenhouse climate variables [13]. These developments motivate evaluating learning-based controllers not only in simulation but also under operational deployment conditions.

Reinforcement learning (RL) formulates control as sequential decision making in which actions are selected to maximize accumulated reward. Q-learning established a model-free framework for learning state-action values [14], and Deep Q-Network (DQN) extended value-based RL by using a deep neural network to approximate action values [15]. Recent greenhouse studies have explored energy-efficient deep RL with robust optimization [16], direct comparisons of RL and model predictive control (MPC) [17], deep-RL-assisted HVAC tuning [18], IoT-integrated RL greenhouse control [19], RL-based MPC for uncertainty handling [20], and broader prediction-and-control approaches for greenhouse temperature [21]. Interactive RL has also been proposed to incorporate grower input, although current evaluation remains simulation-based [22]. Collectively, the literature shows strong interest in learning-based climate control, but also shows that reported performance depends on evaluation setting, objectives, operational constraints, and available actuators.

This paper addresses a narrower practical question: how did a fixed, previously trained DQN deployment behave relative to a rule-based deployment on the same indoor-farming platform? The study does not introduce a new RL algorithm or retrain the policy. Instead, it re-analyzes existing field records from two equal-duration 24-hour deployment windows. The comparison concerns the complete deployed configurations rather than isolated decision algorithms, because the two deployments differed in control-update interval and auxiliary operational constraints. Two research questions are addressed: (RQ1) How did the deployed DQN and rule-based configurations differ in maintaining the selected temperature, relative-humidity, and CO2 evaluation bands? (RQ2) What trade-offs were observed among climate-control performance, measured electrical energy, and actuator operation? The contribution is a configuration-level field comparison that jointly reports environmental compliance, midpoint deviation, excursion magnitude, electrical energy, and actuator behavior while explicitly documenting the practical differences between the deployments.

## 2. Materials and Methods

### 2.1 Experimental Indoor Farming System

The experimental platform was an operational indoor-farming system instrumented with indoor and outdoor environmental sensing, actuator power/status telemetry, an InfluxDB time-series database, and MQTT-based actuator commands. Indoor measurements used by the controllers included temperature,

relative humidity, illuminance, and CO2 concentration; outdoor temperature, relative humidity, and illuminance were also available. The controllable climate devices considered in this paper were an air conditioner (AC), humidifier, and CO2 injection solenoid. Lighting followed the existing photoperiod and its measured state was included in the DQN state. The DQN deployment used a 600 s (10 min) decision interval, while the rule-based deployment used a 300 s (5 min) decision interval. The CO2 injection line was actuated by an electrically controlled solenoid without a flow meter; consequently, gas volume or mass was not measured. Figure 1 summarizes the shared platform and the two controller paths.

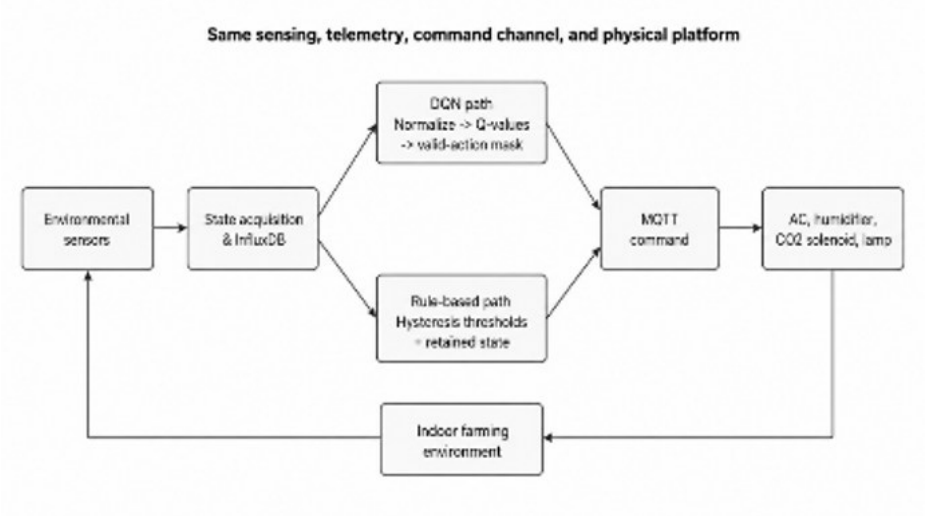



**Fig. 1** ***Experimental architecture of the shared platform and the DQN/rule-based decision paths. Both configurations use the same sensing, telemetry, MQTT command channel, and physical platform; the decision logic and auxiliary constraints differ. Source: developed by the authors from the deployment software and telemetry configuration.***

For evaluation, the selected bands were 20-24 °C for indoor temperature, 60-70% for relative humidity, and 800-1000 ppm for CO2. These are experiment-specific control objectives configured for this platform and are not presented as universal crop-optimal ranges. CO2 performance is reported primarily during lamp-on periods because both controller designs prevent deliberate CO2 injection while the lamp is off.

**Table 1** *Evaluation variables and bands*

| Variable | Evaluation band | Primary analysis period |
|---|---|---|
| Temperature | 20-24 °C | 24 h |
| Relative humidity | 60-70% | 24 h |
| CO2 | 800-1000 ppm | Lamp-on samples |

## 2.2 Control Strategies

### 2.2.1 Rule-Based Controller

The baseline controller used independent hysteresis rules and the measured actuator state as the retained state inside each deadband. The AC was commanded ON when indoor temperature exceeded 24 °C and OFF when temperature was at or below 20 °C. The humidifier was commanded ON below 60% RH and OFF at or above 70% RH. CO2 injection was forced OFF while the lamp was OFF; during lamp-on operation it was commanded ON below 800 ppm and OFF at or above 950 ppm. Inside each deadband, the current actuator state was retained. The rule-based deployment did not use DQN inference, state normalization, DQN action masking, or the DQN AC restart-protection layer.

### 2.2.2 DQN Training and Model Selection

The deployed policy was produced by simulation-based Vanilla DQN training using the same 11-feature state order, min-max normalization limits, 11-64-64-8 network architecture, and eight AC/humidifier/CO2 action combinations used by the deployment artifact. One simulation step represented 15 min and each episode contained 96 steps, corresponding to 24 h. The full configuration specified 3000 training episodes with random episode start time and initial conditions drawn from five scenarios (normal, hot, high-RH, low-CO2, and mixed adverse conditions). Training used valid-action masking and minimum ON/OFF constraints inside the simulator; these training constraints were not identical to the later deployment safety layer described in Section 2.2.3.

Learning used a replay buffer of 50,000 transitions with a 2,000-transition warm-up and mini-batches of 64. The policy network was optimized with Adam at a learning rate of 1e-4 using Huber (Smooth L1) loss, discount factor gamma = 0.95, gradient-norm clipping at 10, and a hard target-network update every 500 gradient steps. Epsilon-greedy exploration decayed from 1.0 to a floor of 0.02 using a per-episode factor of 0.9975. The random seed was 42.

The training reward was multi-objective and can be represented schematically as r_t = B_zone - P_T - P_RH - P_CO2 - P_power - P_switch - P_unsafe. B_zone rewarded operation inside the experiment-specific temperature, RH, and lamp-conditioned CO2 bands. P_T, P_RH, and P_CO2 were quadratic deviation penalties; P_power used normalized nominal actuator power; P_switch penalized actuator-state changes; and P_unsafe penalized operation outside predefined safety bounds. The documented reward did not include a dedicated asymmetric CO2-overshoot term. The nominal-power term was used only during simulation training and is distinct from the measured electrical-energy telemetry reported in Section 3.

**Table 2** *DQN training and model-selection configuration*

| Item | Configuration |
|---|---|
| Algorithm / network | Vanilla DQN; 11-64-64-8; ReLU hidden layers; 8 linear Q-value outputs |
| Training horizon | 3000 episodes; 96 steps/episode; 15 min/step (24 h simulated day) |
| Replay / batch | 50,000 capacity; 2,000 warm-up; batch size 64 |
| Optimizer / loss | Adam, learning rate 1e-4; Huber (Smooth L1); gradient clip 10 |
| Discount / target update | gamma = 0.95; hard target-network update every 500 gradient steps |
| Exploration | epsilon-greedy: 1.0 to 0.02; decay factor 0.9975 per episode |
| Initialization | Random start hour; normal, hot, high-RH, low-CO2, and adverse-mix scenarios |
| Evaluation / seed | Every 50 episodes using 10 greedy evaluation episodes; random seed 42 |
| Reward objective | Multi-objective reward: ideal-zone bonuses; quadratic T/RH/lamp-conditioned CO2 deviation penalties; normalized nominal-power penalty; actuator-switching penalties; unsafe-state penalty |
| Selected checkpoint | Evaluation log last identified the best checkpoint at episode 2650 |
| Export / traceability | best_model.pth -> dqn_weights.npz; exported network reproduced logged deployment Q-values within 3.1e-4 maximum absolute difference on 10 checked deployment records |

The evaluation log evaluated the policy every 50 episodes without exploration and the final best-model flag occurred at episode 2650. The export pipeline converts the selected PyTorch policy checkpoint to NPZ for lightweight deployment. As an artifact-consistency check for this study, forward inference using the supplied NPZ weights, training normalization, and recorded deployment states reproduced the Q-values stored in the deployment log to a maximum absolute difference of 3.1 x 10^-4 across 10 checked records, with identical greedy action indices. This verifies the training-to-export-to-deployment model linkage for the analyzed field record.

### 2.2.3 DQN Policy and Operational Constraints

The deployed DQN received the 11-element state vector s_t = [T_in, RH_in, Lux_in, CO2_in, T_out, RH_out, Lux_out, u_AC, u_H, u_CO2, u_L]. The seven environmental values were min-max normalized and clipped to [0,1], while the four device states remained binary. The network architecture was 11-64-64-8 with ReLU activation in both hidden layers and eight linear outputs representing Q(s,a) for discrete combinations of AC, humidifier, and CO2 states: ALL_OFF, AC_ON, H_ON, CO2_ON, AC_H_ON, AC_CO2_ON, H_CO2_ON, and AC_H_CO2_ON. The policy had been trained previously in simulation. For the field evaluation analyzed here, the network weights were fixed; no online gradient update, replay-buffer learning, or policy retraining occurred. Accordingly, this study evaluates fixed-policy deployment behavior rather than training convergence or reward optimality.

Operational constraints reduced the action set before greedy selection. CO2-ON actions were invalid when the lamp was OFF or when indoor CO2 was at least 1000 ppm. Humidifier-ON actions were invalid when RH was at least 80%. When the lamp was OFF, AC-ON actions were invalid at indoor temperature at or below 22.5 °C, while an emergency guard required an AC-ON action at temperatures of at least 28 °C. AC restart protection also required a minimum OFF duration of three DQN steps (30 min), except when temperature reached 24.8 °C. Duration-related CO2 gas-delivery claims are not used in this paper because the available status telemetry and the absence of a flow meter do not provide an independent measurement of delivered gas quantity. The final DQN action was selected as the highest-Q action among those remaining valid, as expressed in Equation (1).

$$a_t = arg\,max_{a \in A_{valid}(s_t)} Q(s_t, a) \quad (1)$$

where $A_{valid}(s_t)$ denotes the action subset remaining after the deployment constraints are evaluated for the current state.

**Table 3** ***Controller implementation summary***

| Item | DQN | Rule-Based |
|---|---|---|
| Decision interval | 600 s (10 min) | 300 s (5 min) |
| Input | 11-state vector | Current sensor values and actuator states |
| Action representation | 8 AC/H/CO2 combinations | Equivalent AC/H/CO2 final-state combinations |
| Decision logic | Highest Q-value among valid actions | Fixed hysteresis thresholds |
| CO2 at lamp OFF | CO2-ON actions masked | Rule forces CO2 OFF |
| Auxiliary constraints | Hard action constraints + AC restart protection | Hysteresis logic; no DQN safety layer |
| Online learning during test | No | Not applicable |

The comparison therefore represents the two complete deployed configurations. Differences observed in the field can reflect decision logic, update frequency, auxiliary operational constraints, outdoor conditions, and equipment behavior; they are not interpreted as isolated effects of DQN inference versus hysteresis alone.

## 2.3 Experimental Data and Comparison Window

Existing field records were used without additional experiments. The DQN comparison window was 4 July 2026 08:20 to 5 July 2026 08:20, and the rule-based window was 20 June 2026 08:16 to 21 June 2026 08:16. The equal-duration windows contain 720 valid DQN samples and 716 valid rule-based samples for temperature and RH at approximately 2-minute telemetry intervals. CO2 records were available at approximately 5-minute intervals, producing 288 samples per configuration; 192 samples per configuration occurred while the lamp was ON. These telemetry sampling intervals are distinct from the 10-minute DQN and 5-minute rule-based decision intervals. Power and actuator-status logs were analyzed over the same 24-hour boundaries. Missing environmental samples were not imputed for compliance or MAE calculations.

## 2.4 Evaluation Metrics

For variable X, target-band compliance was the percentage of valid samples within the inclusive evaluation limits. Midpoint MAE was calculated relative to 22 °C for temperature, 65% for RH, and 900 ppm for CO2; it therefore measures average deviation from the center of each evaluation band rather than tracking error to a time-varying reference. Upper and lower excursions were calculated from the observed extrema relative to the corresponding band limits. Measured electrical energy was calculated by trapezoidal integration of the recorded power signal. Controller-dependent electrical energy was defined as AC + humidifier + CO2-solenoid electrical energy; lamp energy is reported separately because the photoperiod was common to both windows. CO2-solenoid electrical energy and ON-state telemetry do not quantify injected gas volume or mass. Actuator ON-state duration and observed state transitions were derived from exported telemetry; AC ON-state was derived from the recorded status/current signal rather than from controller-command counts. In Equation (4), P_i is measured electrical power in watts and delta t_i is the elapsed time between consecutive power samples in hours, so E is expressed in kWh. Because each configuration is represented by one continuous, serially correlated 24-hour deployment from a single physical installation, the telemetry samples are not independent experimental replicates and no independent-sample significance test is used. The sample count reflects temporal resolution within each window rather than experimental replication; therefore, the reported compliance, MAE, excursion, energy, and actuator differences are descriptive between-window observations rather than statistically generalizable treatment effects.

$$Compliance_X = \frac{N_{in-band,X}}{N_{valid,X}} \times 100\% \quad (2)$$

$$MAE_X = \frac{1}{N_X} \sum_{i=1}^{N_X} \left| X_i - X_{ref,X} \right| \quad (3)$$

$$E = \frac{1}{1000} \sum_{i=1}^{N-1} \frac{P_i + P_{i+1}}{2} \Delta t_i \quad (4)$$

# 3. Results and Discussion

## 3.1 Outdoor Conditions and Data Coverage

The two 24-hour trials were not simultaneous. During the DQN trial, outdoor temperature averaged 27.43 ± 4.45 °C, compared with 25.15 ± 4.12 °C during the rule-based trial. Outdoor relative humidity averaged 83.23 ± 17.00% for DQN and 79.27 ± 14.04% for rule-based control. Thus, the DQN window was on average 2.28 °C warmer and 3.96 percentage points more humid outdoors. These differences are important when interpreting AC energy and indoor RH; the results describe observed field performance rather than a controlled A/B test.

**Table 4** *Data coverage and outdoor conditions*

| Metric | DQN | Rule-Based |
|---|---|---|
| Evaluation window | 4 Jul 08:20 - 5 Jul 08:20 | 20 Jun 08:16 - 21 Jun 08:16 |
| Temperature/RH samples | 720 / 720 | 716 / 716 |
| CO2 samples (lamp on) | 192 | 192 |
| Outdoor T, mean ± SD (°C) | 27.43 ± 4.45 | 25.15 ± 4.12 |
| Outdoor RH, mean ± SD (%) | 83.23 ± 17.00 | 79.27 ± 14.04 |

## 3.2 Temperature Control Performance

Figure 2 shows the indoor-temperature trajectories. The DQN window contained 79.86% of valid samples within the 20-24 °C band, compared with 77.37% for the rule-based window, a difference of 2.49 percentage

points. Mean temperature was 22.48 °C for DQN and 22.09 °C for rule-based control. Midpoint MAE was 1.10 °C for DQN versus 1.28 °C for rule-based control, approximately 14% lower. The extreme behavior was mixed: DQN reached 26.7 °C, corresponding to a 2.7 °C upper excursion, while rule-based control reached 25.2 °C (1.2 °C upper excursion). Conversely, the DQN minimum was 18.4 °C (1.6 °C lower excursion) compared with 17.3 °C (2.7 °C lower excursion) for the rule-based window. Thus, the DQN deployment showed modestly better average regulation but not uniformly better containment of extremes.

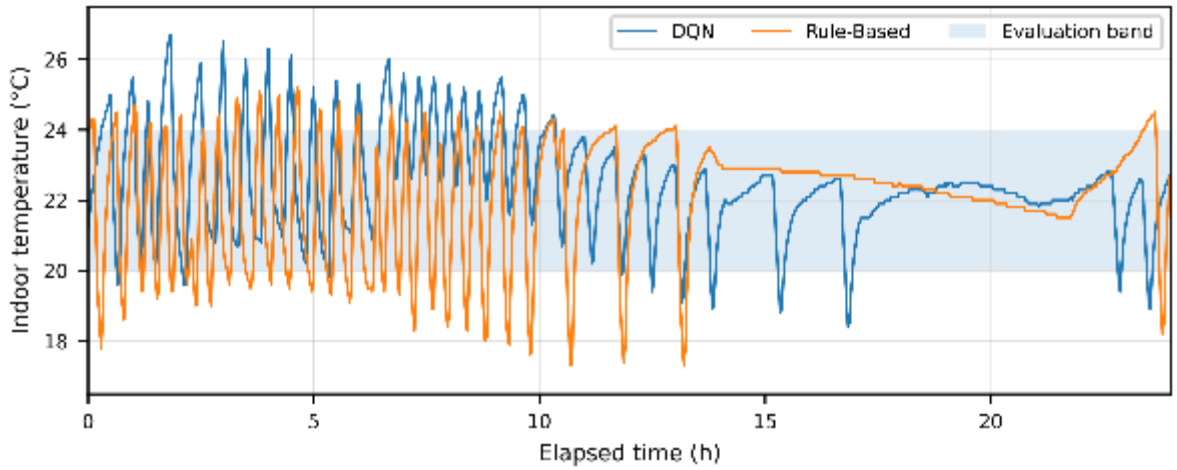



**Fig. 2** ***Indoor temperature over equal-duration 24-hour windows. The shaded region is the 20-24 °C evaluation band. Source: experimental data processed by the authors.***

### 3.3 Relative Humidity Performance

Relative humidity was the least well-controlled variable in both windows (Figure 3). Rule-based control recorded 44.97% in-band samples, compared with 40.69% for DQN. However, the DQN RH distribution was closer to the 65% midpoint: midpoint MAE was 7.89 percentage points for DQN and 10.58 percentage points for rule-based control, a 25.47% lower observed value. The maximum upper excursion was also lower in the DQN window (15.5 versus 19.4 percentage points). Mean RH was 68.54% for DQN and 74.85% for rule-based control. The humidifier was never ON in either 24-hour window, consistent with an environment dominated by excessive rather than deficient humidity. Since the platform did not include a dedicated dehumidifier, neither controller could directly remove moisture. The contrast between compliance and midpoint MAE also illustrates why a single metric is insufficient: the rule-based window entered the band more often, while the DQN distribution was closer to its midpoint overall.

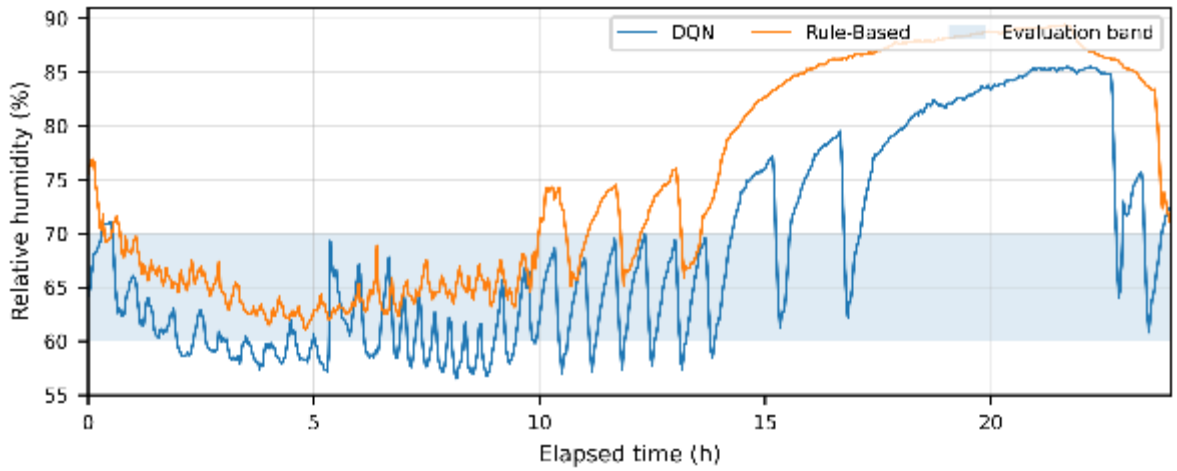



**Fig. 3** ***Indoor relative humidity over equal-duration 24-hour windows. The shaded region is the 60-70% evaluation band. Source: experimental data processed by the authors.***

### 3.4 CO2 Control Performance

CO2 behavior differed substantially between the two deployment windows (Figure 4). During lamp-on periods, the DQN window recorded 45.31% compliance with the 800-1000 ppm band, compared with 18.23% for rule-based control. Midpoint CO2 MAE was 119.31 ppm for DQN and 221.69 ppm for rule-based control, a 46.18% lower observed value in the DQN window. The rule-based lamp-on data remained predominantly below the target band, with a minimum of 537 ppm and a maximum of 869 ppm. The DQN lamp-on mean was 835.09 ppm compared with 678.31 ppm for rule-based control, but the DQN window also reached 1240 ppm, or 240 ppm above the upper evaluation limit. Thus, stronger average tracking in the DQN window was accompanied by a larger upper excursion. Because injection flow was not measured, concentration performance cannot be converted into a comparison of CO2 gas consumption.

The combination of better average tracking and a larger upper excursion is consistent with the controller design. The training reward balanced climate deviation against nominal-power, switching, and unsafe-state penalties rather than minimizing CO2 overshoot alone, and the documented reward did not use a dedicated asymmetric overshoot term. In deployment, decisions were updated every 10 min and CO2-ON actions were blocked once the measured concentration reached 1000 ppm or when the lamp was OFF. This is a reactive sampled-data guard rather than a predictive constraint: after an injection command, CO2 can continue to mix and rise before the next decision. The deployment log illustrates this behavior: CO2 was 799 ppm at 08:22 with a CO2-ON action selected, then reached 1045 ppm at the next decision at 08:32, when CO2-ON actions were masked. Therefore, an upper excursion can occur even though further injection is blocked after the threshold is observed. The available data cannot isolate the relative contributions of gas mixing, sensor response, simulator-to-field mismatch, and other operational differences, so this explanation is treated as a design-consistent mechanism rather than a causal attribution.

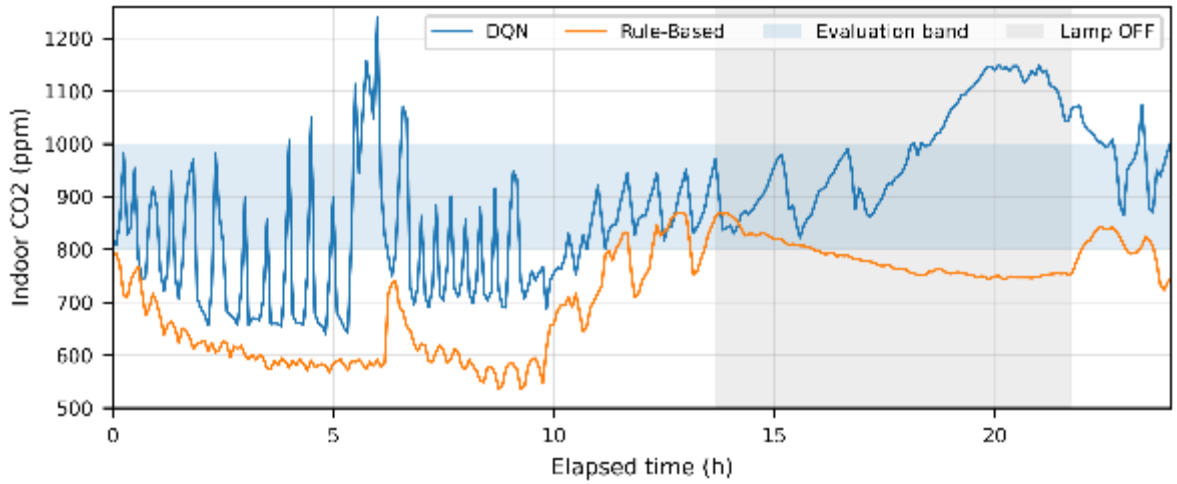



**Fig. 4** *Indoor CO2 concentration over equal-duration 24-hour windows. The 800-1000 ppm evaluation band is shaded blue; light gray background periods denote lamp OFF. Metrics in Table 5 use lamp-on samples. Source: experimental data processed by the authors.*

**Table 5** *Climate-control performance*

| Variable | DQN: compliance; midpoint MAE; range | Rule-Based: compliance; midpoint MAE; range |
|---|---|---|
| Temperature | 79.86%; 1.10 °C; 18.4-26.7 °C | 77.37%; 1.28 °C; 17.3-25.2 °C |
| Relative humidity | 40.69%; 7.89 pp; 56.5-85.5% | 44.97%; 10.58 pp; 61.1-89.4% |
| CO2 (lamp on) | 45.31%; 119.31 ppm; 640-1240 ppm | 18.23%; 221.69 ppm; 537-869 ppm |

## 3.5 Electrical Energy Consumption and Actuator Operation

Figure 5 and Table 6 compare measured electrical energy and actuator telemetry. The DQN window recorded 4.497 kWh in the AC channel versus 3.777 kWh in the rule-based window. Humidifier electrical energy was 0.000 kWh in both windows, CO2-solenoid electrical energy was approximately 0.011 kWh in both, and lamp electrical energy was nearly identical (3.688 versus 3.691 kWh) under the common 16-hour photoperiod. Controller-dependent electrical energy (AC + humidifier + CO2 solenoid) was 4.508 kWh for DQN and 3.788 kWh for rule-based control, a 19.01% higher observed value in the DQN window. Four-actuator total electrical energy was 8.197 versus 7.479 kWh, a 9.60% difference. The additional electrical load in the DQN window was dominated by AC operation, but the warmer DQN outdoor conditions prevent attributing this difference solely to controller logic. The DQN reward included a nominal-power penalty together with climate, switching, and safety terms, so the policy was energy-aware rather than an energy-minimization controller. Accordingly, the 19.01% difference is a between-window observation and should not be interpreted as evidence that DQN is intrinsically less energy efficient than rule-based control.

The AC was observed ON for 6.50 h under DQN and 5.70 h under rule-based control. DQN also produced 18 observed CO2-status transitions compared with 6 for rule-based control. These telemetry counts describe the deployed configurations, not intrinsic switching rates, because the decision intervals differed. For CO2, the exported ON-state duration and solenoid electrical energy are actuator-status/electrical indicators only; without a flow meter they do not measure injected gas quantity.

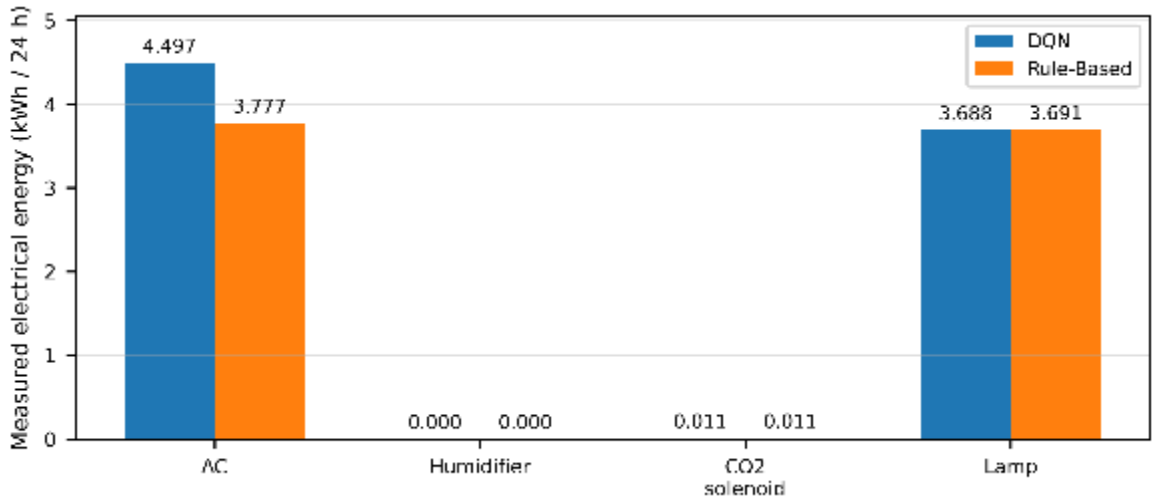



**Fig. 5** ***Measured electrical energy by shared actuator channel during each 24-hour window. Numeric labels show integrated kWh. Source: experimental power telemetry processed by the authors.***

**Table 6** ***Measured electrical energy and actuator operation***

| Metric | DQN | Rule-Based |
|---|---|---|
| AC | 4.497 kWh; 6.50 h; 62 transitions | 3.777 kWh; 5.70 h; 58 transitions |
| Humidifier | 0.000 kWh; 0.00 h; 0 transitions | 0.000 kWh; 0.00 h; 0 transitions |
| CO2 solenoid/status telemetry | 0.011 kWh; 13.87 h; 18 transitions | 0.011 kWh; 13.67 h; 6 transitions |
| Lamp | 3.688 kWh; 16.00 h; 2 transitions | 3.691 kWh; 16.00 h; 2 transitions |
| Controller-dependent electrical energy | 4.508 kWh | 3.788 kWh |
| Four-actuator total electrical energy | 8.197 kWh | 7.479 kWh |

*Note: CO2 ON-state duration is derived from exported status telemetry and must not be interpreted as gas volume or mass.*

**Table 7** ***Summary of observed deployment trade-offs***

| Aspect | Observed comparison | Interpretation |
|---|---|---|
| Temperature occupancy | 79.86% DQN vs 77.37% rule-based | Modest DQN-window advantage |
| Temperature deviation | 1.10 vs 1.28 °C midpoint MAE | DQN window closer to band midpoint |
| Temperature extremes | DQN higher upper excursion; rule-based higher lower excursion | Neither configuration dominated both extremes |
| Relative humidity | Rule-based higher compliance; DQN lower midpoint MAE | Mixed result, strongly limited by no dehumidifier |
| CO2 tracking | DQN higher compliance and lower midpoint MAE | Stronger DQN-window tracking during lamp-on periods |
| CO2 excursion | DQN reached 1240 ppm; rule-based remained below 1000 ppm | Better average tracking accompanied by overshoot in DQN window |
| Electrical energy | 4.508 vs 3.788 kWh controller-dependent energy | Lower observed electrical use in rule-based window; not an isolated algorithm effect because weather/configuration differed |

### 3.6 Overall Trade-off and Limitations

Taken together, the field records do not support a single statement that one controller was uniformly superior. The DQN window showed modestly higher temperature-band occupancy and substantially better lamp-on CO2 tracking, but it also showed a larger CO2 upper excursion and higher controller-dependent electrical energy. The rule-based window used less observed electrical energy and produced fewer CO2-status transitions, while RH results remained mixed because neither configuration had active dehumidification. These observations therefore represent a multi-objective trade-off between climate regulation, excursion containment, electrical energy, and actuator behavior rather than a single-controller advantage.

Generalization is limited by the experimental design. The two configurations were deployed on different calendar days under unmatched outdoor conditions, used different decision intervals (10 min for DQN and 5 min for rule-based control), and did not share identical supervisory constraints. Each configuration was represented by only one continuous 24-hour window, so the telemetry samples are serially correlated observations rather than independent experimental replicates. Measurement and model limitations include the absence of direct CO2 flow measurement, training of the DQN policy in a simplified stochastic simulator whose fidelity was not independently validated in this field comparison, and the absence of active dehumidification. Future work should use repeated or randomized deployments under comparable conditions, a common control interval and supervisory-constraint set, direct CO2 flow measurement, longer observation periods, and bidirectional humidity control.

## 4. Conclusion

This experimental case study compared a fixed DQN deployment with a hysteresis rule-based deployment using equal-duration 24-hour records from the same indoor-farming platform. The DQN window recorded 79.86% temperature compliance and a midpoint MAE of 1.10 °C, compared with 77.37% and 1.28 °C in the rule-based window. During lamp-on periods, CO2 compliance was 45.31% for DQN and 18.23% for rule-based control, while CO2 midpoint MAE was 119.31 versus 221.69 ppm; however, the DQN window also produced an upper excursion to 1240 ppm. RH performance remained limited because the platform lacked active dehumidification. Controller-dependent electrical energy was 4.508 kWh in the DQN window and 3.788 kWh in the rule-based window, with the difference dominated by AC operation. Actual CO2 gas consumption could not be compared because no flow meter was installed. These results answer RQ1 by showing variable-specific differences rather than uniform improvement, and RQ2 by demonstrating an observed climate-electrical-energy-actuation trade-off. They do not establish that DQN is intrinsically more or less energy efficient or generally superior to the rule-based algorithm. Because the deployments differed in date, outdoor conditions, decision interval, and auxiliary operational constraints, the reported values should be interpreted as configuration-level field observations rather than causal algorithm effects. The practical implication is that learning-based climate-control systems should be evaluated with multiple climate, excursion, energy, and actuator metrics under matched and explicitly documented operating conditions.

## Acknowledgement

The authors would like to thank Institut Teknologi Del for providing access to the indoor farming experimental platform and supporting facilities used to obtain the deployment records analyzed in this study.

## Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Author Contribution

The authors confirm contribution to the paper as follows: study conception and design: Raj David Francisco Damanik, Samuel Abednego Sormin; data collection: Samuel Abednego Sormin; analysis and interpretation of results: Raj David Francisco Damanik, Samuel Abednego Sormin, Albert Sagala; draft manuscript

preparation: Raj David Francisco Damanik; supervision and critical review: Albert Sagala. All authors reviewed the results and approved the final version of the manuscript.

## References

[1] Vatistas, C., Avgoustaki, D. D., & Bartzanas, T. (2022). A systematic literature review on controlled-environment agriculture: How vertical farms and greenhouses can influence the sustainability and footprint of urban microclimate with local food production. Atmosphere, 13(8), 1258. https://doi.org/10.3390/atmos13081258

[2] Shamshiri, R. R., Kalantari, F., Ting, K. C., Thorp, K. R., Hameed, I. A., Weltzien, C., Ahmad, D., & Shad, Z. M. (2018). Advances in greenhouse automation and controlled environment agriculture: A transition to plant factories and urban agriculture. International Journal of Agricultural and Biological Engineering, 11(1), 1-22. https://doi.org/10.25165/j.ijabe.20181101.3210

[3] Wang, L., Xiao, M., Guo, X., Yang, Y., Zhang, Z., & Lee, C. (2024). Sensing technologies for outdoor/indoor farming. Biosensors, 14(12), 629. https://doi.org/10.3390/bios14120629

[4] Graamans, L., Baeza, E., van den Dobbelsteen, A., Tsafaras, I., & Stanghellini, C. (2018). Plant factories versus greenhouses: Comparison of resource use efficiency. Agricultural Systems, 160, 31-43. https://doi.org/10.1016/j.agsy.2017.11.003

[5] Cai, W., Bu, K., Zha, L., Zhang, J., Lai, D., & Bao, H. (2025). Energy consumption of plant factory with artificial light: Challenges and opportunities. Renewable and Sustainable Energy Reviews, 210, 115235. https://doi.org/10.1016/j.rser.2024.115235

[6] Ghiasi, M., Wang, Z., Mehrandezh, M., & Paranjape, R. (2024). A systematic review of optimal and practical methods in design, construction, control, energy management and operation of smart greenhouses. IEEE Access, 12, 2830-2853. https://doi.org/10.1109/ACCESS.2023.3346436

[7] Chen, S., Liu, A., Tang, F., Hou, P., Lu, Y., & Yuan, P. (2025). A review of environmental control strategies and models for modern agricultural greenhouses. Sensors, 25(5), 1388. https://doi.org/10.3390/s25051388

[8] Ojo, M. O., & Zahid, A. (2022). Deep learning in controlled environment agriculture: A review of recent advancements, challenges and prospects. Sensors, 22(20), 7965. https://doi.org/10.3390/s22207965

[9] Morales-García, J., Bueno-Crespo, A., Terroso-Sáenz, F., Arcas-Túnez, F., Martínez-España, R., & Cecilia, J. M. (2023). Evaluation of synthetic data generation for intelligent climate control in greenhouses. Applied Intelligence, 53(21), 24765-24781. https://doi.org/10.1007/s10489-023-04783-2

[10] Wang, Y., Chen, Z., Jiang, Y., & Liu, T. (2023). High-order neural-network-based multi-model nonlinear adaptive decoupling control for microclimate environment of plant factory. Sensors, 23(19), 8323. https://doi.org/10.3390/s23198323

[11] Hu, G., & You, F. (2023). Energy management for plant factory with deep learning and predictive control. Chemical Engineering Transactions, 103, 91-96. https://doi.org/10.3303/CET23103016

[12] Chen, W.-H., & You, F. (2024). Decarbonization through smart energy management: Climate control in building-integrated rooftop greenhouses for urban agriculture across various climate conditions. Journal of Cleaner Production, 458, 142544. https://doi.org/10.1016/j.jclepro.2024.142544

[13] Jeon, Y.-J., Kim, J. Y., Hwang, K.-S., Cho, W.-J., Kim, H.-J., & Jung, D.-H. (2024). Machine learning-powered forecasting of climate conditions in smart greenhouse containing netted melons. Agronomy, 14(5), 1070. https://doi.org/10.3390/agronomy14051070

[14] Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. Machine Learning, 8, 279-292. https://doi.org/10.1007/BF00992698

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

[16] Ajagekar, A., Mattson, N. S., & You, F. (2023). Energy-efficient AI-based control of semi-closed greenhouses leveraging robust optimization in deep reinforcement learning. Advances in Applied Energy, 9, 100119. https://doi.org/10.1016/j.adapen.2022.100119
[17] Morcego, B., Yin, W., Boersma, S., van Henten, E., Puig, V., & Sun, C. (2023). Reinforcement learning versus model predictive control on greenhouse climate control. Computers and Electronics in Agriculture, 215, 108372. https://doi.org/10.1016/j.compag.2023.108372
[18] Adesanya, M. A., Obasekore, H., Rabiu, A., Na, W.-H., Ogunlowo, Q. O., Akpenpuun, T. D., Kim, M.-H., Kim, H.-T., Kang, B.-Y., & Lee, H.-W. (2024). Deep reinforcement learning for PID parameter tuning in greenhouse HVAC system energy optimization: A TRNSYS-Python cosimulation approach. Expert Systems with Applications, 252, 124126. https://doi.org/10.1016/j.eswa.2024.124126
[19] Platero-Horcajadas, M., Pardo-Pina, S., Cámara-Zapata, J.-M., Brenes-Carranza, J.-A., & Ferrández-Pastor, F.-J. (2024). Enhancing greenhouse efficiency: Integrating IoT and reinforcement learning for optimized climate control. Sensors, 24(24), 8109. https://doi.org/10.3390/s24248109
[20] Mallick, S., Airaldi, F., Dabiri, A., Sun, C., & De Schutter, B. (2025). Reinforcement learning-based model predictive control for greenhouse climate control. Smart Agricultural Technology, 10, 100751. https://doi.org/10.1016/j.atech.2024.100751
[21] Yu, J., Sun, C., Zhao, J., Ma, L., Zheng, W., Xie, Q., & Wei, X. (2025). Prediction and control of greenhouse temperature: Methods, applications, and future directions. Computers and Electronics in Agriculture, 237, 110603. https://doi.org/10.1016/j.compag.2025.110603
[22] Xiao, M., Lan, J., Yu, J., Ma, W., Xie, Q., & Sun, C. (2026). Grower-in-the-loop interactive reinforcement learning for greenhouse climate control. Computers and Electronics in Agriculture, 242, 111312. https://doi.org/10.1016/j.compag.2025.111312