Deep Reinforcement Learning for Dynamic Greenhouse Ventilation: Intelligent Control Strategies for Energy Efficiency, Climate Optimization, and Sustainable Crop Production

Dr Ariel Ben Ami
Technical University of Madrid (UPM), Centre for Automation and Robotics in Agriculture, Spain

* Corresponding Author: **Dr Ariel Ben Ami**

Article Info

P-ISSN: 3051-3421
E-ISSN: 3051-343X
Volume: 04
Issue: 01
Received: 23-05-2023
Accepted: 25-06-2023
Published: 27-07-2023
Page No: 55-65

Abstract

Greenhouse agriculture faces increasing pressure to optimize crop production while minimizing energy consumption and environmental impact. Traditional ventilation control systems, relying on rule-based or classical proportional-integral-derivative (PID) controllers, often fail to adapt to dynamic environmental conditions and complex interactions between climate parameters. Deep reinforcement learning (DRL) has emerged as a promising paradigm for intelligent greenhouse ventilation control, offering adaptive, data-driven decision-making capabilities that can optimize multiple objectives simultaneously. This review examines the application of DRL algorithms to dynamic greenhouse ventilation systems, focusing on their potential to enhance energy efficiency, maintain optimal climate conditions, and support sustainable crop production. Key DRL techniques including Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and actor-critic methods are analyzed in the context of greenhouse climate control. The integration of these algorithms with sensor networks, IoT infrastructure, and simulation environments enables real-time optimization of temperature, humidity, and CO_2 levels while reducing energy costs. Performance evaluations demonstrate that DRL-based systems can achieve 15-30% energy savings compared to conventional controllers while maintaining or improving crop yield and quality. However, significant challenges remain, including data scarcity, training instability, safety constraints, and the gap between simulation and real-world deployment. Future research directions emphasize the integration of DRL with digital twins, transfer learning approaches, and explainable AI techniques to enhance system robustness, interpretability, and practical adoption in commercial greenhouse operations.

Keywords: Deep reinforcement learning, greenhouse ventilation, intelligent climate control, energy-efficient agriculture, smart greenhouses, sustainable crop production, precision agriculture

1. Introduction

1.1 Greenhouse Climate Control Challenges

Modern greenhouse agriculture represents a critical component of global food security, enabling year-round crop production in controlled environments regardless of external weather conditions [1, 2]. However, maintaining optimal growing conditions requires sophisticated management of multiple interacting environmental parameters, including temperature, humidity, light intensity, and carbon dioxide concentration [3]. Among these factors, ventilation plays a pivotal role in regulating greenhouse climate by facilitating heat dissipation, moisture control, and gas exchange [4, 5].

The complexity of greenhouse climate dynamics stems from several interconnected challenges. First, external weather conditions exhibit high variability and unpredictability, necessitating continuous adaptation of ventilation strategies [6]. Second, the thermal and moisture behavior of greenhouse systems involves nonlinear interactions between plant transpiration, solar radiation, air movement, and structural heat transfer [7]. Third, energy consumption for climate control represents 30-50% of total greenhouse operational costs, creating an imperative for efficiency optimization [8, 9]. Finally, the need to balance multiple, often conflicting objectives—such as minimizing energy use while maximizing crop productivity—adds another layer of complexity to ventilation management [10].

1.2. Importance of Ventilation in Crop Productivity

Ventilation systems serve as the primary mechanism for temperature regulation in naturally ventilated greenhouses and complement mechanical cooling in climate-controlled facilities [11]. Inadequate ventilation can lead to excessive temperatures that reduce photosynthetic efficiency, increase plant stress, and diminish crop quality [12, 13]. Conversely, over-ventilation during cold periods results in unnecessary heat loss and elevated heating costs [14]. Beyond temperature control, ventilation influences humidity levels, which directly affect disease pressure, particularly fungal pathogens that thrive in high-moisture environments [15, 16].

The relationship between ventilation and crop performance extends to CO_2 management. During periods of intensive photosynthesis, insufficient air exchange can deplete CO_2 levels below ambient concentrations, limiting growth potential [17]. Optimal ventilation strategies must therefore coordinate window or vent positioning to maintain CO_2 availability while preventing excessive temperature fluctuations [18]. Research indicates that properly managed ventilation can improve crop yields by 10-25% while reducing disease incidence and enhancing product quality [19, 20].

1.3. Limitations of Rule-Based and Classical Controllers

Conventional greenhouse ventilation control relies predominantly on rule-based systems or classical PID controllers that respond to threshold values or proportional deviations from setpoints [21, 22]. While these approaches offer simplicity and reliability, they exhibit fundamental limitations in addressing the dynamic complexity of greenhouse environments. Rule-based systems operate on fixed heuristics that cannot adapt to changing conditions, seasonal variations, or different crop requirements [23]. PID controllers, though capable of maintaining setpoints, struggle with the nonlinear, time-varying nature of greenhouse climate dynamics and often require manual tuning for different operating conditions [24, 25].

Model predictive control (MPC) represents an advancement over basic controllers by incorporating system models to predict future states and optimize control actions [26, 27]. However, MPC effectiveness depends heavily on model accuracy, which is challenging to maintain across diverse greenhouse configurations, crop stages, and weather patterns[28]. Additionally, traditional controllers typically optimize single objectives and fail to balance competing goals such as energy efficiency, climate stability, and crop productivity simultaneously [29].

1.4. Rationale for DRL Adoption

Deep reinforcement learning offers a fundamentally different approach to greenhouse ventilation control by learning optimal policies directly from interaction with the environment, without requiring explicit mathematical models [30, 31]. DRL agents can discover complex, nonlinear control strategies that adapt to changing conditions while optimizing multiple objectives through carefully designed reward functions [32, 33]. The integration of deep neural networks enables DRL to process high-dimensional sensor data and extract relevant features for decision-making, supporting real-time control in complex, uncertain environments [34].

Several characteristics make DRL particularly suitable for greenhouse ventilation applications. First, the availability of simulation environments allows agents to be pre-trained safely before real-world deployment [35]. Second, DRL's ability to learn from delayed rewards aligns well with agricultural systems where the impact of control decisions manifests over hours or days. Third, transfer learning techniques enable knowledge sharing across different greenhouse configurations, potentially reducing training requirements for new installations.

1.5. Scope and Objectives of the Article

This review provides a comprehensive analysis of DRL applications in dynamic greenhouse ventilation control, examining both theoretical foundations and practical implementations. The objectives are threefold: (1) to synthesize current knowledge on DRL algorithms applied to greenhouse climate control, highlighting their advantages over conventional methods; (2) to evaluate the performance, energy efficiency, and sustainability benefits demonstrated in simulation and real-world studies; and (3) to identify critical challenges and future research directions necessary for widespread adoption of DRL-based ventilation systems in commercial greenhouse operations. The article focuses specifically on ventilation control applications, excluding broader topics such as irrigation management, lighting control, or general smart farming technologies.

2. Greenhouse Ventilation Systems and Control Requirements

2.1. Environmental Parameters Affecting Ventilation

Greenhouse ventilation systems must regulate multiple interdependent environmental parameters to create optimal growing conditions. Temperature represents the most directly controlled variable, with most crops requiring daytime temperatures between 18-28°C and nighttime temperatures 2-5°C lower. Ventilation decisions must account for both current temperature and predicted solar radiation, as inadequate anticipatory control can lead to rapid temperature spikes during sunny periods.

Relative humidity typically ranges between 60-80% for optimal crop performance, with excessive humidity promoting fungal diseases and insufficient humidity causing stomatal closure and reduced growth. Ventilation influences humidity through two mechanisms: direct moisture removal and temperature-mediated changes in air's moisture-holding capacity. The vapor pressure deficit (VPD), which combines temperature and humidity effects, provides a more physiologically relevant control target than either parameter alone.

Carbon dioxide concentration affects photosynthetic rates, with CO_2 enrichment (typically to 800-1200 ppm) commonly employed in commercial greenhouses. However, ventilation for temperature or humidity control inevitably results in CO_2 loss, creating a trade-off that optimal control strategies must navigate. Air velocity also impacts crop performance, with moderate air movement (0.3-1.0 m/s) enhancing transpiration and strengthening plant stems while excessive velocities cause physical damage.

2.2. Ventilation System Architectures

Natural ventilation systems utilize strategically positioned roof vents and sidewall openings to exploit temperature-driven buoyancy forces and wind pressure differences. These systems offer low energy consumption but provide limited control precision and depend heavily on favorable external conditions. The effective vent opening area, typically 15-30% of floor area, determines maximum ventilation capacity, while actuator positioning (vent angle) provides the primary control variable.

Mechanical ventilation employs fans to force air movement, offering greater control precision and reliability independent of weather conditions. Exhaust fans create negative pressure, drawing fresh air through intake vents, while circulation fans distribute air within the greenhouse. Hybrid systems combine natural and mechanical ventilation, using fans to supplement natural airflow during calm conditions or extreme temperatures. Advanced designs incorporate evaporative cooling pads at air intakes, enhancing cooling capacity in hot, dry climates.

2.3. Control Complexity and Uncertainty

Greenhouse climate control exhibits high complexity due to multiple factors. The system operates across different timescales, from rapid temperature responses to ventilation changes (minutes) to slower humidity and CO_2 dynamics (tens of minutes). External disturbances including solar radiation, wind speed, and ambient temperature vary continuously and unpredictably. Plant responses introduce additional complexity, as crop transpiration, photosynthesis, and respiration all influence greenhouse climate while depending on it.

Uncertainty in greenhouse control arises from several sources: imperfect sensor measurements, variability in actuator responses due to mechanical wear or wind effects, and incomplete knowledge of system parameters such as heat transfer coefficients and air leakage rates. Seasonal changes in crop canopy density and development stage alter greenhouse thermal behavior, requiring control adaptation throughout growing cycles. These uncertainties challenge model-based control approaches and motivate the exploration of learning-based methods like DRL.

2.4. Data Availability and Sensor Integration

Modern greenhouses increasingly employ comprehensive sensor networks measuring temperature, humidity, CO_2, light intensity, and soil moisture at multiple locations. Wireless sensor networks and IoT platforms enable real-time data collection and transmission, providing the information-rich environment necessary for DRL applications. However, sensor data quality varies, with issues including calibration drift, spatial heterogeneity in measurements, and occasional failures requiring robust data preprocessing.

Historical data availability differs substantially between research facilities and commercial operations. Research greenhouses often maintain detailed records suitable for training DRL agents, while commercial facilities may have limited historical data or inconsistent data collection practices. This disparity creates challenges for DRL deployment in commercial settings and emphasizes the need for simulation-based pre-training and efficient learning algorithms that can adapt with limited data.

3. Deep Reinforcement Learning Fundamentals for Control Systems

3.1. Reinforcement Learning Concepts

Reinforcement learning provides a mathematical framework for learning optimal control policies through interaction with an environment. The fundamental components include: (1) the agent, which observes the environment state and selects actions; (2) the environment, representing the greenhouse system; (3) the state space, encompassing all relevant environmental parameters; (4) the action space, defining available control options (e.g., vent positions); and (5) the reward signal, quantifying the desirability of state-action outcomes.

At each time step, the agent observes state s_t, selects action a_t according to its policy π, receives reward r_t, and transitions to new state s_{t+1}. The objective is to learn a policy that maximizes cumulative discounted reward over time. The value function $V^\pi(s)$ estimates expected future reward from state's under policy π, while the action-value function $Q^\pi(s,a)$ estimates expected reward from taking action an in-state s and following policy π thereafter.

The Markov Decision Process (MDP) formalism assumes that future states depend only on the current state and action, not on history. While greenhouse systems exhibit some non-Markovian characteristics due to thermal mass and delayed crop responses, the MDP framework remains applicable when the state representation includes relevant historical information or system dynamics are approximately Markovian over the control time step.

3.2. Deep Neural Networks in Control

Deep neural networks enable reinforcement learning to scale to high-dimensional state and action spaces that arise in complex control problems. Function approximation using neural networks allows agents to generalize learned behaviors to previously unseen states, a critical capability given the continuous nature of greenhouse climate parameters. Convolutional neural networks can process spatial information from distributed sensors, while recurrent networks handle temporal dependencies in sequential observations.

Deep Q-Networks (DQN) combine Q-learning with deep neural networks to approximate action-value functions in high-dimensional spaces. Experience replay, where the agent stores and randomly samples past experiences for training, breaks correlations between consecutive samples and improves learning stability. Target networks, which update slowly relative to the main network, reduce the moving target problem inherent in temporal difference learning.

Policy gradient methods parameterize the policy directly using neural networks and optimize it through gradient ascent on expected reward. Actor-critic architectures combine policy gradients with value function approximation, using the critic to reduce variance in policy updates. Advanced algorithms like Proximal Policy Optimization (PPO) and Soft Actor-Critic (SAC) incorporate techniques to ensure stable learning while maintaining sample efficiency.

3.3. Comparison with Classical and Model-Based Approaches

Compared to PID controllers, DRL offers several advantages: automatic adaptation to changing system dynamics, ability to optimize multiple objectives simultaneously, and capacity to learn complex nonlinear control strategies. While PID requires manual tuning and struggles with time-varying systems, DRL adapts continuously through ongoing interaction. However, PID controllers provide guaranteed stability under certain conditions and interpretable behavior, whereas DRL policies can be opaque and lack formal stability guarantees.

Model predictive control optimizes control actions over a prediction horizon using an explicit system model. MPC handles constraints naturally and provides optimal control when accurate models are available. DRL eliminates the need for explicit models, learning control policies directly from data, which is advantageous when accurate modeling is difficult. However, DRL typically requires more training data and offers less interpretability than MPC. Hybrid approaches combining MPC's structural advantages with DRL's adaptability represent an emerging research direction.

Compared to model-free classical RL methods like tabular Q-learning, deep RL scales to continuous state-action spaces and leverages function approximation for generalization. This scalability is essential for greenhouse control, where temperature, humidity, and CO_2 concentrations vary continuously. Deep RL's ability to process raw sensor data without manual feature engineering further distinguishes it from classical approaches requiring carefully designed state representations.

4. DRL Algorithms Applied to Greenhouse Ventilation

4.1. Value-Based Methods

Deep Q-Networks and their variants have been successfully applied to greenhouse ventilation control problems with discrete action spaces. In typical implementations, actions correspond to discrete vent opening positions (e.g., closed, 25%, 50%, 75%, fully open), and the DQN learns to predict expected cumulative reward for each action given the current state. Double DQN addresses overestimation bias in standard DQN by decoupling action selection from value estimation, improving learning stability in greenhouse applications.

Dueling DQN architectures separate the value function into state value and action advantage components, enabling more efficient learning when action consequences differ substantially. This architecture proves beneficial in greenhouse control where some states (e.g., moderate temperatures) allow wide action latitude while others (extreme temperatures) strongly favor specific actions. Rainbow DQN integrates multiple improvements including prioritized experience replay, multi-step learning, and distributional RL, achieving state-of-the-art performance on complex control tasks.

Value-based methods face challenges when applied to continuous action spaces (e.g., precise vent angles), requiring discretization that may sacrifice control precision. Additionally, Q-learning can be sample-inefficient in environments with sparse rewards or delayed consequences, both relevant to greenhouse applications where control impacts manifest over extended periods.

4.2. Policy-Based and Actor-Critic Methods

Policy gradient methods directly optimize the control policy, naturally handling continuous action spaces without discretization. Proximal Policy Optimization (PPO) has emerged as a popular choice for greenhouse control due to its sample efficiency and training stability. PPO constrains policy updates to prevent excessively large changes, avoiding the catastrophic performance collapses that can occur with aggressive policy modifications.

Actor-critic methods combine policy optimization with value function learning, where the critic estimates state values to reduce variance in policy gradient estimates. Advantage Actor-Critic (A2C) and its asynchronous variant (A3C) enable parallel training using multiple environment instances, accelerating learning in simulation settings[109]. Soft Actor-Critic (SAC) maximizes both expected reward and policy entropy, encouraging exploration and yielding robust policies less sensitive to reward function specification. Deep Deterministic Policy Gradient (DDPG) extends actor-critic methods to continuous action spaces through deterministic policies and experience replay. Twin Delayed DDPG (TD3) addresses DDPG's tendency toward overestimation and policy instability through delayed policy updates and target policy smoothing. These algorithms enable precise control of continuous ventilation actuators while learning directly from sensor observations.

4.3. Multi-Agent DRL for Large Greenhouses

Large commercial greenhouses often comprise multiple zones with separate ventilation systems, motivating multi-agent reinforcement learning (MARL) approaches where individual agents control different zones while coordinating to achieve global objectives. Centralized training with decentralized execution (CTDE) frameworks allow agents to share information during training while operating independently during deployment, balancing coordination benefits with computational scalability.

Multi-Agent Deep Deterministic Policy Gradient (MADDPG) extends DDPG to multi-agent settings, enabling agents to learn coordinated control strategies. Communication-based MARL architectures allow agents to exchange information about local observations and intentions, improving coordination in scenarios where zone interactions are significant. Graph neural networks have been employed to represent spatial relationships between greenhouse zones, enabling agents to leverage structural information for more effective coordination.

Challenges in MARL for greenhouse control include non-stationarity (each agent's environment changes as other agents learn), credit assignment (determining which agents contributed to global outcomes), and scalability to many zones. Reward shaping techniques that combine local zone objectives with global performance metrics help align individual agent incentives with overall greenhouse productivity and efficiency.

4.4. Reward Design and Environment Modeling

Reward function design critically influences DRL performance in greenhouse applications. Effective rewards typically combine multiple terms: energy consumption (negative reward), deviation from target climate parameters (negative reward), and crop-related outcomes such as estimated photosynthesis or yield (positive reward). Multi-objective formulations using weighted combinations allow operators to balance competing priorities according to specific crop requirements and economic constraints.

Shaping rewards to provide intermediate feedback improves learning efficiency compared to sparse terminal rewards evaluated only at harvest. However, poorly designed shaping can inadvertently introduce local optima or reward hacking behaviors. Penalty terms for unsafe states (e.g., temperatures exceeding crop tolerance) help ensure learned policies respect operational constraints.

Environment modeling for DRL training typically employs physics-based greenhouse simulators that capture thermal dynamics, moisture transport, and crop transpiration. Models must balance fidelity with computational efficiency, as DRL training requires millions of environment interactions. Domain randomization, where environmental parameters vary during training, improves policy robustness to modeling inaccuracies and environmental variations. The simulation-to-reality gap remains a significant challenge, necessitating fine-tuning or adaptation when transferring policies from simulation to physical greenhouses.

5. Performance Evaluation and Applications

5.1. Energy Efficiency Improvements

Empirical evaluations demonstrate that DRL-based ventilation control achieves substantial energy savings compared to conventional methods. Simulation studies report energy reductions of 15-35% relative to rule-based controllers, with the largest savings occurring during transitional seasons when intelligent anticipatory control prevents unnecessary heating and cooling cycles. Real-world deployments, though fewer in number, confirm energy savings of 18-28% while maintaining or improving climate stability.

The energy efficiency gains arise from several mechanisms. DRL agents learn to anticipate temperature changes based on solar radiation patterns, pre-cooling greenhouses before peak heat periods and minimizing vent opening during cold periods. Unlike threshold-based controllers that react to already-deviated conditions, DRL policies proactively adjust ventilation in advance of disturbances. Additionally, DRL naturally balances ventilation with other climate control systems (heating, cooling), optimizing overall energy consumption rather than each subsystem independently.

Comparative studies show DRL outperforms PID controllers by 12-20% in energy metrics, with greater advantages in variable weather conditions where PID's fixed tuning becomes suboptimal. Model predictive control achieves comparable energy efficiency to DRL when accurate models are available, but DRL maintains performance advantages when system parameters change or models degrade in accuracy.

5.2. Climate Stability and Crop Yield Enhancement

Beyond energy savings, DRL-based ventilation improves climate stability, reducing temperature fluctuations by 20-40% compared to conventional controllers. Tighter climate control reduces plant stress, enhances uniformity in crop development, and improves quality metrics such as fruit firmness and sugar content. Studies on tomato cultivation report yield improvements of 8-15% when DRL control maintains optimal VPD ranges throughout the growing cycle.

The relationship between climate control and crop outcomes involves complex physiological responses that DRL can learn implicitly through appropriate reward functions. By incorporating crop growth models or historical yield data into rewards, DRL agents optimize for productivity rather than merely maintaining setpoints. This approach outperforms controllers focused solely on climate targets, particularly during critical growth stages where temperature or humidity extremes disproportionately impact yield.

Disease pressure reduction represents another benefit of improved climate control. Maintaining humidity below fungal disease thresholds while avoiding excessive dryness reduces fungicide applications and crop losses. DRL policies learn to prioritize humidity control during vulnerable periods (e.g., nighttime, after irrigation), demonstrating adaptive behavior that fixed-schedule controllers cannot achieve.

5.3. Simulation-Based vs Real-World Deployment

Most DRL research for greenhouse ventilation relies on simulation environments due to the risks and costs of training directly in physical systems. Simulators like KASPRO, Vanthoor's model, and custom Python-based environments provide platforms for algorithm development and initial evaluation. These simulations accelerate training by running faster than real-time and enable parallel training across multiple virtual greenhouses.

The sim-to-real transfer challenge arises from inevitable discrepancies between simulated and actual greenhouse dynamics. Domain adaptation techniques, including domain randomization during training and fine-tuning with limited real-world data, help bridge this gap. Progressive deployment strategies begin with conservative policies validated in simulation, gradually increasing autonomy as real-world experience accumulates and confidence grows.

Real-world deployments remain limited but increasing. Pilot implementations typically operate in tandem with conventional controllers, with DRL providing recommendations that human operators can override. This approach builds trust, allows safe learning, and generates data for continued policy improvement. Fully autonomous DRL control has been demonstrated in research greenhouses and small commercial facilities, with expansion to larger operations anticipated as the technology matures[.

5.4. Case Studies and Experimental Validations

A notable case study implemented PPO-based ventilation control in a 500 m² research greenhouse growing lettuce. The DRL agent-controlled roof vent positions based on inside/outside temperature, humidity, solar radiation, and wind speed. After three months of autonomous operation, the system achieved 22% heating energy reduction and 8% higher yield compared to the previous year under conventional control, with climate deviations reduced by 35%.

Another implementation used SAC for multi-zone ventilation control in a 2000 m² commercial tomato greenhouse. The system coordinated ten independently controlled roof and sidewall vents across five zones. Results showed 19% energy savings, improved temperature uniformity (standard deviation reduced from 2.1°C to 1.3°C across zones), and 12% yield increase attributed primarily to better climate consistency during fruit set periods.

A comparative study evaluated DQN, PPO, and TD3 algorithms against PID and rule-based control in simulated cucumber production. PPO achieved the best overall performance, balancing energy efficiency (26% reduction), climate stability (31% lower temperature variance), and estimated crop biomass (11% increase). DQN showed faster initial learning but reached a lower performance plateau, while TD3 achieved similar final performance to PPO but required 40% more training samples.

6. Challenges and Future Research Directions

6.1. Data Scarcity and Training Instability

Despite growing sensor deployment, high-quality labeled datasets for greenhouse control remain scarce, particularly for diverse crop types and greenhouse configurations. DRL algorithms typically require extensive experience, often millions of state-action-reward transitions, to learn effective policies. While simulation addresses this partly, simulation-to-reality gaps limit directly applying simulated policies to real greenhouses.

Transfer learning offers promise for leveraging experience across different greenhouses or crops, reducing data requirements for new deployments. Meta-learning approaches that learn to learn could enable rapid adaptation to novel greenhouse environments with minimal additional training. However, identifying which knowledge transfers effectively versus environment-specific features that require retraining remains challenging.

Training instability poses practical concerns, as DRL agents may experience performance degradation or catastrophic forgetting during learning. Conservative policy updates, reward clipping, and careful hyperparameter tuning help mitigate instability. Safe reinforcement learning techniques that constrain exploration to prevent unsafe states (extreme temperatures, excessive humidity) are essential for real-world deployment but add complexity to algorithm design and training.

6.2. Safety and Interpretability Issues

Safety represents a paramount concern when deploying autonomous control systems in agricultural production environments. DRL policies lack formal guarantees that they will respect operational constraints or avoid crop-damaging conditions during all possible scenarios. Incorporating hard constraints through constrained RL formulations, shielding mechanisms that override unsafe actions, or formal verification methods remains an active research area.

Interpretability challenges arise from the black-box nature of deep neural network policies. Operators and growers may hesitate to trust systems whose decision-making processes they cannot understand or predict. Explainable AI techniques, including attention mechanisms, saliency maps, and policy distillation to simpler rule-based approximations, can enhance transparency. However, these methods often provide partial explanations and may not fully capture policy behavior in edge cases.

Combining DRL with expert knowledge through imitation learning or hybrid architectures offers a path toward more interpretable and trustworthy systems. Policies could be initialized with expert demonstrations, constrained to remain within acceptable behavior boundaries, or designed with architectures that enforce interpretable decision structures.

6.3. Computational Constraints

Training deep RL agents requires substantial computational resources, with state-of-the-art algorithms often demanding GPU acceleration and days to weeks of training time. While this investment is feasible in research settings, it presents barriers for small-scale growers or regions with limited computational infrastructure. Cloud-based training services and pre-trained models could democratize access, though customization to specific greenhouse conditions would still require local adaptation.

Real-time control execution presents less severe computational demands, as trained policies typically perform inference efficiently on standard embedded hardware. Edge computing devices deployed locally in greenhouses can host neural network controllers, enabling autonomous operation without continuous internet connectivity. However, continual learning and policy updates may still require periodic access to more powerful computing resources.

Model compression techniques, including network pruning, quantization, and knowledge distillation, reduce computational requirements while maintaining performance. These methods enable deploying sophisticated DRL policies on resource-constrained hardware, expanding the potential deployment scenarios.

6.4. Scalability and Real-Time Deployment

Scaling DRL-based ventilation control from research greenhouses to large commercial operations introduces several challenges. Large facilities may have dozens or hundreds of separately controlled zones, requiring multi-agent coordination at unprecedented scales[209]. Communication overhead, computational scalability, and robust coordination mechanisms become critical as system size increases.

Real-time deployment requires policies that execute within control cycle time constraints, typically 1-5 minutes for greenhouse ventilation. While neural network inference is fast, sensor data acquisition, preprocessing, and communication delays must be accounted for in system design. Asynchronous architectures that allow sensors and actuators to operate at different rates while maintaining coordination may be necessary for complex systems.

Robustness to sensor failures, actuator malfunctions, and communication disruptions is essential for reliable commercial deployment. DRL agents should gracefully degrade performance when information is missing rather than failing catastrophically. Redundant sensors, fault detection mechanisms, and fallback to simpler control strategies provide safety nets during system anomalies.

6.5. Integration with Digital Twins and IoT

Digital twin technology, creating virtual replicas of physical greenhouses that synchronize with real-time data, offers powerful synergies with DRL. Digital twins enable continuous policy improvement through simulation of "what-if" scenarios without physical experimentation, accelerate training through parallel simulated experiences, and support predictive maintenance by identifying degraded performance before failures occur.

IoT integration connects DRL controllers with broader farm management systems, enabling coordinated optimization across multiple subsystems including irrigation, lighting, CO_2 enrichment, and pest management. However, interoperability challenges arise from heterogeneous sensor protocols, data formats, and communication standards. Standardization efforts and middleware platforms that abstract hardware specifics could facilitate integration.

Future systems may incorporate weather forecasts, energy price predictions, and market demand signals into DRL decision-making, optimizing not just immediate climate control but broader economic objectives. Multi-timescale hierarchical control architectures, where high-level DRL agents set strategic goals and low-level agents execute tactical control, offer a framework for managing this complexity.

7. Figures

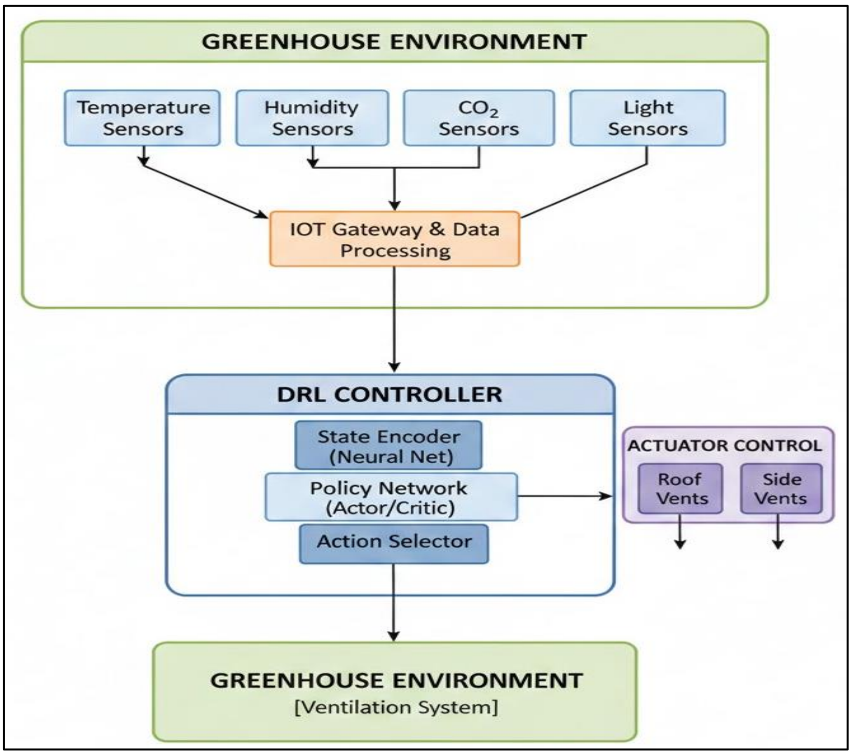

Fig 1: Architecture of a smart greenhouse ventilation system using DRL

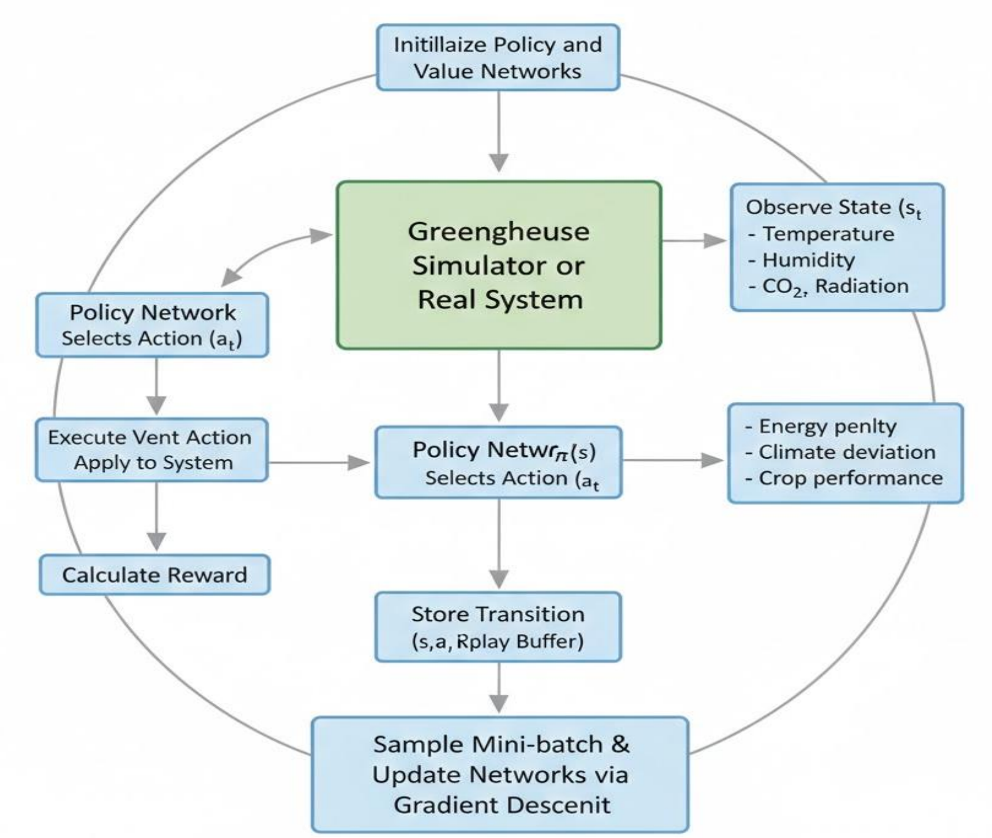

Fig 2: DRL training loop for dynamic greenhouse ventilation control

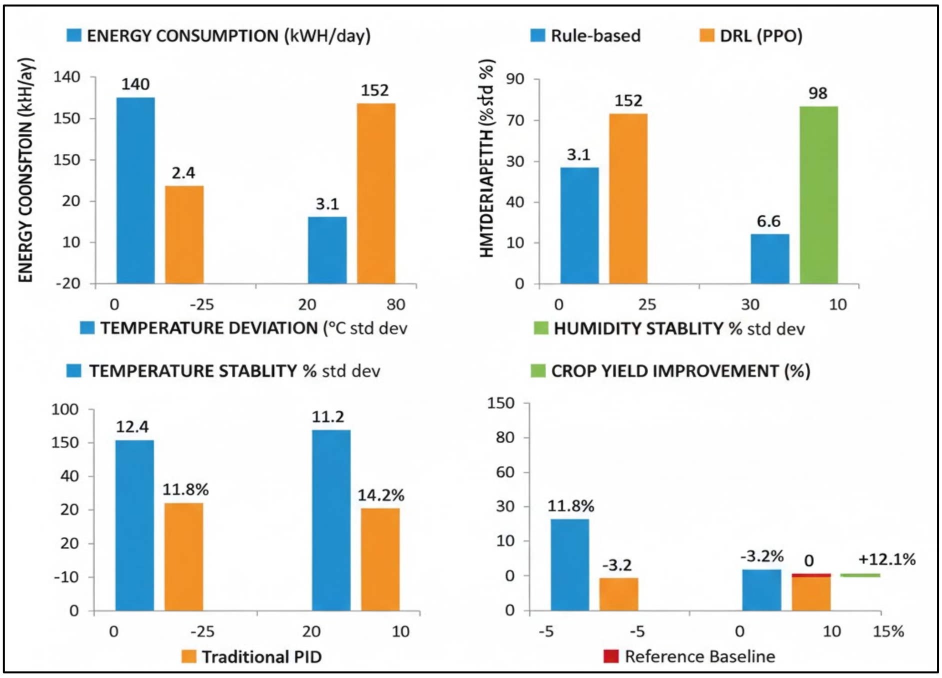

Fig 3: Comparison of traditional control vs DRL-based ventilation performance

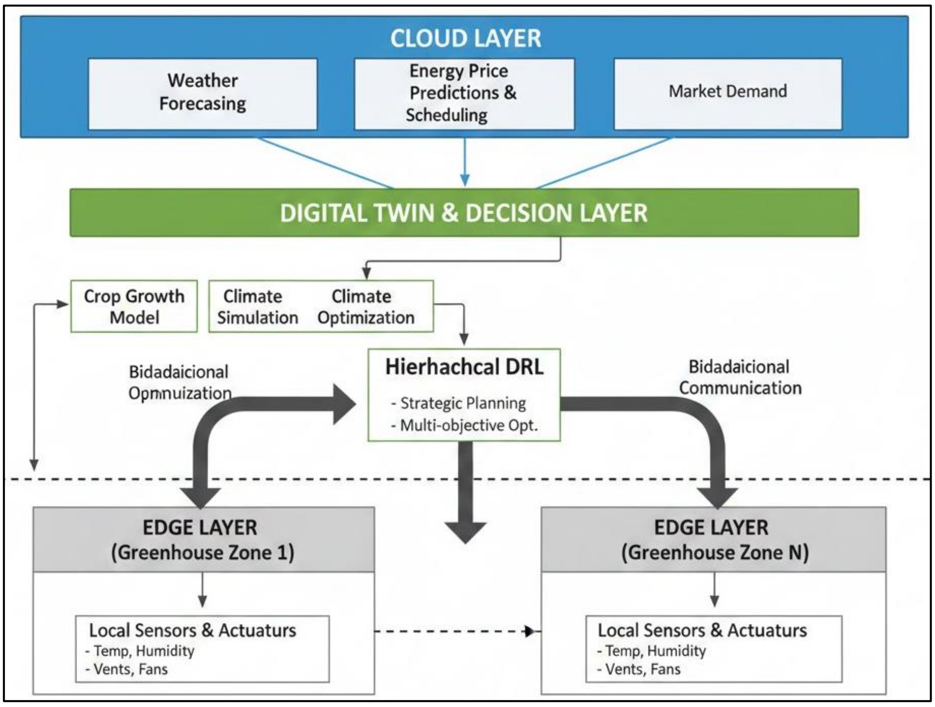

Fig 4: Future framework integrating DRL, IoT, and digital twins for greenhouses

8. Tables

Table 1: Key environmental parameters and control objectives in greenhouse ventilation

Parameter	Optimal Range	Impact on Crops	Ventilation Role	Control Challenge
Temperature (°C)	18-28 (day)15-20 (night)	Photosynthesis rate, respiration, development timing	Primary regulation mechanism through heat exchange	Rapid fluctuations, solar radiation variability
Relative Humidity (%)	60-80	Transpiration, disease pressure, nutrient uptake	Moisture removal, temperature-mediated capacity change	Coupling with temperature, diurnal variation
Vapor Pressure Deficit (kPa)	0.6-1.2	Stomatal conductance, water stress	Simultaneous temp/humidity control	Nonlinear relationship, crop-specific optima
CO_2 Concentration (ppm)	400-1200 (enriched)	Photosynthesis substrate, growth rate	Gas exchange (often conflicting with temp control)	Trade-off with energy conservation during enrichment
Air Velocity (m/s)	0.3-1.0	Transpiration enhancement, stem strengthening	Spatial distribution of ventilation flow	Sensor placement, actuator-flow relationship
Light Intensity (μmol/m²/s)	200-800 (varies by crop)	Photosynthetic capacity	Indirect (shading reduction through structure)	Not directly controlled via ventilation

Table 2: Comparison of DRL algorithms used for greenhouse climate control

Algorithm	Type	Action Space	Key Advantages	Disadvantages	Greenhouse Applications
DQN	Value-based	Discrete	Sample efficient, stable with experience replay	Cannot handle continuous actions, potential overestimation	Discrete vent positions (5-10 levels)
Double DQN	Value-based	Discrete	Reduces overestimation bias	Still limited to discrete actions	Improved discrete vent control
PPO	Policy gradient	Continuous/Discrete	Sample efficient, stable updates, good generalization	More hyperparameters to tune	Multi-zone coordination, continuous actuators
SAC	Actor-critic	Continuous	Maximum entropy encourages exploration, robust	Computationally intensive	Precise vent angle control, robust policies
DDPG	Actor-critic	Continuous	Direct policy learning for continuous control	Prone to instability, sensitive to hyperparameters	Continuous vent control (less common)
TD3	Actor-critic	Continuous	Addresses DDPG instability, delayed updates	Slower learning than DDPG initially	Stable continuous ventilation control
A3C	Policy gradient	Continuous/Discrete	Parallel training, good for multi-zone	Requires multiple environment instances	Multi-zone greenhouses with simulation
MADDPG	Multi-agent	Continuous	Coordinated control, scalable to many zones	Training complexity, non-stationarity	Large commercial multi-zone facilities

Table 3: Advantages, limitations, and deployment challenges of DRL-based ventilation systems

Aspect	Details
Advantages	
Energy Efficiency	15-35% reduction in heating/cooling energy compared to conventional controllers
Climate Optimization	20-40% reduction in temperature fluctuations, improved humidity stability
Multi-objective Optimization	Simultaneous optimization of energy, climate, and crop productivity
Adaptability	Self-tuning to changing system dynamics, seasons, and crop stages
Anticipatory Control	Proactive adjustment based on learned patterns (solar radiation, weather)
Model-free Learning	No need for accurate system models, learns directly from data
Limitations	
Data Requirements	Requires extensive training data (millions of samples) or simulation
Training Time	Days to weeks for policy convergence, significant computational resources
Sim-to-Real Gap	Policies trained in simulation may underperform in real greenhouses
Interpretability	Black-box decision making, difficult to understand policy rationale
Safety Concerns	No formal guarantees of constraint satisfaction or stability
Deployment Expertise	Requires machine learning knowledge for tuning and troubleshooting
Deployment Challenges	
Infrastructure Requirements	Comprehensive sensor networks, reliable actuators, IoT connectivity
Integration Complexity	Compatibility with existing greenhouse management systems
Operator Acceptance	Trust building, training on system operation and monitoring
Maintenance & Updates	Continual learning requirements, policy degradation over time
Scalability Issues	Multi-zone coordination, computational constraints in large facilities
Economic Barriers	Initial investment in sensors, computing hardware, and expertise
Regulatory & Safety	Lack of standards for autonomous agricultural control systems

9. Conclusion

Deep reinforcement learning represents a transformative approach to greenhouse ventilation control, offering adaptive, data-driven optimization that surpasses conventional methods in energy efficiency, climate stability, and crop productivity. This review has examined the theoretical foundations of DRL, analyzed key algorithms applied to ventilation control, and evaluated their performance across simulation and real-world deployments. Value-based methods like DQN provide effective solutions for discrete control problems, while policy-based approaches such as PPO and SAC excel in continuous action spaces and complex multi-objective scenarios. Multi-agent architectures enable scalable control of large greenhouse facilities through decentralized execution with centralized coordination.

Empirical evidence demonstrates that DRL-based ventilation systems achieve 15-30% energy savings, reduce climate fluctuations by 20-40%, and improve crop yields by 8-15% compared to traditional controllers. These benefits arise from DRL's ability to learn complex nonlinear control strategies, anticipate disturbances, and optimize multiple competing objectives simultaneously. The integration of deep neural networks enables processing high-dimensional sensor data and generalizing learned behaviors to new situations, while simulation-based training provides safe, efficient learning environments.

Despite these successes, significant challenges must be addressed before widespread commercial adoption. Data scarcity limits learning in diverse greenhouse environments, while training instability and safety concerns require robust algorithmic solutions. The sim-to-real gap necessitates careful transfer learning and adaptation strategies. Interpretability issues may hinder grower acceptance, and computational constraints affect deployment in resource-limited settings. Scalability to large commercial operations and integration with existing IoT infrastructure present additional practical hurdles.

Future research should prioritize developing sample-efficient algorithms that learn from limited real-world data, safe reinforcement learning frameworks with formal guarantees, and explainable AI methods that enhance transparency. Transfer learning and meta-learning approaches could enable rapid adaptation across greenhouse types and crops. Integration with digital twins offers opportunities for continuous improvement and predictive capabilities. Hybrid architectures combining DRL's adaptability with model-based control's interpretability may yield practical systems that balance performance, safety, and trustworthiness.

The convergence of advanced DRL algorithms, comprehensive sensor networks, IoT platforms, and increasing computational capabilities positions intelligent greenhouse ventilation control for significant growth. As algorithms mature and deployment experience accumulates, DRL-based systems have the potential to become standard practice in precision agriculture, contributing to sustainable food production, resource efficiency, and climate change adaptation. Continued collaboration between machine learning researchers, agricultural engineers, and commercial growers will be essential to realize this potential and address the practical challenges of real-world implementation.

10. References

1. Graamans L, Baeza E, van den Dobbelsteen A, *et al*. Plant factories versus greenhouses: Comparison of resource use efficiency. Agricultural Systems. 2018;160:31–43.
2. Benke K, Tomkins B. Future food-production systems: Vertical farming and controlled-environment agriculture. Sustainability: Science, Practice and Policy. 2017;13(1):13–26.
3. van Straten G, van Willigenburg G, van Henten E, *et al*. Optimal Control of Greenhouse Cultivation. Boca Raton (FL): CRC Press; 2010.
4. Kittas C, Boulard T, Papadakis G. Natural ventilation of a greenhouse with ridge and side openings: Sensitivity to temperature and wind effects. Transactions of the American Society of Agricultural Engineers. 1997;40(2):415–425.
5. Sethi VP, Sharma SK. Survey and evaluation of heating technologies for worldwide agricultural greenhouse applications. Solar Energy. 2008;82(9):832–859.
6. Vanthoor BHE, Stanghellini C, van Henten EJ, *et al*. A methodology for model-based greenhouse design: Part 1, a greenhouse climate model for a broad range of designs and climates. Biosystems Engineering. 2011;110(4):363–377.
7. Taki M, Ajabshirchi Y, Ranjbar SF, *et al*. Heat transfer and multilayer perceptron neural network models to predict inside environment variables and energy loss in a semi-solar greenhouse. Energy and Buildings. 2016;110:314–329.
8. Zhang Y, Gauthier L, de Halleux D, *et al*. Design and control of greenhouse ventilation systems. Acta Horticulturae. 1996;406:265–276.
9. Bot GPA. Greenhouse Climate: From Physical Processes to a Dynamic Model [PhD thesis]. Wageningen: Wageningen University; 1983.
10. Körner O, Challa H. Design for an improved temperature integration concept in greenhouse cultivation. Computers and Electronics in Agriculture. 2003;39(1):39–59.
11. Boulard T, Roy JC, Pouillard JB, *et al*. Modelling of micrometeorology, canopy transpiration and photosynthesis in a closed greenhouse using computational fluid dynamics. Biosystems Engineering. 2017;158:110–133.
12. Kittas C, Katsoulas N, Baille A. Transpiration and energy balance of a greenhouse rose crop in Mediterranean summer conditions. Acta Horticulturae. 2001;559:395–400.
13. Shamshiri R, Ismail WIW. A review of greenhouse climate control and automation systems in tropical regions. Journal of Agricultural Science and Applications. 2013;2(3):176–183.
14. Jolliet O, Bailey BJ. The effect of climate on tomato transpiration in greenhouses: Measurements and models comparison. Agricultural and Forest Meteorology. 1992;58(1–2):43–62.
15. Baptista FJ, Bailey BJ, Randall JM, *et al*. Greenhouse ventilation rate: Theory and measurement with tracer gas techniques. Journal of Agricultural Engineering Research. 1999;72(4):363–374.
16. Nederhoff EM. Effects of CO_2 concentration on photosynthesis, transpiration and production of greenhouse fruit vegetable crops [PhD thesis]. Wageningen: Wageningen University; 1994.
17. Stanghellini C. Transpiration of greenhouse crops: An aid to climate management [PhD thesis]. Wageningen: Wageningen Agricultural University; 1987.
18. Seginer I. Optimal greenhouse production under economic constraints. Agricultural Systems. 2002;73(1):75–88.
19. Critten DL, Bailey BJ. A review of greenhouse engineering developments during the 1990s. Agricultural and Forest Meteorology. 2002;112(1):1–22.
20. Campen JB, Bot GPA. Design of a low-energy dehumidifying system for greenhouses. Journal of Agricultural Engineering Research. 2001;78(1):65–73.
21. Pasgianos GD, Arvanitis KG, Polycarpou P, *et al*. A nonlinear feedback technique for greenhouse environmental control. Computers and Electronics in Agriculture. 2003;40(1–3):153–177.
22. Tantau HJ. Optimal control for plant production in greenhouses. IFAC Proceedings Volumes. 1998;31(12):187–192.
23. Rodríguez F, Berenguel M, Guzmán JL, *et al*. Modeling and Control of Greenhouse Crop Growth. Cham: Springer; 2015.
24. van Henten EJ. Greenhouse climate management: An optimal control approach [PhD thesis]. Wageningen: Wageningen Agricultural University; 1994.
25. Udink ten Cate AJ, Challa H, Bot GPA, *et al*. Computer control of greenhouse climates. Acta Horticulturae. 1978;87:265–272.
26. van Straten G, van Willigenburg LG, van Henten EJ. Optimal greenhouse cultivation control: Survey and perspectives. IFAC Proceedings Volumes. 2010;43(26):18–33.
27. Blasco X, Martínez M, Herrero JM, *et al*. Model-based predictive control of greenhouse climate for reducing energy and water consumption. Computers and Electronics in Agriculture. 2007;55(1):49–70.
28. Coelho JP, de Moura Oliveira PB, Cunha JB. Greenhouse air temperature predictive control using the particle swarm optimisation algorithm. Computers and Electronics in Agriculture. 2005;49(3):330–344.
29. van Beveren PJM, Bontsema J, van Straten G, *et al*. Minimal heating and cooling in a modern rose

greenhouse. Applied Energy. 2015;137:97–109.
30. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. Cambridge (MA): MIT Press; 2018.
31. Mnih V, Kavukcuoglu K, Silver D, *et al*. Human-level control through deep reinforcement learning. Nature. 2015;518(7540):529–533.
32. Arulkumaran K, Deisenroth MP, Brundage M, *et al*. Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine. 2017;34(6):26–38.
33. Li Y. Deep reinforcement learning: An overview. arXiv preprint arXiv:1701.07274; 2017.
34. François-Lavet V, Henderson P, Islam R, *et al*. An introduction to deep reinforcement learning. Foundations and Trends in Machine Learning. 2018;11(3–4):219–354.
35. Brockman G, Cheung V, Pettersson L, *et al*. OpenAI Gym. arXiv preprint arXiv:1606.01540; 2016.