Advances in Applied Energy 9 (2023) 100119

Contents lists available at ScienceDirect

Advances in Applied Energy

journal homepage: www.elsevier.com/locate/adapen

Energy-efficient AI-based Control of Semi-closed Greenhouses Leveraging Robust Optimization in Deep Reinforcement Learning

Akshay Ajagekar[a], Neil S. Mattson[b], Fengqi You[a,c,*]

[a] Systems Engineering, Cornell University, Ithaca, New York 14853, USA
[b] School of Integrative Plant Sciences, Cornell University, Ithaca, New York 14853, USA
[c] Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York 14853, USA

ARTICLE INFO

Key words:
Artificial intelligence
climate control
deep reinforcement learning
robust optimization
energy efficiency
greenhouse

ABSTRACT

As greenhouses are being widely adopted worldwide, it is important to improve the energy efficiency of the control systems while accurately regulating their indoor climate to realize sustainable agricultural practices for food production. In this work, we propose an artificial intelligence (AI)-based control framework that combines deep reinforcement learning techniques to generate insights into greenhouse operation combined with robust optimization to produce energy-efficient controls by hedging against associated uncertainties. The proposed control strategy is capable of learning from historical greenhouse climate trajectories while adapting to current climatic conditions and disturbances like time-varying crop growth and outdoor weather. We evaluate the performance of the proposed AI-based control strategy against state-of-the-art model-based and model-free approaches like certainty-equivalent model predictive control, robust model predictive control (RMPC), and deep deterministic policy gradient. Based on the computational results obtained for the tomato crop's greenhouse climate control case study, the proposed control technique demonstrates a significant reduction in energy consumption of 57% over traditional control techniques. The AI-based control framework also produces robust controls that are not overly conservative, with an improvement in deviation from setpoints of over 26.8% as compared to the baseline control approach RMPC.

1. Introduction

Global food production systems account for up to one-third of all greenhouse gas emissions [1]. As the demand for healthy and fresh food is projected to increase 60% by the year 2050 due to increasing global population [2], reducing the carbon footprint of agricultural practices is central to limiting climate change. Greenhouse farming offers a resource-efficient method (reduced water, fertilizer, and transportation) for crop production and has experienced a significant increase in its adoption worldwide [3]. Controlled environment agriculture (CEA) in greenhouses can further improve crop yields while allowing for food production near locations of consumption despite crop growth prohibiting climates and seasons. Although greenhouse farming is critical for ensuring food safety, it is also energy-intensive compared to other agricultural sectors due to the necessity of maintaining the climate in CEA by frequently regulating environmental parameters such as lighting, temperature, and relative humidity [4]. The development of energy-efficient high-tech greenhouses is the key step towards achieving sustainable food production and enabling food security [5].

Costs associated with energy consumption and labor are the primary contributors to the overall operating costs incurred in crop production with greenhouses [6]. Energy utilization by control systems to maintain suitable climate conditions in the greenhouse make up over 20% of the greenhouse farming costs and may even increase for greenhouse structures located in extreme climates [7]. Indoor air temperature, relative humidity, carbon dioxide (CO_2) density, and lighting intensity are some of the vital environmental parameters that need to be monitored in a greenhouse environment to prevent any undue crop stress and optimize yield [8]. Their regulation is achieved through mechanical systems typically guided by the control strategies derived from growers' expertise and implemented in a greenhouse computerized climate control system (ex. implementing thresholds for heating and ventilation and shading or lighting) [9]. To reduce energy consumption and its accompanying costs, efficient operation of mechanical systems is of utmost importance. However, the complex interactions of the greenhouse microclimate variables with these systems, the crop, and the outdoor environment constitute a difficulty in formulating an energy-efficient control strategy for greenhouse climate control [10]. Artificial intelligence (AI) is a promising solution for realizing sustainable greenhouse farming through

* Corresponding author.
E-mail address: fengqi.you@cornell.edu (F. You).

https://doi.org/10.1016/j.adapen.2022.100119
Received 12 November 2022; Received in revised form 8 December 2022; Accepted 21 December 2022
Available online 22 December 2022

combining horticulture knowledge with predictability, thus leading to a reduction in energy consumption and labor costs facilitated by automation [11,12]. AI can help CEA in greenhouses with various aspects like pest and disease detection, maximizing space utilization, and accurately forecasting crop yields. In addition to utilizing AI for greenhouse climate control and monitoring crop quality, tasks like water management for irrigation can be tackled with AI to improve the crop yield and energy efficiency of traditional greenhouses [13].

Conventional control systems implemented in greenhouses adopt feedback control structure wherein the required controls are computed using sensor measurements of the environmental parameters [14]. Proportional-integral-derivative (PID) is one such control technique that minimizes the deviation of measured variables from desired setpoints over time and has been widely used in greenhouses due to its high reliability [15,16]. Model predictive control (MPC) is a sophisticated approach for greenhouse control and has been extensively investigated owing to its ability to optimize specific control objectives subject to constraints enforced on the environmental parameters as well as the actuators [17–19]. MPC has been demonstrated to lower energy consumption in greenhouses without significantly impacting the accuracy of the controls [20]. However, model-based approaches like MPC assume the availability of a state-space model that accurately describes the system dynamics governed by the involved physical phenomena. Within a greenhouse, the complex interactions among its individual components, controls, and environmental disturbances like outside weather conditions make it difficult to obtain a state-space model from first principles [21]. Additionally, utilizing approximation techniques to model the dependence of greenhouse controls on phenomena like crop growth and heat flow in different regions of the greenhouse may produce errors with model-based approaches. AI techniques like reinforcement learning (RL) which is a feedback-based learning approach that allows for the adoption of a model-free approach in greenhouses and have demonstrated efficacy in agricultural tasks like irrigation [22], water management [23], and crop management [24]. Deep reinforcement learning (DRL), which combines the capabilities of deep learning with RL, has also been successfully applied to climate control in greenhouses [25]. Deep learning allows for hierarchical feature learning which captures both local and inter relationships in the data. DRL offers an opportunity to design a controller for greenhouse indoor climate without assuming a prior knowledge of the underlying system dynamics. This, combined with its learning capabilities for estimating the cumulative future energy consumption or costs given the current greenhouse state and control pair makes DRL a powerful technique to adopt for control applications in greenhouses. Among the various climate control strategies investigated for greenhouse [10], uncertainties associated with time-varying crop growth parameters, sensor measurement errors, as well as external disturbances like outdoor weather conditions are not always soundly tackled. Such uncertainties may lower the precision of the obtained controls [26], consequently affecting the energy efficiency of the control system and harming crop growth in the greenhouse. Control techniques like robust model predictive control (RMPC) which performs robust optimization over a fixed horizon have been shown to exhibit economic efficiency owing to their capability of handling weather uncertainties [27,28]. The advantages of tackling uncertainties with RMPC are especially apparent for climate control in greenhouses situated in areas of harsh weather conditions [29]. However, the reliance of common robust control strategies on an accurate state-space model is an additional limiting factor during their implementation in a practical greenhouse setting. As a result, there arises a need to develop an AI-based model-free approach for regulating climate in greenhouses that produces precise and robust energy-efficient controls for realizing sustainable greenhouses.

There are several research challenges associated with developing an AI-based control strategy that utilizes DRL to yield robust controls for energy-efficient CEA operations of the greenhouse. The first challenge lies in constructing an adaptive DRL-based control strategy that can adjust to time-varying nonlinearities present in complex environments by learning from recent experiences or observations. The variable amounts of CO_2 consumption from crops in greenhouses over their growing cycle and its interactions with the energy required to pump CO_2 into the greenhouse is one such example of time-dependent nonlinearity. Feedback control techniques are capable of handling the system nonlinearities [30,31], but they do not account for system variability over time and may even utilize unsubstantiated nonlinear models. The lack of provision for strictly enforcing constraints on the different controls that comply with the actuators present in the greenhouse is another challenge faced by common DRL-based approaches applied to greenhouse climate control [25]. A further challenge lies in ensuring that the AI-based control strategy can hedge against various uncertainties associated with varying outdoor weather conditions and produce robust controls. Controls that are robust to such external disturbances may be too conservative [32], leading to excess energy consumption in a greenhouse. Therefore, the final challenge is to establish the robustness of the control strategy employed in a greenhouse without compromising its energy efficiency.

In this work, we propose an energy-efficient AI-based control framework for greenhouse climate control that leverages DRL for automatic control along with robust optimization to produce safe and robust controls over the crop's growing cycle. A DRL strategy, namely, Q-learning, is employed to extract the dependence of energy consumption on greenhouse climate states and applied controls from historically recorded observations. This learning process is performed by utilizing a neural network as a function approximator which is trained in an off-policy manner without the need for additional neural networks to approximate the climate control policy. The trained function approximator estimates the subsequent energy consumption based on the inputs like greenhouse climate variables, external weather conditions, and applied controls for regulating indoor climate. This is followed by formulating an optimization problem that serves as a surrogate for the trained neural network, thus enabling actuator constraints on greenhouse controls to ensure their safety. In order to address the robustness considerations of the greenhouse control systems, a robust formulation is derived by treating future weather disturbances as uncertain variables and constructing appropriate uncertainty sets from historical weather observations. The proposed control strategy then employs the function approximator and the robust formulation to yield controls that minimize energy consumption in a greenhouse while adapting to current observations by learning from recently recorded experiences in an online manner. A case study of greenhouse climate control for the tomato crop over its growing cycle is presented to validate the performance and energy efficiency of the proposed AI-based control framework. We also provide detailed comparisons against state-of-the-art control approaches like certainty equivalent MPC (CEMPC), RMPC, and deep deterministic policy gradient (DDPG) for greenhouse climate control. The main contributions of this work are summarised as follows:

- A novel AI-based control strategy that leverages a model-free DRL approach to learn insights from previous greenhouse climate trajectories coupled with robust optimization to produce energy-efficient controls by hedging against weather uncertainties.
- Optimal controls and CEA operations are produced by a novel integration of formulating robust optimization problem with learning complex nonlinear relationships between greenhouse climate, controls, and subsequent energy consumption.
- Energy-efficiency capabilities of the proposed AI-based control strategy realized without compromising the robustness of the obtained controls are demonstrated for climate control in a greenhouse in New York. An improvement of over 57% in energy consumption is observed against conventional model-based MPC and model-free DRL control strategies.

The remainder of this paper is structured as follows. A brief background on RL is presented in the Appendix. The greenhouse energy simulation constructed in this work is described in Section 2. The proposed AI-based energy-efficient control strategy along with a case study on greenhouse climate control for the tomato crop, are presented in Sections 2 and 3, respectively. Finally, conclusions are drawn in Section 4.

2. RO-DRL framework: An AI-based greenhouse climate control and energy optimization system

We propose a novel control strategy that leverages robust optimization (RO) in DRL for autonomous greenhouse climate control. This controller strategy is hereafter referred to as the RO-DRL controller. The design of the RO-DRL control strategy can be divided into two phases. The primary phase involves training a DRL agent in an offline setting to capture the cumulative benefits of the control action at a given state, meaning the trained agent is able to estimate the subsequent overall energy consumption incurred in the future for any given climate state. This is followed by the formulation of a mixed-integer linear programming (MILP) optimization problem as a surrogate for the trained DRL agent. A robust formulation of this problem is further adopted under the RO paradigm to hedge against external weather uncertainties and enforce actuator constraints on the greenhouse controls. The implementation of the RO-DRL controller is, however, performed in an adaptive manner to learn and generate insights from recent greenhouse climate trajectories. An overview of the flow of data in the proposed AI-based controller for greenhouse climate control is depicted in Fig. 1. The design and implementation phases of the RO-DRL controller are described in the following subsections.

2.1. Energy modeling for a semi-closed greenhouse

For the purposes of controller design and validation, it is important to develop a digital twin for crop production in a semi-closed greenhouse that is equipped with mechanical systems for climate control and is subject to varying weather conditions. Based on the Vanthoor's greenhouse model [33], we construct a digital twin of the CEA by building upon the energy simulation in [34] and simulating the exchanges of heat, carbohydrates, and CO_2 in a heated, ventilated single-zone greenhouse with a floor area of 250m^2. The heat transfer processes include exchanges between different components comprising the greenhouse, like the cover, vegetation, floor, tray, mat, and different soil layers through conduction, convection, and radiation. The transfer of carbohydrate mass between the fruits, leaves, and stems of the crop is simulated with a crop growth model and is also influenced by water evaporation from the leaf surface [34]. On the other hand, the exchange of CO_2 is governed by the photosynthesis process. Crop growth and photosynthesis models for tomatoes are considered in the designed greenhouse digital twin. It should also be noted that the greenhouse climate and crop growth are influenced by external environmental stimuli since this is not a closed system. Specifically, a semi-closed greenhouse is influenced by the outside air temperature, wind speed during ventilation, and solar radiation, allowing for better climate control than conventional open greenhouses and less energy-intensive operation compared to closed food production systems like plant factories [35]. Furthermore, we implement provisions for various actuators in the greenhouse digital twin. They include a heating and air conditioning system for the addition or removal of heat, a CO_2 pumping system for enrichment, humidification, and lighting systems. Control of supplemental lighting directly influencing the photosynthesis model to account for low solar radiation levels is also implemented in this greenhouse simulation.

At any time t, the temperatures of greenhouse components, humidity, and CO_2 concentration, in the greenhouse are measured. The mass of carbohydrates per unit area in the fruit, leaves, and stem, as well as their relative growth, averaged over five days, is also recorded. The state vector $x_t \in \mathbb{R}_+^{21}$ captures these measurements. External disturbances, including the ambient air temperature, sky temperature, solar radiation, wind speed, and external relative humidity, are also measured and can be represented by $w_t \in \mathbb{R}_+^5$. As previously mentioned, the control vector $u_t \in \mathbb{R}^4$ comprises of heating/cooling power, humidification rate, CO_2 enrichment, and power consumed by supplemental lighting. A pictorial description of the greenhouse components, along with the states, controls, and disturbances considered in the designed greenhouse digital twin, is provided in Fig. 2.

2.2. Estimating overall energy consumption of a semi-closed greenhouse

Real-world problems like greenhouse climate control comprise of large state and continuous action spaces. Value-based methods like Q-learning are typically adopted for such complex problems owing to their ability to solve harder problems. However, this involves discretizing the action space since such methods are not suitable for continuous action or

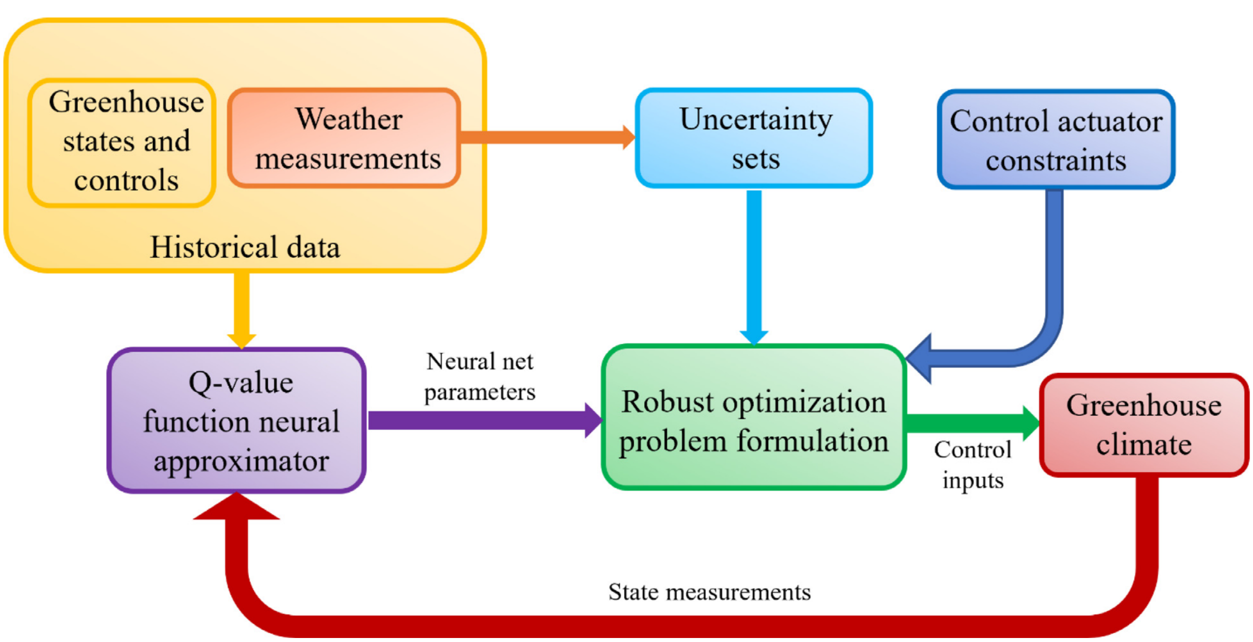

Fig. 1. An overview of the proposed RO-DRL control framework for greenhouse climate control and energy optimization.

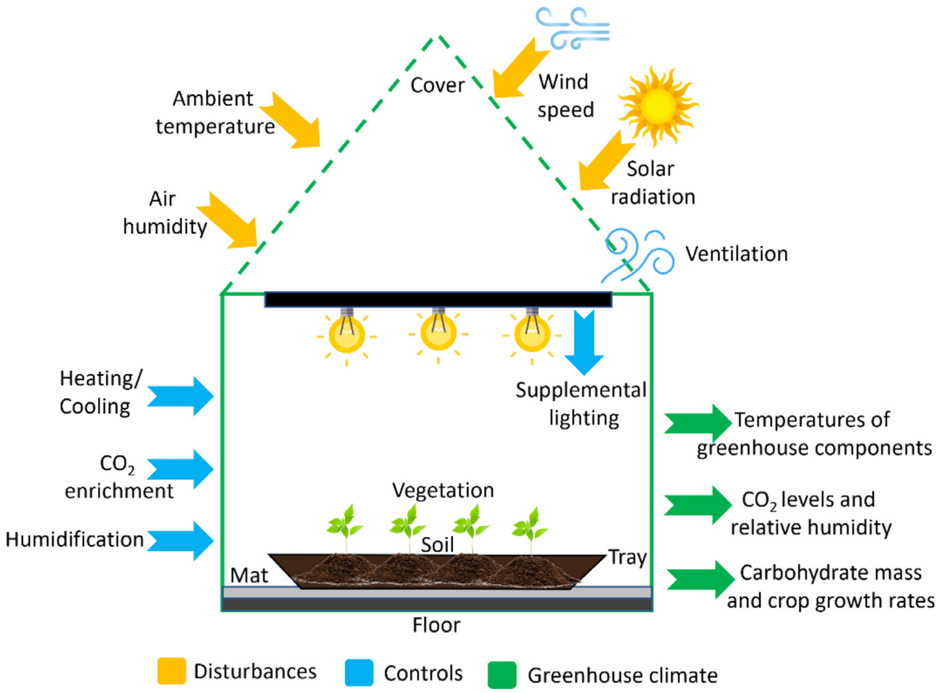

Fig. 2. Greenhouse simulation depicting the external disturbances, control actuators, and greenhouse states along with its components.

control spaces. Previous works in literature on greenhouse control make use of policy gradient methods to tackle continuous action spaces [25]. To overcome difficulties brought forth by policy gradient methods, we utilize a variant of off-policy Q-learning capable of tackling continuous action spaces by leveraging optimization strategies based on gradient ascent.

Q-learning is an off-policy algorithm meaning that the agent can learn from data without considering the behavior of the policy used for data collection [36]. Similar to deep Q-network [37] described in Section A1 of the Appendix, we approximate the value function with a neural network. The offline learning phase assumes the construction of a feedforward neural network consisting of ReLU and linear activations represented by $Q_\theta(s,a)$ that is parameterized by θ and the availability of dataset D consisting of transitions $(s,a,s',r) \in D$ recorded previously with any behavioral policy. The elements in the observation $s = [s_1, \dots, s_n]$, $\forall s_i \in (x, w)$ are a subset of greenhouse states and disturbances, while the action $a = [a_1, \dots, a_n]$, $\forall a_i \in u$ is a subset of available controls. The subset notation is especially useful in specific cases wherein the observations and actions are independent; for example, the greenhouse's temperature is not influenced by the amount of CO_2 supplied. Similarly, rewards r can also be computed from(s,a,s') depending on the specific control application. The greenhouse climate states and observed external weather conditions along with the applied greenhouse controls constitute the input to the neural network, while the output of this value function approximator provides an estimate of subsequent energy consumption over a future fixed horizon. The objective of offline learning is to obtain a trained value function that approximately satisfies Bellman's equation in Eq. (1). For this, the value function $Q_\theta(s,a)$ can then be updated using the target in Eq. (2).

$$Q(s,a) = \mathbb{E}_{s'}\left[r + \gamma Q\left(s', \pi\left(s'\right)\right)\right] \tag{1}$$

$$r + \gamma Q_\theta\left(s', \pi\left(s'\right)\right), \quad \pi\left(s'\right) = \arg\max_a Q_\theta\left(s', a\right) \tag{2}$$

$$\min \frac{1}{|D|} \sum_{(s,a,s',r)\in D} \left[Q_\theta(s,a) - r - \gamma \max_{a'} Q_\theta\left(s', a'\right)\right]^2 \tag{3}$$

Q-learning approaches like DQN are typically applied to RL tasks with discrete action spaces making it easier to extract the optimal actions for a particular state. For continuous action spaces, actor-critic methods are commonly applied where the policy π is parameterized by the actor network [36]. This is done to simplify the action selection procedure for computing $\pi(s')$, which denotes the computed controls with s' representing the greenhouse climate and outdoor weather at the next time step. However, such methods are not always guaranteed to perform well as their success depends on the quality of the actor network in terms of its depth and number of parameters. Here, we employ a gradient ascent method to obtain $\max_a Q_\theta(s',a)$ for computing targets given in Eqs. (1) and (2), which is consistent with the required objective of minimum energy consumption. Performing gradient ascent requires the computation of the gradients $\nabla_a Q_\theta(s',a)$. These gradients can be easily estimated by performing a forward pass through the neural network $Q_\theta(s,a)$ and backpropagating through the network. The learning rule $a \leftarrow a + \eta \cdot \nabla_a Q(s,a)$ can be used to yield $\max_a Q_\theta(s',a)$ after the convergence criterion is satisfied. The gradient update rules and convergence criterion may vary with the choice of optimization algorithms used like L-BFGS or Adam [38]. The overall loss between the current value function estimates and the appropriate targets computed with gradient ascent is minimized as shown in Eq. (3) to yield the trained value function. This optimization procedure can also be performed over smaller batches $B \subseteq D$ for computational efficiency in case of larger transition dataset sizes (Fig. 3).

2.3. Robust optimization formulation to minimize greenhouse energy consumption

Although the trained value function in conjunction with the gradient ascent-based action selection scheme can be directly implemented for control applications [39], it poses certain limitations in terms of the performance and validity of the obtained controls. The training of the value function over a fixed dataset may lead to extrapolation errors [40], resulting in erroneous value estimates for out-of-distribution observation-action pairs. This is especially relevant for greenhouse control, where the previously recorded transitions could lie in specific periods like summer, meaning that the insights extracted from the data recorded in summer may not be useful for the controller implemented during periods of low temperatures. Furthermore, it is not straightforward to impose control actuator constraints with the gradient ascent

Q-learning based DRL Algorithm

Initialize
  Initialize Agent : $M,\ \theta,\ \varepsilon,\ \gamma$, batch size
Repeat (for each episode)
  Reset greenhouse environment states and cumulative returns R
  Repeat (for each step of episode)
    Choose action u for current state x using policy derived from $Q_\theta(x,u)$
    $\Pr\left\{u = \text{Gradient-ascent}\left(Q_\theta(x,u)\right)\right\} = 1-\varepsilon$ (Epsilon greedy approach)
    Apply optimal action u^* and observe next state x' and reward r
    Record (x, u, r, x') in memory M
    $R \leftarrow R + r$
    if size(M) exceeds batch size **then**
      $u' = \text{Gradient-ascent}\left(Q_\theta(x',u)\right)$
      $Q^*(x,u) = r + \gamma \cdot Q_\theta(x',u')$
      $\theta \leftarrow \arg\min_\theta \left(Q_\theta(x,u) - Q^*(x,u)\right)$
    end
  end
end
Define Gradient-Ascent$\left(Q_\theta(x,u)\right)$
  Compute gradients : $\nabla_u Q_\theta(x,u)$
  Apply gradients to action u to maximize Q-value: $u \leftarrow u + \eta \cdot \nabla_u Q_\theta(x,u)$
  Return $u^* = \arg\max_u Q_\theta(x,u)$
end

Fig. 3. Learning strategy utilized for offline training phase of the proposed RO-DRL control framework for CEA operations and energy optimization.

approach for extracting actions or controls from the trained value function. To overcome these limitations, we formulate an MILP optimization problem as a surrogate to the trained value function network allowing us to easily incorporate control constraints. This is followed by formulating a robust counterpart to account for the value estimation errors associated with out-of-distribution external climate disturbances. It should be noted that formulating a MILP optimization problem is possible for neural networks with linear and ReLU activations. Other nonlinear activation functions may lead to nonconvex optimization problems resulting in additional complexities during the solution process.

The main idea behind extracting an MILP formulation from the trained value function is to introduce binary variables for each node in the network with ReLU activation, followed by applying the big-M method. We introduce a variable vector (w_v, u_v) to represent the relevant external disturbances and greenhouse controls that serve as inputs when training the value function in an offline manner with $w_v \subseteq s$ and $u_v \equiv a$. For a value function network consisting of L layers, including the hidden and output layers, each layer $1 \le l \le L$ comprises of n_l units. Let $R \subseteq \{1, ..., L\}$ be the layers of the network for which the ReLU activation is applied. For each layer l of the neural network, θ_l^w and θ_l^b form the weight matrix, and the bias vector used to compute the layer outputs. For any given greenhouse climate state x, the robust optimization problem can be formulated as $RO_\theta(x)$ below. The objective function in Eq. (4) indicates the value function network output representing the subsequent cumulative rewards over a fixed horizon and corresponds to maximizing total reward or minimizing overall energy consumption.

The variable vector y_l represents the value of each node in the l'th layer. Additional continuous variable vectors $q_l \in (\mathbb{R})_+^{n_l}$ and binary variable vectors $z_l \in \{0,1\}^{n_l}$ are introduced for layers with ReLU activation [41]. For the nodes with ReLU activation y_l, q_l indicate the positive and negative terms of the node output while z_l represents whether the ReLU is activated. Eqs. (5) and (6) establish the node output computation in a feedforward neural network for layers with ReLU and linear activations, respectively. Constraints (8)-(10) are the big-M constraints wherein the lower bounds LB and upper bounds UB on layer outputs $\theta^w y + \theta^b$ assume finite values.

$RO_\theta(x)$:

$$\max_{y,z,q,G,h} y_L \tag{4}$$

$$s.t.\ \theta_l^w y_{l-1} + \theta_l^b = y_l - q_l, \quad \forall l \in R \tag{5}$$

$$\theta_l^w y_{l-1} + \theta_l^b = y_l, \quad \forall l \in \{1, ..., L\} \backslash R \tag{6}$$

$$y_0 = \left(x, w_v, G w_v + h\right), \quad \forall w_v \in U \tag{7}$$

$$0 \le y_l^k \le U B_l^k z_l^k, \quad \forall l \in R,\ \forall k = \left\{1, ..., n_l\right\} \tag{8}$$

$$0 \le q_l^k \le -L B_l^k \left(1 - z_l^k\right), \quad \forall l \in R,\ \forall k = \left\{1, ..., n_l\right\} \tag{9}$$

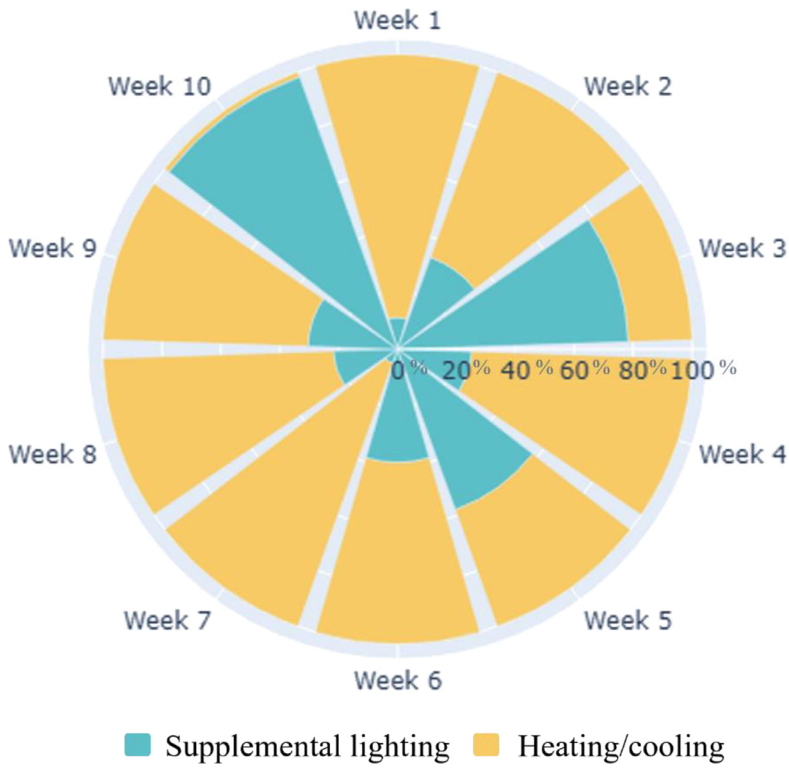

Fig. 4. The proportion of heating/cooling power and supplemental lighting (in percentage) consumed with the proposed RO-DRL control approach.

$$z_l^k \in \{0,1\}, \quad \forall l \in R, \ \forall k = \{1, ..., n_l\} \tag{10}$$

$$F_u\left[Gw_v + h\right] \le f_u, \quad \forall w_v \in U \tag{11}$$

To construct a robust formulation, we incorporate an affine disturbance feedback (ADF) policy in $RO_\theta(x)$ such that $u_v = Gw_v + h$ as reflected by constraints (7) and (11). G and h are the unbounded variable vectors $G \in (\mathbb{R})^{|u_v|\times|w_v|}$, $h \in (\mathbb{R})^{|u_v|}$ in the affine policy. Eq. (7) sets the value function network inputs while (11) imposes the control actuator constraints in the greenhouse dictated by the parameters F_u and f_u. Furthermore, U represents the uncertainty set for external disturbances. Various types of uncertainty sets can be created for this problem [42] depending on the problem structure [43], greenhouse climate requirements, and conservativeness of the obtained controls.

2.4. Implementation of the AI-based energy-efficient controller for CEA

The secondary phase of the RO-DRL controller comprises an adaptive approach to learn from recent greenhouse climate trajectories and utilize corresponding insights through a solution to the robust optimization problem. To ensure that the proposed controller optimizes energy consumption while satisfying climate conditions like temperature and relative humidity setpoints, and CO_2 thresholds, it is important to specify the reward function for model-free approaches accurately. Here, we use a reward function that directly translates to the MPC objective commonly employed over a fixed horizon, as shown in Eq. (12), where ξ_s and Q represents the deviation of the s'th state variable from its corresponding setpoint and a user-defined non-negative parameter for scaling the reward, respectively. The term $E_i(u_i)$ in the reward function denotes the energy consumption incurred by applying the control action u_i in a greenhouse. A consistent reward function is used throughout the offline learning phase, as well as the adaptive implementation of the RO-DRL controller.

$$r_t = -\sum_i E_i(u_i) - \sum_s Q_s \xi_s^2 \tag{12}$$

The performance of the proposed RO-DRL controller is directly dependent on the quality of value network parameters which is influenced by the dataset and algorithm used for offline learning. During the implementation of the RO-DRL controller for greenhouse climate control, we can perform learning steps to update the value function estimator with previously unseen transitions to accurately estimate energy consumption levels associated with greenhouse controls. To this end, an empty buffer B_o is initialized to record new transitions. Over a period of τ timesteps, the transitions (s, a, s', r) are added to B_o. After every τ timesteps, the value function network is updated according to Eqs. (2)-(3) to yield improved estimates of the value function for out-of-distribution observation-action pairs. This is followed by reformulation of $RO_\theta(x)$ to $RO_{\theta'}(x)$ with updated value network parameters θ'.

3. Application to a climate-controlled greenhouse for tomato crop growth

In this section, the applicability of the proposed DRL-based control framework that leverages robust optimization is demonstrated through the application to greenhouse climate control over the growing cycle of the tomato crop. We also evaluate the performance of the proposed method in terms of both energy consumption as well as its ability to regulate the greenhouse environmental parameters. Indoor greenhouse climate conditions that are conducive to the growth of tomatoes are selected as the setpoints and thresholds for different climatic variables. Specifically, the day and night air temperatures are required to be maintained at 25°C and 20°C, respectively, for greenhouse tomato varieties [44]. The humidity levels inside the greenhouse should also be kept at 70-85% relative to the air temperature during the day and night, respectively. In addition to the temperature and relative humidity setpoints, the CO_2 density in the greenhouse is modulated to vary between 600 ppm and 1100 ppm to promote photosynthesis in tomato crops [45]. The rate of photosynthesis in tomato crops is also affected by lighting levels and requires 400-500 μmol/m^{2}s light intensity to promote the growth process. The viability of the proposed RO-DRL framework is validated

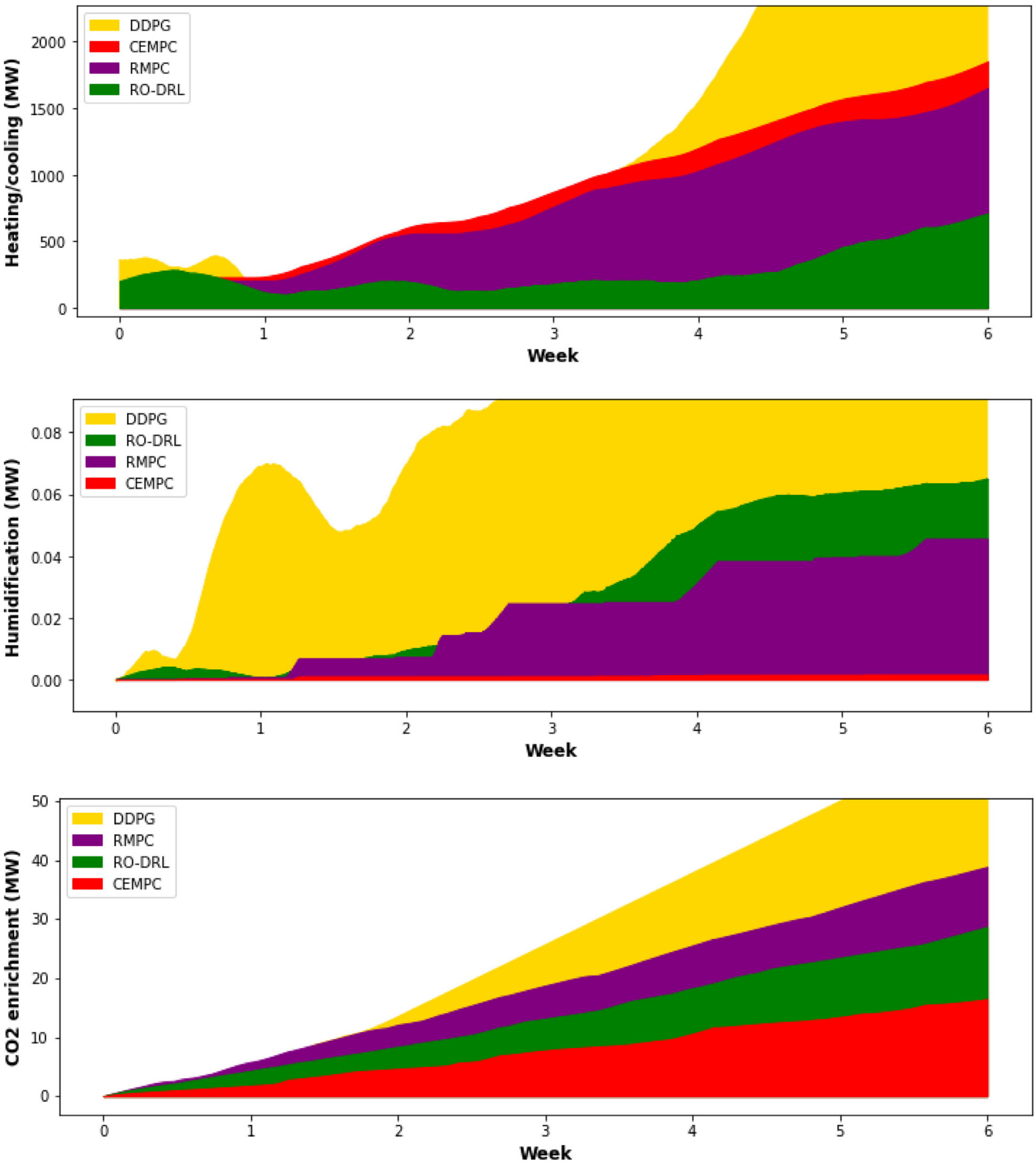

Fig. 5. Power supplied (in MW) for operating various controls. These include addition or removal of heat from the greenhouse, humidification to regulate water vapor density levels, and pumping CO_2 into the greenhouse.

against commonly used state-of-the-art model-based control strategies like MPC and model-free reinforcement learning.

3.1. AI framework and baseline configuration

As the initial phase of the proposed RO-DRL control strategy learns from historical climate trajectories, we perform a simulation of the greenhouse climate controlled with a proportional rule-based controller over a period of 90 days from January to March in the year 2021. All greenhouse climate states, external weather disturbances, and controls applied are recorded over a resolution of 60 seconds resulting in a dataset D containing 129,600 transitions. The greenhouse climate states include measurements of the temperature of various greenhouse components, water vapor density, CO_2 levels within the greenhouse, as well as mass of carbohydrates measured in different crop regions. The weather disturbances considered for the RO-DRL controller design comprise of external air temperature, relative humidity, and amount of solar radiation, while all the greenhouse controls like addition or removal of heat, humidification, CO_2 enrichment, and supplemental lighting, constitute the control vector considered here. A neural network with the following configuration of (128:ReLU, 64:ReLU, 16:ReLU) describing the number of neurons in the hidden layers is constructed with one neuron and linear

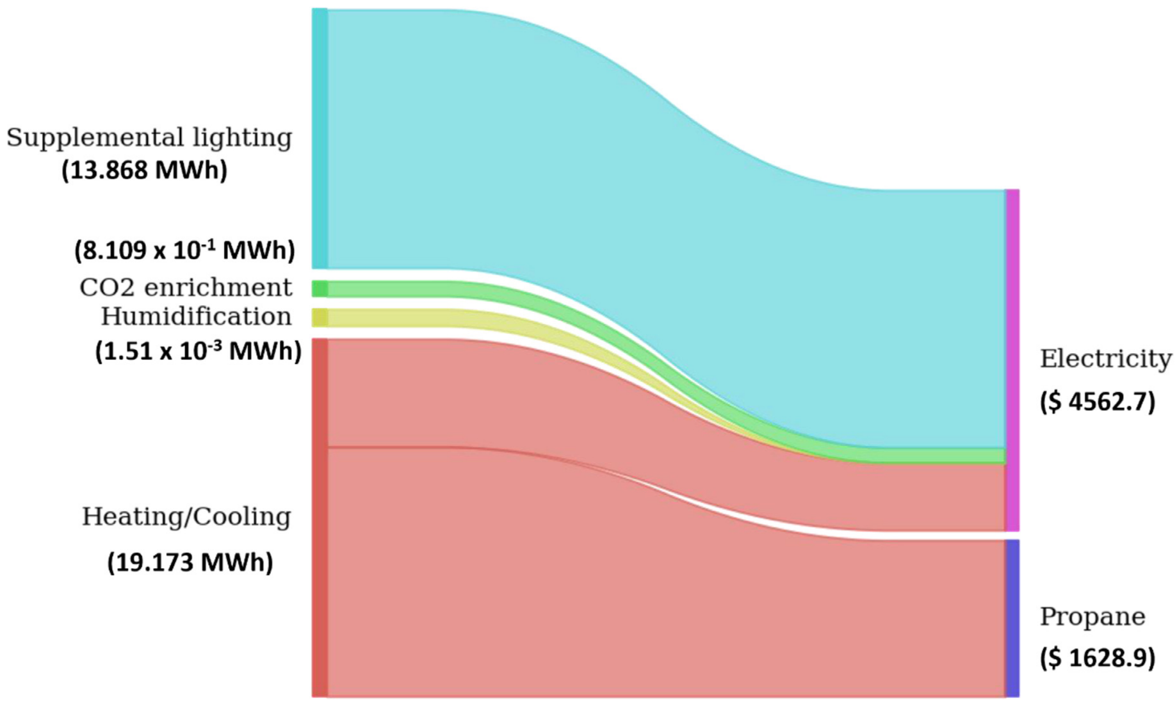

Fig. 6. Relationship between energy consumption incurred over the growth cycle of ten weeks associated with different controls in the greenhouse obtained with the RO-DRL control strategy and the costs incurred by applying them for greenhouse climate control.

activation in the output layer. This neural network represents the value function or the subsequent energy consumption over a fixed horizon for a pair of the indoor-outdoor climate state vector and applied control vector. This value function network is trained with the offline learning strategy until the temporal difference (TD) error does not change beyond one percent in consecutive iterations. The Adam optimizer is used to extract continuous actions from the value function estimator while computing targets for Q-learning.

For evaluation purposes, we simulate the operation of climate control in the greenhouse for the growth of tomatoes from July to September over a period of ten weeks. This period is chosen for evaluation due to the distinction in weather patterns as compared to those recorded in the dataset *D*. In addition to RO-DRL as the control strategy, we employ two MPC-based methods, namely, CEMPC and RMPC as the baselines. A model-free DRL technique DDPG is also chosen as a baseline control technique. It should be noted that the system dynamics of the CEA and crop growth in the greenhouse are treated as unknown throughout this work. However, as applying MPC-based methods requires a state-space model, we construct a discrete-time time-invariant system represented by $x_{k+1} = Ax_k + Bu_k + Cw_k$. This system can be obtained by system identification performed through the linear fitting of the historical greenhouse climate trajectories in dataset *D* [46]. CEMPC assumes a nominal value for the external disturbances and results in solving a deterministic optimization problem over a control horizon. On the other hand, RMPC uses an ADF policy with uncertainty sets constructed to tackle the uncertain variables corresponding to external air temperature, relative humidity, and solar radiation in the optimization problem. For each disturbance variable, we construct box uncertainty sets *U* for use in the robust optimization problem arising in RO-DRL as well as RMPC. A visual depiction of the uncertainty set utilized for RMPC and the RO-DRL controller is presented in Fig. A1 of the Appendix. Additionally, a plot representing the performance comparison of the offline learning phase of the proposed AI-based control framework with another state-of-the-art DRL technique is provided in the Appendix to justify the use of the offline training technique used here. DDPG uses the same configuration of the neural network described above for the value function approximator and a linear policy for a fair comparison with RO-DRL and RMPC. As DDPG does not allow for enforcing complex control actuator constraints, we perform clipping of the policy outputs to restrict the controls within appropriate bounds. All neural networks in the RO-DRL strategy and DDPG are constructed and trained with the PyTorch deep learning framework using a RTX 3090 GPU. During each step of the greenhouse climate control, the robust optimization problem formulated in RO-DRL is solved with the Gurobi 9.5 solver with an Intel i9-10920X processor operating at 3.5Ghz. Optimization problems arising in CEMPC and RMPC are also solved with the same computing environment.

3.2. Energy efficiency

Over the course of the ten-week tomato crop's growth cycle, we measured the energy consumption associated with each control in the 250m^2 greenhouse obtained with the RO-DRL controller and the baseline control techniques. These include measurements of the power supplied to the control actuators for heating or cooling the greenhouse, humidification, CO_2 enrichment, and supplemental lighting at each control interval of 60 seconds recorded during the simulated greenhouse operation. The overall energy consumption associated with controlling the greenhouse climate incurred with the proposed AI-based approach is 33,854 kWh. A significant portion of this consumption is due to the heating and cooling required to maintain greenhouse indoor air temperature and comprises of 19,173 kWh, which constitute more than 56.6% of the overall energy consumed. On the other hand, supplementary lighting is another major contributor to energy consumption, making up for approximately 40.9% of the overall values. The amount of energy utilized for humidification and pumping of CO_2 into the greenhouse accounts for only 2.40% of the overall usage. The proportion of energy utilization by different controls obtained with the proposed RO-DRL control framework is consistent with a typical greenhouse operated by growers with conventional control approaches. For comparison purposes, we consider a semi-closed greenhouse in New York equipped with efficient lighting and control systems. Actuators like heating and cooling systems,

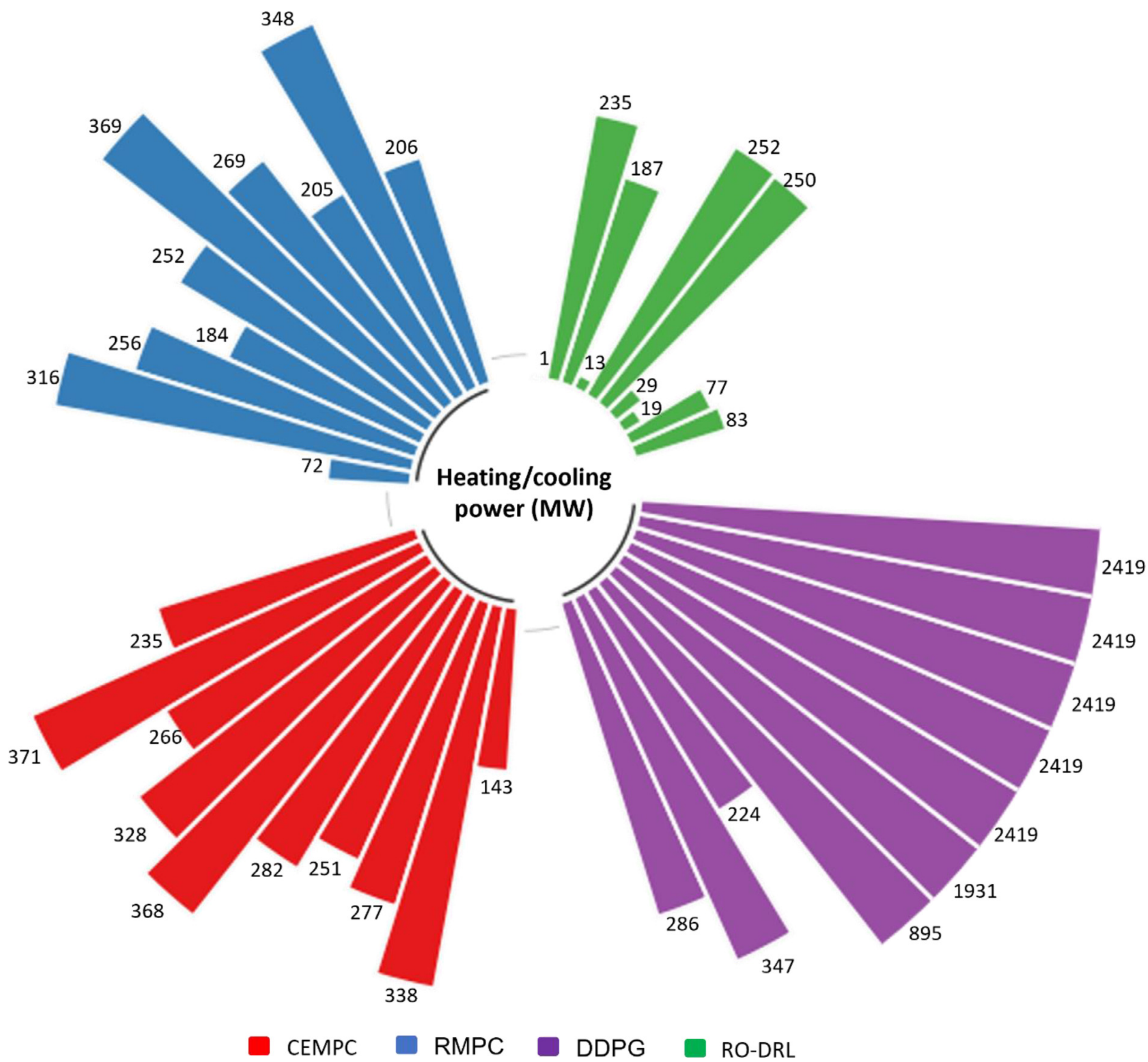

Fig. 7. Power supply in (MW) associated with supplying heating/cooling power in the greenhouse over a growing cycle of ten weeks incurred with different control strategies. Direction of the polar plot is counterclockwise from week 1 to week 10.

growth lights, vents, and others are precisely controlled by experienced growers to maximize plant productivity while minimizing energy usage [47]. Energy usage incurred in this greenhouse over a growing period of ten weeks is 79,258 kWh, while a reduction of over 57% is observed with the AI-based control strategy. Supplying or removing heat from the greenhouse and providing supplemental lighting are responsible for the majority of the energy consumption during climate control in greenhouses. To visualize their contributions, we plot the proportion of energy consumed for heating the greenhouse and maintaining optimal lighting conditions over the crop growth cycle in Fig. 4. The proportion of energy consumed by additional lighting is more than 75% of the overall energy usage for only two weeks of the growth period.

The performance of the RO-DRL controller in terms of energy savings is compared against the baseline control techniques to demonstrate its efficacy. We visualize the power consumption associated with heating and cooling, humidification, and CO_2 supplementation recorded with different control strategies in Fig. 5. All measurements associated with the control actuators in the greenhouse are measured in terms of energy supplied over a fixed interval of 60 seconds. The germination and early growth stage, along with the vegetative period for the tomato crop, account for the first six weeks since cultivation [48] and are presented in this plot. As evident from Fig. 5, DDPG exhibits the worst energy efficiency by producing large controls that lead to high energy consumption. By the sixth week, more than 1,500 MW of power has already been consumed by both CEMPC and RMPC techniques. Based on the lower energy usage corresponding to the humidification in the greenhouse and the relative humidity-temperature relationship, it can be inferred that CEMPC and RMPC choose to regulate the humidity levels by supplying or removing heat. In contrast, the proposed AI-based RO-DRL control strategy learns the complex nonlinear relationship between the climate states of air temperature and water vapor density with the corresponding controls and yields energy-efficient controls to achieve the desired greenhouse climate. Maintaining the CO_2 levels in the greenhouse is especially complex due to the time-varying CO_2 release from crops over their growing cycle. CEMPC incurs the lowest energy consumption to regulate CO_2 levels, however, it fails to achieve greenhouse climate conditions that promote crop growth, as shown in the CO_2 control profile in Fig. 9. Although both RO-DRL and RMPC are able to regulate appropriate CO_2 levels, the AI-based controller demonstrates energy efficiency for CO_2 control. The daily average power consumption incurred over ten weeks with different control strategies are also reported in Table A1 of the Appendix. The objective specified for all control techniques

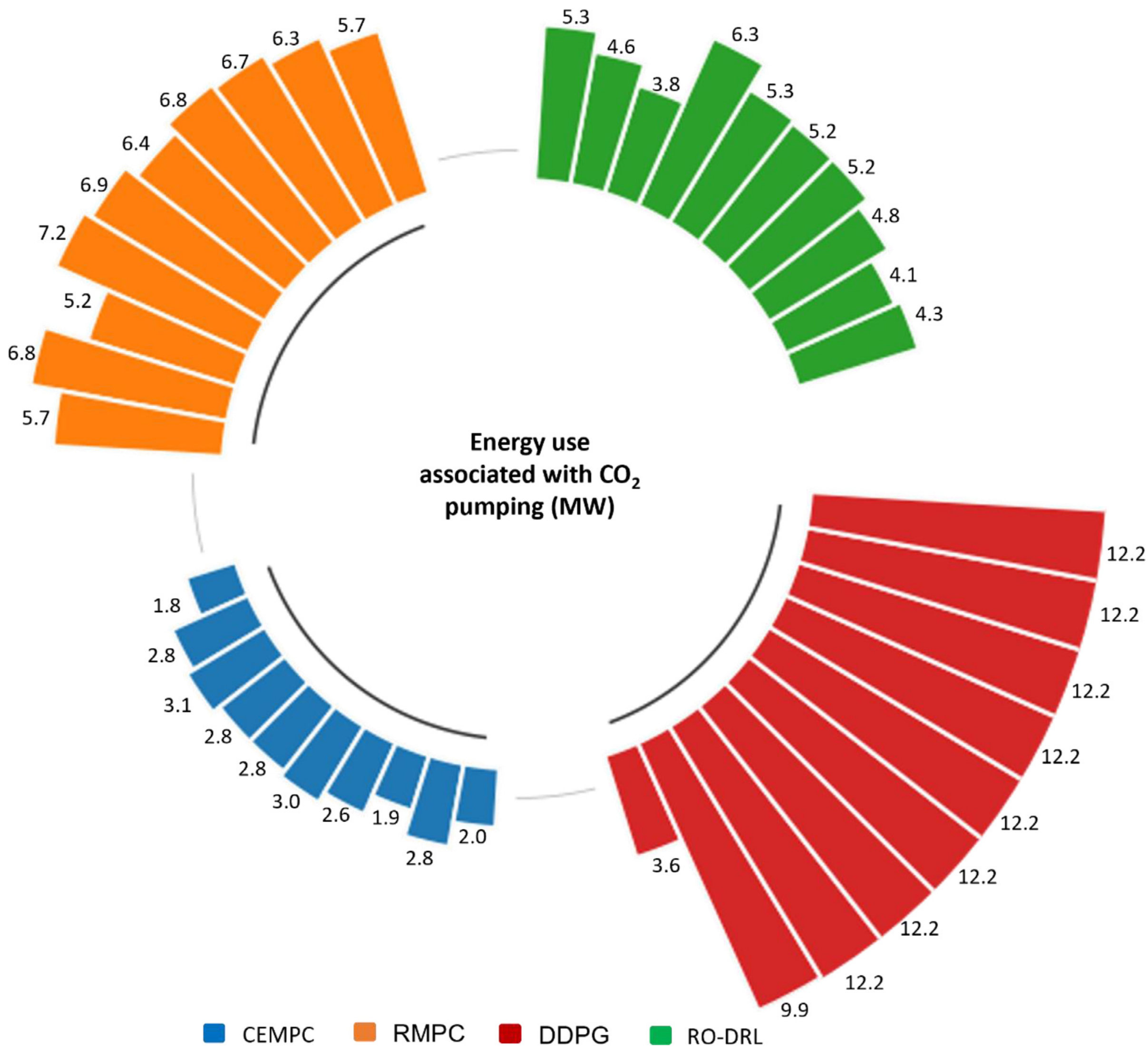

Fig. 8. Power supply in (MW) associated with pumping carbon dioxide into the greenhouse over a growing cycle of ten weeks incurred with different control strategies. The direction of the polar plot is counterclockwise from week 1 to week 10.

validated for this case study, including the proposed RO-DRL approach and other baselines, is to minimize the overall energy consumption associated with the controls for greenhouse climate regulation. However, it is important to note that this objective can be updated to include costs associated with energy usage by updating the reward function in the DRL-based training accordingly. We further provide an estimate of the costs of operating the greenhouse with the proposed AI-based control strategy, as shown in Fig. 6. Supply and removal of heat from the greenhouse requires the use of fuel like natural gas and propane to generate the required heat and electricity from the power grid to operate the mechanical systems for heat transfer. The requirement of propane fuel is estimated by assuming a conversion efficiency of 90% in the furnaces, which incurs a cost of $1628.9 for one growing cycle of the tomato crop in a greenhouse climate maintained with the RO-DRL control approach. The electricity consumption for pumping CO_2 for the enrichment and water vapor for humidification constitutes only $202.3 for one growing season. A significant portion of the electricity consumption in a greenhouse account for the supplementary lighting controlled with our AI-based approach leading to an overall electric expenditure of $4562.7 over the crop's growth period.

A breakdown of the energy consumption is further analyzed to gain insights into the energy efficiency capabilities of the RO-DRL controller in contrast to other state-of-the-art model-based and model-free control approaches. The levels of heating/cooling power required to maintain indoor greenhouse temperature for each week of the crop growth cycle are shown in Fig. 7, while Fig. 8 presents the weekly consumption levels associated with CO_2 enrichment for different control strategies. Energy consumed by supplemental lighting with different control strategies did not exhibit any significant variation. For both heating/cooling power and power required for CO_2 supply, DDPG yields the highest possible control levels set by the corresponding actuators for a significant period of the tomato growth cycle. This can be attributed to the limiting learning behavior of DDPG, which performs poorly when previously unseen climatic conditions are observed. As seen in Fig. 7 and Fig. 8, the weekly energy consumptions for both CEMPC and RMPC follow a common trend owing to their utilization of identical state-space models for greenhouse climate control. However, a significant disparity in the trend followed by the weekly energy consumption with the RO-DRL controller can be observed. Such incongruity of the power consumption trend compared to model-based approaches combined with the lower energy usage substantiates the ability of the proposed AI-based energy-efficient control strategy to adapt to recent climatic conditions and disturbances despite the lack of an accurate model describing the system dynamics of greenhouse climate and crop growth.

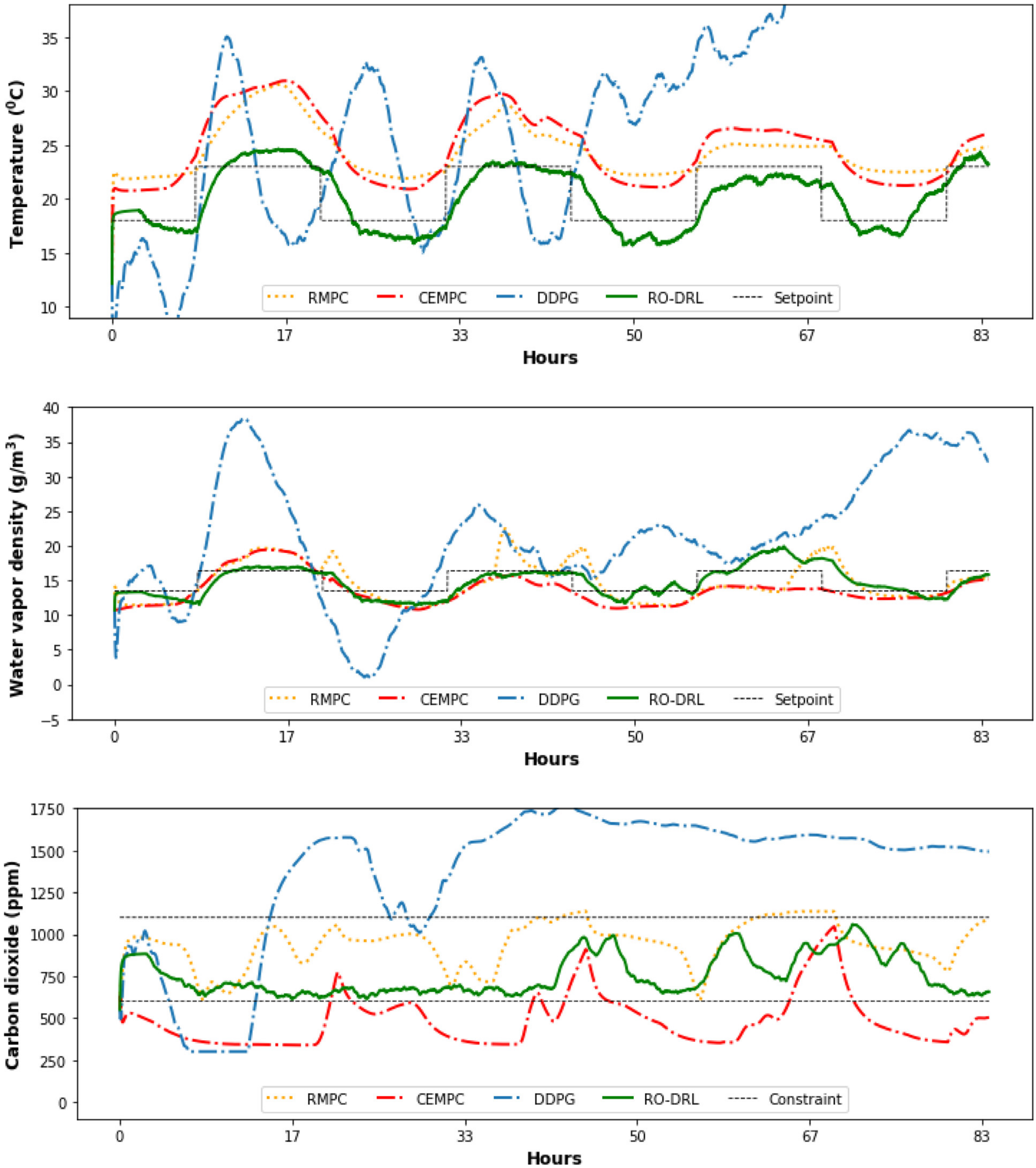

Fig. 9. Control profiles for carbon dioxide, relative humidity, and air temperature levels within the greenhouse obtained with different control strategies over half a week or 84 hours.

3.3. Robustness of the AI-based control strategy

In addition to ensuring that the controls obtained with the specific climate control strategy in a greenhouse exhibit low energy consumption, it is necessary to ensure that the climate conditions within the greenhouse are regulated such that they are conducive for crop growth. As the greenhouse climate and the employed control system is subject to uncertainties like time-varying crop growth and external weather conditions, we validate the robustness of the controls obtained with the proposed RO-DRL control strategy as well as the CEMPC, RMPC, and DDPG approaches. The control profiles for greenhouse indoor air temperature, water vapor density, and CO_2 density are plotted in Fig. 9 to visualize whether the setpoints and thresholds set for optimal crop growth are maintained accordingly. It should be noted that the temperature and relative humidity setpoints are dynamic and vary during the day and night, while thresholds for CO_2 levels are expected to be satisfied over the growth period. As evident from the temperature control profile, only the proposed RO-DRL controller is able to maintain the air temperature at the given setpoints. Although CEMPC and RMPC can follow the required temperature trend, they tend to overshoot the temperature by 5-7°C, which may be detrimental to crop growth. Although the controls obtained with MPC-based approaches are obtained by solving an optimization problem over a fixed horizon, the observed behavior of greenhouse states in the digital twin is inconsistent with that of the

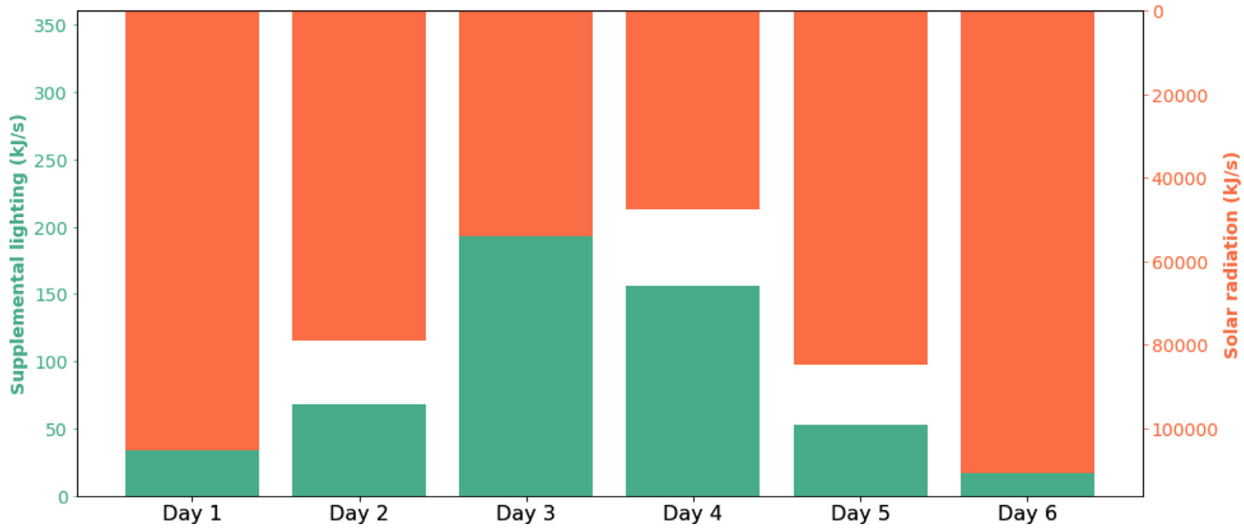

Fig. 10. Supplemental lighting provided in the greenhouse over the first week of July, along with the corresponding observed solar diffuse and direct radiation levels. It should be noted that the axis scale for solar radiation is inverted.

state-space model for greenhouse dynamics. Such inconsistencies can be attributed to the modeling errors for the state-space model, which approximates the actual greenhouse behavior while substantiating the inherent limitations of model-based approaches like CEMPC and RMPC. The model inconsistencies combined with the soft constraints utilized in the MPC formulation that allow deviation from setpoints at the expense of lower energy consumption can lead to the observed overshoot in indoor air temperature with CEMPC and RMPC. On the other hand, water vapor levels within the greenhouse are maintained in a fitting manner with RO-DRL, CEMPC, and RMPC control approaches, with slight overshoots observed with RMPC. Controlling the CO_2 levels in the greenhouse within the required thresholds is a complex task and is observed to be tackled only with uncertainty handling approaches of RMPC and RO-DRL. The constraint satisfaction capabilities of the proposed RO-DRL approach, coupled with its energy efficiency validated in the previous subsection, demonstrate its ability to reduce the conservatism of the controls as compared to those obtained with RMPC. The DDPG control approach not only incurs high energy consumption but also is unable to maintain greenhouse climate. To measure the performance improvement of the proposed AI-based control strategy over CEMPC and RMPC, we compute the average deviations from the setpoints for temperature and water vapor density, and percentage constraint violations for CO_2 levels in the greenhouse. These deviations and violation percentages are reported for each month of the growth phase of the tomato crop in Table A2 of the Appendix. The average percentage deviations observed with RO-DRL are significantly lower than both CEMPC and RMPC with the exception of water vapor density in the month of August. In terms of regulating CO_2 density in the greenhouse within required thresholds, RO-DRL exhibits the lowest percentage of constraint violations over the entirety of the duration.

The adaptive nature of the proposed RO-DRL control method is further verified by analyzing the energy consumed by supplemental lighting with the external disturbance of solar radiation. A rule-based practice for turning on additional lighting in case of insufficient light from the sun is typically employed in greenhouses [49]. A similar trend is observed with the energy supplied for artificial lighting in the greenhouse to make up for low solar radiation, as shown in Fig. 10. Controlling the amount of lighting is especially important for commercial greenhouses that use shade curtains to regulate temperature on sunny days [50]. Ensuring supplemental lighting provision despite high solar radiation levels is challenging in such scenarios. As mentioned in Section 3.1, the historical dataset is recorded over periods of January to March, during which sunlight levels are low, while the simulation conducted for validation is conducted from July to September with higher levels of recorded solar radiation. Despite such out-of-domain uncertainties, the proposed AI-based climate control technique is able to learn from and adapt to newly observed greenhouse climate and weather transitions leading to an energy-efficient and robust control strategy for supplementary lighting control.

4. Conclusion

In this work, we developed an AI-based control framework that leverages DRL for automatic greenhouse climate control with robust optimization to produce energy-efficient controls that help regulate the indoor climate that is conducive to crop growth. A DRL-based offline learning strategy was utilized to extract the complex nonlinear relationship between the greenhouse climate states, controls, external disturbances, and the subsequent energy consumption over a fixed horizon. A robust optimization problem was formulated to minimize overall energy usage with the help of a surrogate for the neural network as a value function approximator trained during the offline learning phase. The RO-DRL controller was further trained in an adaptive manner during its implementation. The viability and performance efficiency of the proposed RO-DRL controller was established with a case study on greenhouse climate control for the tomato crop. Finally, detailed comparisons with model-based control techniques like CEMPC and RMPC, along with model-free DRL approach DDPG, were presented to validate the superiority of the proposed AI-based control framework in terms of energy efficiency and robustness of the obtained controls. The computational results showed that the RO-DRL controller yields over 55% reduction in energy consumption associated with control systems in the greenhouse. Properties of the proposed controller, like its ability to hedge against time-varying uncertainties and produce less conservative controls as compared to RMPC, were also substantiated through the conducted computational experiments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The authors acknowledge support from Cornell Institute for Digital Agriculture (CIDA) for resources utilized in this research.

Appendix

In this section, we provide additional information relevant to the AI-based control framework for CEA proposed in this work. Brief preliminaries on RL and DRL are presented in Section A1 which includes definitions for common RL concepts and background on state-of-the-art DRL techniques. The experimental settings used for the robust control approaches, RMPC and RO-DRL, are provided in Section A2. This section also includes the training comparisons for the offline training phase in RO-DRL, wherein the proposed controller learns from historical greenhouse climate trajectories. Numerical results for the daily average levels of energy consumed in the greenhouse with different control strategies are presented in Section A3.

A1. Background on deep reinforcement learning

RL is a machine learning paradigm that deals with intelligent agents maximizing cumulative reward by taking corrective actions in an environment [36]. From an optimal control perspective, RL provides a model-free framework for solving problems stated as Markov decision processes (MDP). A general RL problem described as an MDP consists of a set of states S, set of actions A, reward function r, discount factor $\gamma \in [0,1]$, and transition dynamics. This problem can be formalized as a discrete-time stochastic control process where an RL agent interacts with its environment at any timestep t by selecting an action $u_t \in A$ after receiving a state $x_t \in S$. This causes the agent to receive a reward r_t and environment state transitions to $x_{t+1} \in S$. Given a state, the RL agent selects an action to perform that is dictated by the control strategy termed as policy π. A deterministic policy provides a mapping from states to actions $\pi : S \to A$, while a stochastic policy is represented by the probability distribution over the set of states and actions. $\pi : S \times A \to [0,1]$. The goal of RL is to learn an optimal policyπ^*that maximizes the expected return defined as cumulative discounted reward in Eq. (A1). The cumulative discounted reward denoted as state-value function $V^\pi(s)$ is the expected return when starting in state x following policyπ subsequently and is also commonly referred to as the V-value function. The optimal expected return is governed by the optimal policyπ^* and can be defined as shown in Eq. (A2). Unavailability of transition dynamics in a typical RL problem allows us to construct state-action value $Q^\pi(x,u)$ defined in Eq. (A3). The optimal policy in Eq. (A4) can be obtained by optimizing the state-action value or Q-value greedily at every state.

$$V^\pi(x) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_{t+1} | x, \pi\right] \tag{A1}$$

$$V^*(x) = \max_\pi V^\pi(x) \quad \forall x \in S \tag{A2}$$

$$Q^\pi(x,u) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_{t+1} | x, u, \pi\right] \tag{A3}$$

$$\pi^*(x) = \arg\max_{u \in A} Q^\pi(x,u) \tag{A4}$$

The use of neural networks to approximate either state value function$V^\pi(s)$, state-action value$Q^\pi(x,u)$, policyπ, or the transition dynamics is referred to as DRL [51]. They can be parameterized by the weights of the deep neural network and is particularly important for scaling up prior state-of-the-art algorithms in RL to higher-dimensional problems. Some of the most popular value-based methods for DRL that aim to build a value function are based on Q-learning [52]. Q-learning based methods like deep Q-network (DQN) [37] and double Q-learning [53] demonstrate exceptional performance with high dimensional sensory states and actions. An alternate class of DRL methods termed as policy gradient methods optimize a performance objective by finding a good policy. DDPG is one such DRL technique that extends the DQN to continuous spaces [54] and is commonly used to handle continuous action spaces.

Table A1
Experimental settings for the case studies conducted, including the greenhouse climate setpoints and the neural network configuration for networks used in RO-DRL and DDPG control strategies

Greenhouse climate setpoints	
Night temperature	25°C
Day temperature	20°C
Relative humidity	80%
CO_2 concentration	600-1100ppm
Lighting level	400-500μmol/m^{2}s
Neural network configuration	
Input size	8
Layer 1 size (activation)	128 (ReLU)
Layer 2 size (activation)	64 (ReLU)
Layer 3 size (activation)	16 (ReLU)
Output size	1

A2. Experimental settings and training performance comparison of the AI framework for CEA

Fig. A1 a depicts the box uncertainty set constructed for use with the robust approaches considered in this work, RMPC, and RO-DRL control techniques. The bounds for each uncertain variable, including external relative humidity, external air temperature, and amount of solar radiation, are selected based on historical weather data. Fig. A1.b visualizes the training profile for the offline training phase of the DRL agent performed with gradient ascent-based Q-learning in the proposed AI-based controller. A comparison against the SARSA approach is also conducted to demonstrate the efficiency of the training strategy in learning from historical greenhouse climate trajectories employed for the RO-DRL controller. Table A1 also summarizes the experimental settings used to conduct the computational experiments, including the control setpoints for different greenhouse climate states as well as the neural network configuration used to represent the Q-value function in RO-DRL and DDPG.

A3. Energy consumption in the greenhouse over the crop growth

We further present numerical results concerning the crop growth and climate control simulations conducted in this work with the greenhouse digital twin and various control strategies. The daily average power consumption for each of the ten weeks is reported in Table A2. These can be utilized to track energy consumption during different phases of

Table A2
Daily average power consumption in the greenhouse over a growing period of 10 weeks from July to September, along with the overall energy consumption incurred with different control strategies.

	CEMPC (kW)	RMPC (kW)	DDPG (kW)	RO-DRL (kW)
July	33,912	30,255	41,411	12,525
	53,463	50,726	51,058	11,711
	38,490	30,336	33,755	3,396
	47,330	39,428	129,665	5,002
August	52,991	53,752	277,631	36,516
	40,827	37,084	347,354	36,801
	36,343	27,375	347,363	2,798
	39,879	37,429	347,365	27,401
September	48,760	46,221	347,365	34,242
	20,726	11,256	347,363	916

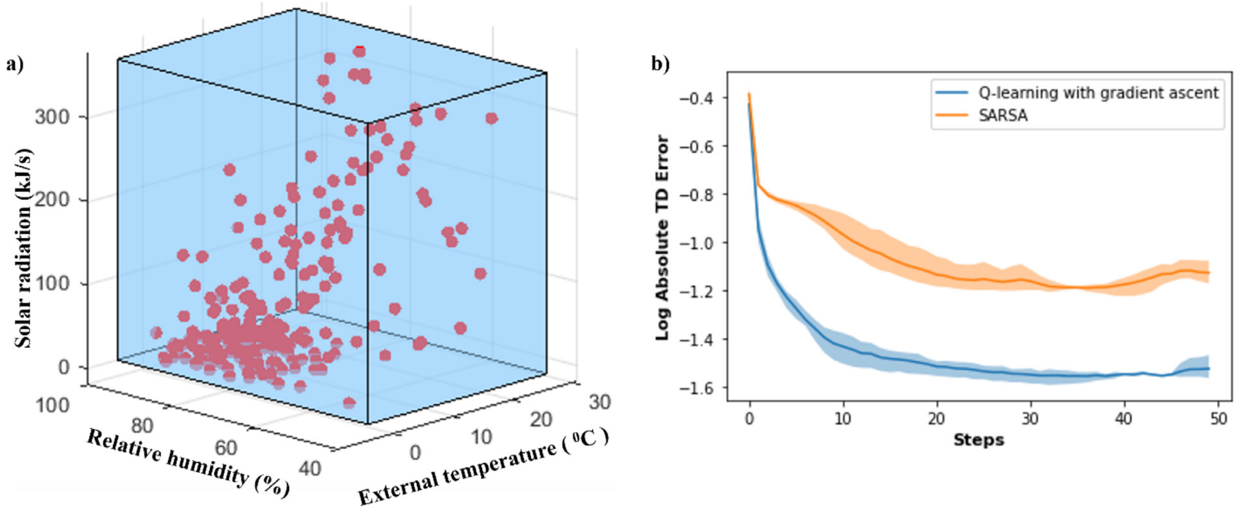

Fig. A1. a) Visualization of the uncertainty set used for RMPC and RO-DRL for greenhouse climate control simulation and b) learning performance comparison between the offline learning strategy utilized in the proposed RO-DRL framework and the SARSA algorithm

Table A3
Daily average power consumption associated with various control actions in the greenhouse, along with the percentage deviations from setpoints for the months of July, August, and September with different control strategies. Greenhouse internal temperature, water vapor density, and CO_2 density are measured to compute the percentage deviation from setpoints and constraint violations.

	CEMPC		RMPC		DDPG		RO-DRL	
	Energy (kW)	Deviation (%)	Energy (kW)	Deviation (%)	Energy (kW)	Deviation (%)	Energy (kW)	Deviation (%)
Heating/Cooling	42,920	22.00	36,772	20.30	62,611	39.61	7,504	6.88
	42,135	21.29	37,987	20.60	328,162	194.08	25,144	12.39
	34,393	27.66	27,836	27.41	345,593	224.51	16,872	17.82
Humidi-fication	0.05	12.98	1.11	13.76	5.42	40.44	1.77	7.27
	0.03	8.67	1.45	13.24	22.43	171.01	1.12	9.69
	0.17	10.29	3.08	26.65	28.16	232.37	0.67	7.58
CO_2 enrichment*	377.3	86.02	911.9	12.02	1353.9	87.48	652.3	0.03
	374.0	84.05	920.1	21.57	1736.6	100	732.9	1.38
	348.5	81.80	897.8	7.65	1735.8	100	704.9	0

*Since CO_2 levels are maintained between a specific interval of 600-1100ppm, the value of percent deviations indicates the average constraint violation percentage for each month.

crop growth. Similarly, such average values for the three months of the growing cycle from July to September are also provided in Table A3. Table A3 also reports the percentage deviation of greenhouse climate variables from the setpoints for air temperature and water vapor density in the greenhouse achieved with different control strategies. Percentage constraint violations are reported for CO_2 density in the greenhouse as regulating CO_2 levels within specific thresholds is required.

References

[1] Gilbert N. One-third of our greenhouse gas emissions come from agriculture. Nature 2012.
[2] Alexandratos N, Bruinsma J. World agriculture towards 2030/2050. Food and Agriculture Organization of the United Nations; 2012.
[3] Castilla N. Greenhouse technology and management. Cabi; 2013.
[4] Iddio E, Wang L, Thomas Y, McMorrow G, Denzer A. Energy efficient operation and modeling for greenhouses: A literature review. Renewable and Sustainable Energy Reviews 2020;117:109480.
[5] Hu G, You F. Renewable energy-powered semi-closed greenhouse for sustainable crop production using model predictive control and machine learning for energy management. Renewable and Sustainable Energy Reviews 2022;168:112790.
[6] Taki M, Rohani A, Rahmati-Joneidabad M. Solar thermal simulation and applications in greenhouse. Information Processing in Agriculture 2018;5(1):83–113.
[7] Nicholson CF, Harbick K, Gómez MI, Mattson NS. An economic and environmental comparison of conventional and controlled environment agriculture (CEA) supply chains for leaf lettuce to US cities. In: Food supply chains in cities. Springer; 2020. p. 33–68.
[8] Ahamed MS, Guo H, Taylor L, Tanino K. Heating demand and economic feasibility analysis for year-round vegetable production in Canadian Prairies greenhouses. Information Processing in Agriculture 2019;6(1):81–90.
[9] Shamshiri R. A Review of Greenhouse Climate Control and Automation Systems in Tropical Regions. Journal of Agricultural Science and Applications 2013;2:175–82.
[10] Zhang S, Guo Y, Zhao H, Wang Y, Chow D, Fang Y. Methodologies of control strategies for improving energy efficiency in agricultural greenhouses. Journal of Cleaner Production 2020;274:122695.
[11] Hemming S, de Zwart F, Elings A, Righini I, Petropoulou A. Remote control of greenhouse vegetable production with artificial intelligence—greenhouse climate, irrigation, and crop production. Sensors 2019;19(8):1807.
[12] González-Briones A, Chamoso P, Yoe H, Corchado JM. GreenVMAS: virtual organization based platform for heating greenhouses using waste energy from power plants. Sensors 2018;18(3):861.
[13] Hemming S, Zwart Fd, Elings A, Petropoulou A, Righini I. Cherry tomato production in intelligent greenhouses—Sensors and AI for control of climate, irrigation, crop yield, and quality. Sensors 2020;20(22):6430.
[14] Hu H, Xu L, Zhu B, Wei R. A compatible control algorithm for greenhouse environment control based on MOCC strategy. Sensors 2011;11(3):3281–302.
[15] Hu H, Xu L, Goodman ED, Zeng S. NSGA-II-based nonlinear PID controller tuning of greenhouse climate for reducing costs and improving performances. Neural Computing and Applications 2014;24(3):927–36.
[16] Su Y, Yu Q, Zeng L. Parameter self-tuning pid control for greenhouse climate control problem. IEEE Access 2020;8:186157–71.
[17] El Ghoumari MY, Tantau HJ, Serrano J. Non-linear constrained MPC: Real-time implementation of greenhouse air temperature control. Computers and Electronics in Agriculture 2005;49(3):345–56.
[18] 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):330–44.
[19] Gil JD, Álvarez JD, Roca L, Sánchez-Molina JA, Berenguel M, Rodríguez F. Optimal thermal energy management of a distributed energy system comprising a solar membrane distillation plant and a greenhouse. Energy Conversion and Management 2019;198:111791.
[20] Ferreira PM, Ruano AE. Discrete model-based greenhouse environmental control using the branch & bound algorithm. IFAC Proceedings Volumes 2008;41(2):2937–43.

[21] Fitz-Rodríguez E, Kubota C, Giacomelli GA, Tignor ME, Wilson SB, McMahon M. Dynamic modeling and simulation of greenhouse environments under several scenarios: A web-based application. Computers and electronics in agriculture 2010;70(1):105–16.
[22] Chen M, et al. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. Agricultural Water Management 2021;250:106838.
[23] Bu F, Wang X. A smart agriculture IoT system based on deep reinforcement learning. Future Generation Computer Systems 2019;99:500–7.
[24] Gautron R, Maillard O-A, Preux P, Corbeels M, Sabbadin R. Reinforcement learning for crop management support: Review, prospects and challenges. Computers and Electronics in Agriculture 2022;200:107182.
[25] L. Wang, X. He, and D. Luo, "Deep reinforcement learning for greenhouse climate control," in *2020 IEEE International Conference on Knowledge Graph (ICKG)*, 9-11 Aug. 2020 2020, pp. 474-480, doi:10.1109/ICBK50248.2020.00073.
[26] Chen J, Yang J, Zhao J, Xu F, Shen Z, Zhang L. Energy demand forecasting of the greenhouses using nonlinear models based on model optimized prediction method. Neurocomputing 2016;174:1087–100.
[27] Chen W-H, You F. Semiclosed greenhouse climate control under uncertainty via machine learning and data-driven robust model predictive control. IEEE Transactions on Control Systems Technology 2021;30(3):1186–97.
[28] Chen L, Du S, He Y, Liang M, Xu D. Robust model predictive control for greenhouse temperature based on particle swarm optimization. Information Processing in Agriculture 2018;5(3):329–38.
[29] Chen W-H, You F. Smart greenhouse control under harsh climate conditions based on data-driven robust model predictive control with principal component analysis and kernel density estimation. Journal of Process Control 2021;107:103–13.
[30] Pasgianos GD, Arvanitis KG, Polycarpou P, Sigrimis N. A nonlinear feedback technique for greenhouse environmental control. Computers and Electronics in Agriculture 2003;40(1):153–77.
[31] Chen W-H, Mattson NS, You F. Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi–closed greenhouse. Applied Energy 2022;320:119334.
[32] Roos E, den Hertog D. Reducing conservatism in robust optimization. INFORMS Journal on Computing 2020;32(4):1109–27.
[33] Vanthoor BH. A model-based greenhouse design method. Wageningen University and Research; 2011.
[34] EECi/GES: First release of GES software. (2019). Zenodo. [Online]. Available: https://www.eeci.cam.ac.uk/research/greenhouse-energy-simulation.
[35] S. Le Quillec, B. Albert, D. Lesourd, D. Loda, R. Barette, and E. Brajeul, "Benefits of a semi-closed greenhouse for tomato production in the West of France," in *International Symposium on New Technologies and Management for Greenhouses-GreenSys2015 1170*, 2015, pp. 883-888.
[36] Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018.
[37] Mnih V, et al. Human-level control through deep reinforcement learning. Nature 2015;518(7540):529–33.
[38] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.
[39] M. Ryu, Y. Chow, R. Anderson, C. Tjandraatmadja, and C. Boutilier, "CAQL: Continuous action Q-learning," arXiv preprint arXiv:1909.12397, 2019.
[40] S. Fujimoto, D. Meger, and D. Precup, "Off-policy deep reinforcement learning without exploration," in *International Conference on Machine Learning*, 2019: PMLR, pp. 2052-2062.
[41] Fischetti M, Jo J. Deep neural networks and mixed integer linear optimization. Constraints 2018;23(3):296–309.
[42] Li Z, Ding R, Floudas CA. A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization. Industrial & engineering chemistry research 2011;50(18):10567–603.
[43] Ning C, You F. Data-Driven Adaptive Nested Robust Optimization: General Modeling Framework and Efficient Computational Algorithm for Decision Making under Uncertainty. AIChE Journal 2017;63:3790–817.
[44] Buschermohle MJ, Grandle GF. Controlling the environment in greenhouses used for tomato production. Agricultural Extension Service, The University of Tennessee 2002:1–10.
[45] Shamshiri RR, Jones JW, Thorp KR, Ahmad D, Man HC, Taheri S. Review of optimum temperature, humidity, and vapour pressure deficit for microclimate evaluation and control in greenhouse cultivation of tomato: a review. International agrophysics 2018;32(2):287–302.
[46] Ljung L. Perspectives on system identification. Annual Reviews in Control 2010;34(1):1–12.
[47] Energy use and savings in greenhouses and growth chambers [Online] Available: cals.cornell.edu/agricultural-experiment-station/greenhouses-growth-chambers/sustainable-greenhouses-growth-chambers/energy-use-savings-greenhouses-growth-chambers.
[48] García ML, Medrano E, Sánchez-Guerrero MC, Lorenzo P. Climatic effects of two cooling systems in greenhouses in the Mediterranean area: External mobile shading and fog system. Biosystems Engineering 2011;108(2):133–43.
[49] M. Dorais, "The use of supplemental lighting for vegetable crop production: light intensity, crop response, nutrition, crop management, cultural practices," in *Canadian Greenhouse Conference*, 2003, vol. 9.
[50] K. Harbick, L. D Albright, and N. S Mattson, "Electrical savings comparison of supplemental lighting control systems in greenhouse environments," presented at the 2016 ASABE Annual International Meeting, St. Joseph, MI, 2016. [Online]. Available: https://elibrary.asabe.org/abstract.asp?aid=47203&t=5.
[51] Y. Li, "Deep reinforcement learning: An overview," 2017.
[52] Watkins CJCH, Dayan P. Q-learning. Machine Learning 1992;8(3):279–92.
[53] H. v. Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-Learning," presented at the Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, Arizona, 2016.
[54] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," arXiv preprint arXiv:1509.02971, 2015.