

Article

Not peer-reviewed version

Enhancing Greenhouse Efficiency: Integrating IoT and Reinforcement Learning for Optimised Climate Control

[Manuel Platero-Horcajadas](#) , [Sofia Pardo-Pina](#) , [José-María Cámara-Zapata](#) , [Jose A Brenes](#) ,
[Francisco-Javier Ferrández-Pastor](#) *

Posted Date: 30 July 2024

doi: 10.20944/preprints202407.2247.v1

Keywords: smart agriculture; reinforcement learning; IoT; greenhouse energy management



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Enhancing Greenhouse Efficiency: Integrating IoT and Reinforcement Learning for Optimised Climate Control

Manuel Platero-Horcajadas ¹, José-M. Cámara-Zapata ², Sofia Pardo-Pina ², José-Antonio Brenes-Carranza ³ and Francisco-Javier Ferrández-Pastor ^{1,*} 

¹ I2RC, University of Alicante, 03690 Alicante Spain, fjferran@ua.es

² Institute for Agri-Food and Agro-Environmental Research and Innovation, Miguel Hernandez University, CIAGRO-UMH, Ctra Beniel, sn, 03312 Orihuela, Spain; jm.camara@umh.es

³ University of Costa Rica, San José, Costa Rica joseantonio.brenes@ucr.ac.cr

* Correspondence: fjferran@ua.es

Abstract: Automated systems, regulated by algorithmic protocols and predefined set-points for feedback control, require the oversight and fine-tuning of skilled technicians. This necessity is particularly pronounced in automated greenhouses, where optimal environmental conditions depend on the specialized knowledge of dedicated technicians, emphasizing the need for expert involvement during installation and maintenance. To address these challenges, this study proposes the integration of data acquisition technologies using Internet of Things (IoT) protocols and optimization services via Reinforcement Learning (RL) methodologies. The proposed model was tested in an industrial production greenhouse for the cultivation of industrial hemp, applying adapted strategies to the crop, and was guided by an agronomic technician knowledgeable about the plant. The expertise of this technician was crucial in transferring the RL model to a real-world automated greenhouse equipped with IoT technology. The study concludes that the integration of IoT and RL technologies is effective, validating the model's ability to manage and optimize greenhouse operations efficiently and adapt to different types of crops. Moreover, this integration not only enhances operational efficiency but also reduces the need for constant human intervention, thereby minimizing labor costs and increasing scalability for larger agricultural enterprises. Furthermore, the RL-based control has demonstrated its ability to maintain selected temperatures and achieve energy savings compared to classical control methods.

Keywords: smart agriculture; reinforcement learning; IoT; greenhouse energy management

1. Introduction

Automated greenhouse systems require the meticulous configuration of diverse parameters to guarantee that the actuators execute the designated functions accurately. This configuration entails the determination of appropriate set-points for various subsystems including climate, lighting, and irrigation control. Typically, these parameters and control rules are reactive. In recent years, numerous studies have advanced models based on set-point selection strategies employing various heuristic or artificial intelligence paradigms, coupled with the implementation of predictive solutions. The application of reinforcement learning (RL) to greenhouse management has gained significant attention in recent years due to its potential to optimize energy usage and automate various control processes. This section provides an overview of the most relevant studies, categorized by their approaches and contributions.

2. Literature Review

The use of reinforcement learning (RL) in the energy management of greenhouses and precision agriculture techniques has evolved significantly. This analysis is based on the references provided, which highlight key contributions in this field. Future research directions are also discussed.

Reinforcement learning (RL) has become a powerful tool in optimizing energy management in greenhouses and precision agriculture. This paper reviews key contributions to the field, analyses their contributions, and discusses future research directions.

Kiumarsi et al. [1] present a comprehensive survey and implementation guidelines for optimal and autonomous control using RL. They discuss applications in complex systems, emphasizing their potential to enhance energy efficiency and reduce operational costs in greenhouses. Perera et al. [2] review various applications of RL in energy management, from generation to consumption. They highlight specific cases of intelligent greenhouses that use RL to optimize energy use. Wang et al. [3] present an RL controller based on recurring neural networks based on long-term memory (LSTM) for microgrid management, with potential applications in greenhouses.

Kazmi et al. [4] explore the use of deep RL for optimal control in hybrid energy systems of buildings, applicable to energy management in greenhouses. Ruelens et al. [5] examine the application of RL in electric water heaters, with direct implications for energy management in greenhouses.

Zhang et al. [6] discuss the developments and future challenges in precision agriculture, including the integration of RL for better resource management in greenhouses. Liu et al. [6] provide a comprehensive review of RL applications in smart agriculture, highlighting various use cases and the potential benefits of RL in optimizing agricultural processes, including greenhouse energy management.

Mason et al. [7] review the applications of RL in smart grids, discussing how these methods can be applied to improve energy management and efficiency in interconnected systems such as greenhouses. Hosseinloo et al. [8] explore data-driven predictive control using RL for energy efficiency and comfort management in buildings, with potential applications in greenhouses to optimize climate control systems. Sun et al. [9] present a study on multi-agent RL for integrated energy management of interconnected microgrids, applicable to complex greenhouse energy systems. Alani et al. [10] discuss the opportunities and challenges of RL-based energy management for smart homes and buildings, with insights applicable to greenhouse energy systems.

Fu et al. [11] survey the applications of RL in building energy management, providing information relevant to greenhouse energy systems. Mauree et al. [12] review data-driven and machine learning models for building energy performance prediction, fault detection, and optimization, applicable to greenhouse energy management. Kazmi et al. [13] discuss the application of multi-agent RL for building energy management, with potential benefits for greenhouse systems. Yang et al. [14] explore energy optimization in smart home buildings using deep RL, providing information for greenhouse energy management.

Ruelens et al. [15] discuss learning sequential decision making for optimal control of thermally activated resources, applicable to greenhouse energy systems. Vázquez-Canteli and Henze [16] discuss the integration of reinforcement learning with predictive control of the model for the response to demand in buildings, which can be applied to optimize energy management in greenhouses. Sutton and Barto [17] discuss the application of RL in real-world games, providing insights that can be adapted to complex energy management systems in greenhouses. Peters et al. [18] discuss a reinforcement learning approach to autonomous vehicles, which can provide insights for autonomous control in greenhouse energy systems. Kar et al. [19] present QD-learning, a collaborative Q-learning approach that can be applied to cooperative energy management strategies in greenhouses.

Sierla et al. [20] review reinforcement learning applications in urban energy systems, which can provide valuable information for energy management in greenhouse environments. Vázquez-Canteli and Nagy [21] discuss reinforcement learning for demand response, which is highly relevant for dynamic energy management in greenhouses. Mauree et al. [12] review assessment methods for urban environments, providing methodologies that can be adapted to evaluate energy performance in greenhouses.

2.1. Future Research Directions

Despite significant advancements, several areas require further exploration to fully implement RL in greenhouse energy management.

- **Integration with IoT:** The integration of RL with IoT devices can enhance real-time data acquisition and decision-making processes in greenhouses. Future research should focus on developing seamless IoT-RL integration frameworks.
- **Scalability:** Research on scaling RL solutions to larger, more complex greenhouse systems is necessary to ensure widespread adoption. Studies should address computational challenges and the ability to handle large datasets.
- **Interdisciplinary Approaches:** Combining RL with other AI techniques, such as genetic algorithms and fuzzy logic, could yield more robust energy management solutions. Exploration of hybrid models that leverage the strengths of different AI paradigms is essential.
- **Environmental Adaptability:** Developing RL algorithms capable of adapting to diverse environmental conditions will be crucial for global applications. This includes designing algorithms that can learn and adapt to changing weather patterns, pest infestations, and other environmental variables.
- **Economic Viability:** Studies on the cost-effectiveness of RL implementations in greenhouses can drive commercial interest and investment. Future research should focus on performing cost-benefit analyses and developing business models that highlight the economic advantages of RL-based energy management systems.
- **User-Friendly Interfaces:** Developing user-friendly interfaces and control systems for greenhouse operators is vital for the practical implementation of RL. Research should focus on creating intuitive dashboards and control panels that allow operators to easily interact with and oversee RL systems.
- **Sustainability Metrics:** Future work should also explore the development of sustainability metrics that RL systems can optimize. This includes not only energy efficiency but also water usage, pesticide application, and overall environmental impact.
- **Policy and Regulatory Compliance:** Research should address how RL systems can be designed to comply with local and international policies and regulations concerning energy usage and environmental protection.
- **Data Privacy and Security:** With the increasing use of IoT and RL, ensuring data privacy and security is essential. Future research should develop robust security protocols to protect sensitive data in greenhouse management systems.
- **Real-World Case Studies:** Conducting real-world case studies and pilot projects can provide valuable insights into the practical challenges and benefits of implementing RL in greenhouses. These studies can help refine RL models and identify best practices for successful adoption.

2.2. Literature review conclusion

The RL paradigm has shown significant potential to optimize energy management in greenhouses and precision agriculture techniques. From optimal autonomous control to deep RL, the reviewed references indicate a growing trend towards intelligent, adaptive solutions that promote energy efficiency and sustainability. Future research should focus on integrating RL with IoT, scalability, interdisciplinary approaches, environmental adaptability, economic viability, user-friendly interfaces, sustainability metrics, policy compliance, data privacy and security, and real-world case studies to further advance this technological evolution.

In conclusion, the field of reinforcement learning for greenhouse management has made significant progress, yet challenges remain. Future research should focus on improving scalability, integrating advanced technologies, and developing hybrid models to fully realize the potential of RL in this domain.

3. Model proposed

The Reinforcement Learning (RL) problem involves a digital agent exploring an environment to achieve a specific goal. In the field of automated greenhouses, it is about managing environmental conditions and crop growth while optimizing resources and inputs. RL is based on the hypothesis that all goals can be characterized by maximizing the expected cumulative reward. The agent must learn to perceive and manipulate the state of the environment through its actions to optimize this reward. In the model presented in this work, the agent performs control actions to adjust the set points, ensuring that the internal conditions of the greenhouse stay within the predefined maximum and minimum limits. By doing so, the agent optimizes the management of resources such as water, energy, and inputs, leading to improved plant growth. Figure 1 shows the application scenario of the RL paradigm.

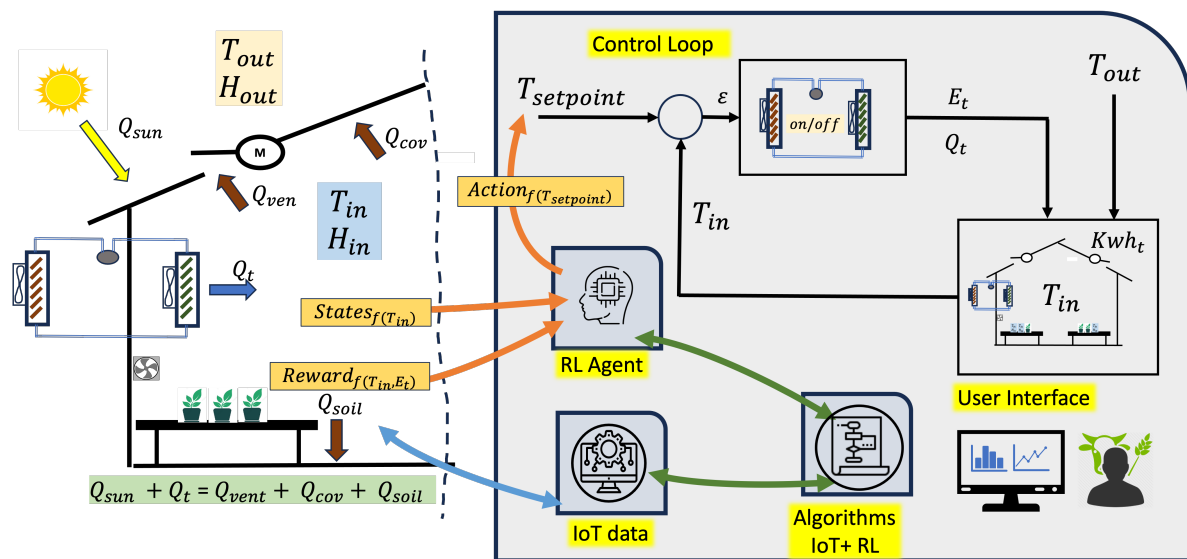


Figure 1. RL paradigm in greenhouse proposed

The system must be able to collect relevant data, establish the necessary sensors and communication protocols, and manage them appropriately using the IOT paradigm.

The proposed model uses a layered architecture (Figure 2) to integrate the services and functionalities of the platform (IoT + RL). In an agricultural facility, whether newly constructed or already operational, monitoring and actuation devices are installed and interfaced with the processing and control layer. The following provides a detailed description of each layer:

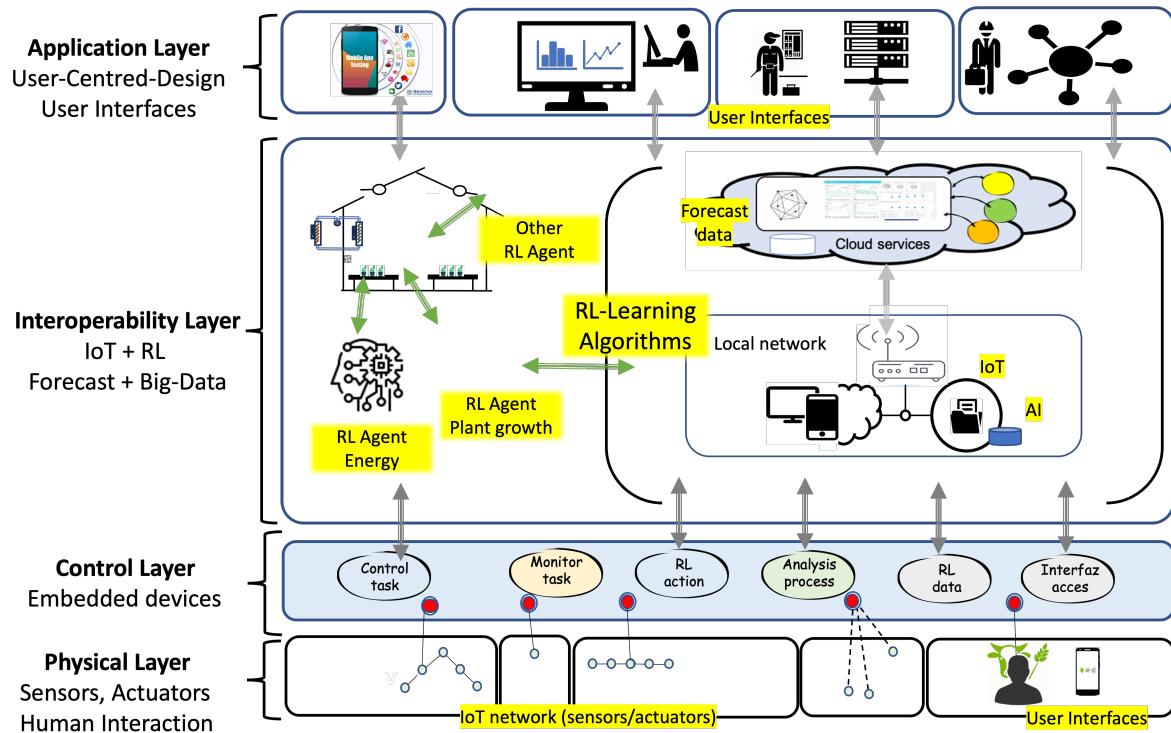


Figure 2. Layered technological architecture. Relationship between IoT, RL, digital platform and different interfaces

In the **physical layer**, IoT sensors/actuators, human-machine interfaces (HMIs) and machine-to-machine interfaces (MMI) are installed. The farmer and technicians can act by entering data and requests. The processing layer filters data, executes control actions, and communicates with the upper layer.

The **control layer** embedded devices receive data from the upper layer to manage the different processes and perform control and maintenance actions. This layer must provide the necessary support to communicate the data to the sensor/actuators installed.

The **interoperability layer** integrates different RL agents connected to digital twin algorithms. The model proposes the development of RL agents to optimize the use of set-point values in various control loops within an automated facility. The RL agent modifies the set-point values to achieve one or more specified objectives, such as maintaining environmental conditions, improving electrical consumption, reducing water use, and optimizing renewable energy use. This process operates automatically and can be analyzed through an interface where users can test control strategies. This interface enables simulations on a digital twin model that replicates different states based on the knowledge acquired from the analyzed data obtained from IoT network sensors. In this layer, the deployment of RL agents, the development of digital twin functionalities, and the implementation of user interfaces are executed.

In the **application layer**, applications are designed and developed on different platforms (mobile phones, business networks, computers, etc.). In each of them, a relationship is defined between the user and the type of access allowed.

3.1. Development in an Automated Greenhouse

Greenhouses provide a controlled environment for plant cultivation, allowing for improved growth and productivity compared to traditional open-field agriculture. However, managing a greenhouse efficiently involves complex decision-making processes to balance resource usage, such as water and energy, with the optimal growing conditions for plants. Traditional management methods

often rely on predefined schedules and heuristics, which may not adapt well to dynamic environmental conditions and changing plant needs.

Reinforcement learning (RL) offers a promising solution to this challenge by enabling systems to learn optimal strategies through interactions with the environment. RL algorithms can adapt to changing conditions and learn from experience, making them well suited to the dynamic and complex nature of greenhouse management. By continuously adjusting actions based on observed outcomes, RL can optimize resource usage while maintaining or improving crop yields.

One of the scenarios where the model can be applied is in energy management. In the case of use of this work, the model is based on a reinforcement learning algorithm in greenhouse energy management. Specifically, the model employs a prediction model to forecast the greenhouse’s environmental conditions throughout the day. Based on these predictions, they implement a reinforcement learning (RL) algorithm that rewards minimal energy usage to regulate the greenhouse’s temperature. The algorithm refines the temperature control loop by optimizing the choice of the set-point input for the existing regulation loop. In essence, the algorithm adjusts the set point value to optimize energy consumption while maintaining the selected minimum and maximum temperature thresholds.

Incorporating climate predictions into the RL model can further enhance its performance by allowing anticipatory adjustments.

Reinforcement learning (RL) offers a promising approach to energy consumption by learning from the environment and making data-driven decisions. This study focuses on the Q-Learning algorithm, a model-free RL method, to determine the connection and disconnection of the air conditioning system.

3.2. Methodology

The proposed system uses the Q-Learning algorithm to make watering decisions based on real-time data from IoT sensors. The states, variables, actions, and reward function are designed to reflect the dynamics of the greenhouse environment. The methods are indicated in Table 1

Table 1. This is a table caption.

ine Phase	Description
ine Phase 1	Control strategies. Analysis of environmental and control variables with the agronomic expert. Examination of all possible strategies that can be part of optimization. Strategies for choosing set-points.
ine Phase 2	IoT infrastructure. Design of the IoT infrastructure needed for the greenhouse to carry out the control and data set generation. Deployment of various embedded systems interconnected with the required IoT technologies.
ine Phase 3	Data-set generation. Sensors generate data that are analyzed to determine greenhouse behavior models. Each greenhouse has specific characteristics that must be taken into account when applying the model. Data set are captured in normalized format (csv, json)
ine Phase 4	Digital model and an RL algorithm. The objective is to maintain the greenhouse temperature between the minimum and maximum limits by optimizing the connection and disconnect of the air conditioning system. Implementation by comparing all strategies with those proposed in the model to determine their effectiveness. To act on the installation, a digital model is previously created on which the RL algorithm begins its calculations. This model will be adjusted to the reality of the behavior in the greenhouse application. Theoretical analysis validates reward strategies and policies that will be applied in an applied way.
ine Phase 5	Training and Evaluation Train the RL agent using the constructed data set and the results of the analysis, iteratively updating the policy based on observed rewards and state transitions. With the data set obtained with the capture of IoT data and the results of the theoretical simulations, the policies, reward functions, and actions most appropriate to the type of greenhouse and installation are promoted.
ine	

4. Phase 1. Control Strategies (set point in the environmental regulation loop)

Reinforcement Learning (RL) techniques can optimize air conditioning connection strategies by making decisions based on continuous feedback from plants and the environment. Various sensors (energy consumption, temperature, humidity, weather data, etc.), actuators, and a data processing unit must be used. In Table 2 different strategies and their relationship with IoT technologies are shown.

Table 2. Energy Control Strategies. Standard control vs. RL control.

Strategy	Description	IoT sensors
ine Set-point selec- tion (standard con- trol)	This is the simplest and most commonly used strategy. The technician selects the temperature, and the maximum and min- imum values	temperature sensors.
ine Set-point adjust- ment with RL algo- rithm (RL actions) to optimize energy con- sumption	The values assigned at the set point are adjusted and modified by predicting expected conditions in the greenhouse. These changes are made at scheduled sampling times	Temperature, en- ergy consumption, weather forecast, and temperature prediction inside the greenhouse.

The selection of the appropriate irrigation strategy for greenhouse cultivation depends on various factors, including crop type, environmental conditions, and resource availability. Strategies based on substrate moisture, evapotranspiration, and VPD offer customized approaches to optimize water use, improve crop performance, and promote sustainable agricultural practices.

5. Phase 2. IoT infrastructure

The IoT infrastructure is designed with the strategy of being interoperable with existing green-
house subsystems: climate, lighting, irrigation, etc. To achieve this, it is designed and developed at
two levels.

The Figure 3 describes the Internet of Things (IoT) network. Here is the identification of the
components according to the numbers in the image: The sequence of events and data sequence are
listed below,

- Climate energy data is captured and stored in the system.
- Sensors collect environmental conditions.
- The data is sent to the gateway, which then transmits it to the local server and cloud services.
- The data is stored on the local server. The datasets are created with the main variables
- Local and cloud services allow remote monitoring and control of the system.
- The prediction of the environmental conditions in the next few hours is obtained
- With the data and predictions the learning algorithm is executed
- The algorithm proposes a modification of the set-point by increasing or decreasing its value,
using the values of the reinforcements or Q values

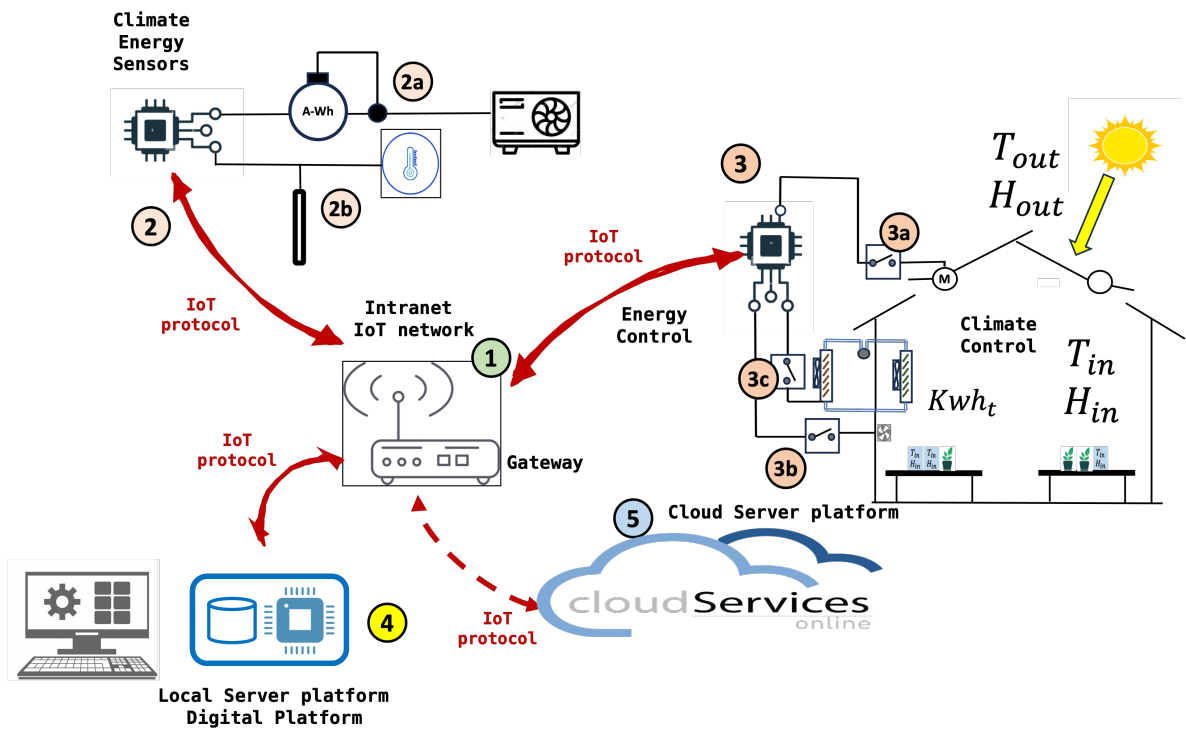


Figure 3. IoT platform. Basic infrastructure for data capture, analysis and management services

Table 3. Identification of components in the IoT network for a solar energy management system.

Number	Component Description
1	IoT Gateway or Router: This device acts as a central communication point between various sensors and devices and the cloud server. It uses IoT protocols to transmit data.
2	Sensors and Meters: These devices collect data from different sources:
2a	Energy meter (A-Wh) that measures the amount of energy consumed.
2b	Temperature sensor (thermostat) that measures the ambient temperature.
3	Solar Energy Controller: This component receives data from the sensors and manages the distribution of energy.
3a	Inverter that converts solar energy from direct current (DC) to alternating current (AC).
3b	Batteries for energy storage.
3c	Switches and fuses for protection.
4	Local Server or Database: Stores and processes data locally. It is where all the data is collected for processing before being sent to the cloud.
5	Cloud Services: The data is sent to the cloud for additional storage and analysis. Cloud services can provide interfaces to monitor and control the system.
M	Monitoring Computer: Allows users to interact with the system, probably through a graphical interface for real-time monitoring and control of the system.

6. Phase 3. Data set generation

Data sets generated from IoT sensors provide a rich source of information for decision makers. These data sets enable real-time monitoring and analysis, allowing for informed decision-making based on accurate and up-to-date information. In greenhouse facilities, IoT sensor traffic data can be used to optimize traffic flow, reduce congestion, and improve public transportation systems. Generating high-

quality datasets from IoT sensor data requires addressing issues related to data quality and consistency. Sensors can produce noisy or incomplete data, which can affect the accuracy and reliability of the data sets. Implementing robust data pre-processing and cleaning techniques is essential to ensure the integrity of the data sets.

Temperature and Humidity Sensors: inside and outside the greenhouse. Weather forecast and energy consumption. The Methodology for Data Capture is:

- **Sensor Placement.** Strategically place sensors to capture a representative sample of environmental conditions within the greenhouse. This includes placing sensors at various depths and locations throughout the greenhouse to monitor microclimates.
- **Data Logging and Transmission.** Utilise data loggers and wireless networks to ensure continuous data capture and transmission. This includes setting up a reliable network infrastructure that can handle the data volume and frequency required for RL applications.
- **Data storage.** Implement a centralized data storage solution, preferably cloud-based, to store the large volumes of data generated by the sensors. Ensure the storage system supports efficient data retrieval and processing.
- **Data preprocessing: cleaning, and normalisation.** Address issues related to missing values, sensor malfunctions, and noise in the data. Techniques such as interpolation and filtering should be applied to ensure data quality and consistency. Normalise sensor data to a common scale to facilitate accurate analysis and model training. This step is crucial for integrating diverse data types into a cohesive dataset.

6.1. *Materials and methods*

Data capture was carried out on a hemp crop in a technical greenhouse (Figure 4) with the following characteristics: Glass greenhouse of 50 m² (5 x 10 m) each. All of them have an automatic climate and irrigation control system with the following equipment:

- Humidifier with osmosis water mist.
- Air conditioners to heat and cool the modules. Twin, Triple Mitsubishi PUHZ-P200YKA Three-phase Classic Inverter Nominal cooling capacity (Min. - Max.) kW 19.00 Nominal heat capacity (Min. - Max.) kW 22.40
- Thermal shading screen.
- Extractor fan and zenithal opening windows with anti-trip mesh.
- Artificial light lamps to increase net assimilation.
- Micro-sprinkler and drip or flood irrigation system.
- Temperature and humidity probes.
- Electrical, compressed air, mains water and osmosed water connections.
- Embedded device (raspberrypi4) that deploys an intranet (WiFi, Bluetooth Low Energy) for communication, monitoring, and control.
- Electric energy meter in three-phase and single-phase circuits (shelly 3 EM). This consumption meter communicates with the embedded system through the WiFi and IP protocol.
- Communication to Web servers to obtain the temperature prediction in the greenhouse area.

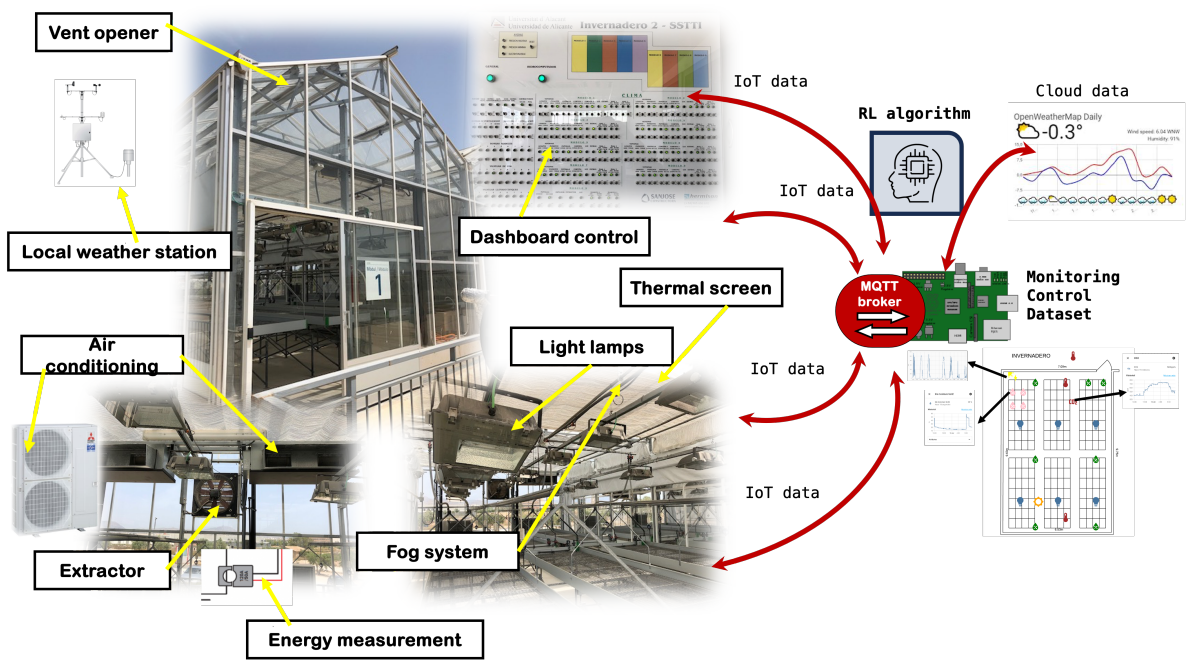


Figure 4. Industrial greenhouse used to deploy the IoT data. A control intranet has been deployed with embedded devices and communications based on IoT protocols (WiFi and MQTT)

This work develops and tests a low-cost sensor/actuator network platform, based on the Internet of Things, integrating machine-to-machine and human-machine interface protocols used in [22]. The system integrated Internet of Things (IoT) paradigms, ubiquitous sensor networks, and edge computing to create a smart agricultural environment. Implementation (Figure ??) included machine-to-machine and human-machine interface protocols to enable seamless communication and control processes considering a precision agriculture scenario.

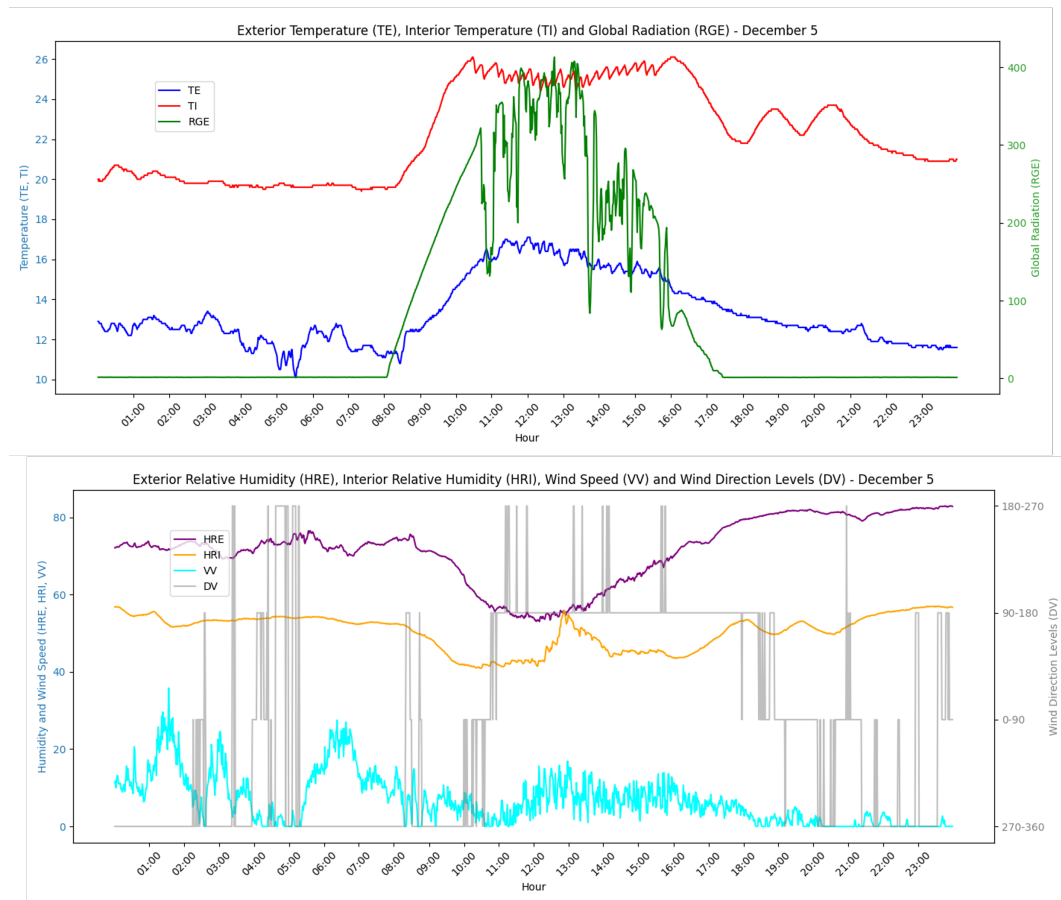


Figure 5. Example of daily data obtained in the dataset. The data is obtained daily every minute and stored in a dataset to obtain temperature prediction models inside the greenhouse.

7. Phase 4. Digital model and RL algorithm

RL involves an agent learning to make sequential decisions through interaction with its environment to maximize cumulative rewards. In the context of greenhouse control, RL offers a promising framework to optimize energy use while maintaining the configured maximum and minimum temperature conditions. The theoretical model serves to support decision making at a given moment. The objective is to start from prior knowledge that allows the system to initiate decision making more effectively. From the initial configuration of the RL algorithm, the system will adapt to the specific conditions of the installation, optimizing control as feedback information is obtained on the evolution of the system, compared to its previous behavior. It is therefore about introducing a finer regulation that is capable of improving the existing one by adapting to the specific conditions of the installation.

7.1. Greenhouse model based in differential equations.

There is knowledge and mathematical models of the behavior of temperature in a greenhouse. These models provide a good starting point for defining reward policies and functions. For all these reasons, the RL model is based on the knowledge already acquired to, from there, introduce improvements and optimize current operation. The behavior of the temperature inside a greenhouse is influenced by several factors, including the external temperature, solar radiation, and the thermal dynamics of the greenhouse itself. Here, we present a mathematical model to simulate the dynamics of the greenhouse temperature.

The thermal model of a greenhouse when both heating and cooling are applied can be described by the following differential equation.

$$C_{\text{greenhouse}} \frac{dT_{\text{in}}(t)}{dt} = Q_{\text{solar}}(t) + Q_{\text{heat}}(t) - Q_{\text{cool}}(t) - UA(T_{\text{in}}(t) - T_{\text{out}}(t)) \quad (1)$$

To discretize this equation for algorithmic application, we consider a time step Δt . The temperature change over the time step from t to $t + \Delta t$ can be approximated as:

$$\frac{dT_{\text{in}}(t)}{dt} \approx \frac{T_{\text{in}}(t + \Delta t) - T_{\text{in}}(t)}{\Delta t} \quad (2)$$

Substituting this into the differential equation gives:

$$C_{\text{greenhouse}} \frac{T_{\text{in}}(t + \Delta t) - T_{\text{in}}(t)}{\Delta t} = Q_{\text{solar}}(t) + Q_{\text{heat}}(t) - Q_{\text{cool}}(t) - UA(T_{\text{in}}(t) - T_{\text{out}}(t)) \quad (3)$$

Solving for $T_{\text{in}}(t + \Delta t)$, we get the following:

$$T_{\text{in}}(t + \Delta t) = T_{\text{in}}(t) + \frac{\Delta t}{C_{\text{greenhouse}}} [Q_{\text{solar}}(t) + Q_{\text{heat}}(t) - Q_{\text{cool}}(t) - UA(T_{\text{in}}(t) - T_{\text{out}}(t))] \quad (4)$$

where:

- $C_{\text{greenhouse}}$ is the heat capacity of the greenhouse (J/°C),
- $T_{\text{in}}(t)$ is the inside temperature of the greenhouse at time t (°C),
- $T_{\text{in}}(t + \Delta t)$ is the inside temperature of the greenhouse at time $t + \Delta t$ (°C),
- $T_{\text{out}}(t)$ is the outside temperature at time t (°C),
- $Q_{\text{solar}}(t)$ is the solar radiation entering the greenhouse at time t (W),
- $Q_{\text{heat}}(t)$ is the heating power applied to the greenhouse at time t (W),
- $Q_{\text{cool}}(t)$ is the cooling power applied to the greenhouse at time t (W),
- U is the overall heat transfer coefficient (W/m²°C),
- A is the surface area of the greenhouse (m²),
- Δt is the time step (s).

This discretized equation can be used to iteratively compute the inside temperature of the greenhouse in discrete time steps for the purpose of applying control algorithms. This model is simple and the results give us a first analysis of trends and future strategies. In the scenario of an RL algorithm, the actions taken depend on the value of the set temperature. The actions that can be taken are shown in the Table 4.

Table 4. Actions example taken by RL algorithm using differential equations. Values can be modified depending on the needs of the installation (cooling or heating). The number of actions can also be modified.

ine Action 1	$T_{\text{setpoint}} = T_{\text{setpoint}} + 2$
ine Action 2	$T_{\text{setpoint}} = T_{\text{setpoint}} + 1$
ine Action 3	$T_{\text{setpoint}} = T_{\text{setpoint}}$
ine Action 4	$T_{\text{setpoint}} = T_{\text{setpoint}} - 1$
ine Action 5	$T_{\text{setpoint}} = T_{\text{setpoint}} - 2$
ine	

7.2. Greenhouse model based in predictions

This model has two parts. In the first case, the behavior of the greenhouse temperature in the next moments of time is predicted on the basis of IoT data and the machine learning paradigm. Once the future value of the temperature inside the greenhouse is calculated, the best possible action for regulation is taken.

Table 5. Actions example taken by RL algorithm using temperature (T_{in}) prediction. The control values are limited by T_{max} and T_{min} . Values can be modified depending on the needs of the installation (cooling or heating). The number of actions can also be modified.

ine Action 1	$T_{setpoint} = T_{in_predicted} + 2$
ine Action 2	$T_{setpoint} = T_{in_predicted} + 1$
ine Action 3	$T_{setpoint} = T_{in_predicted}$
ine Action 4	$T_{setpoint} = T_{in_predicted} - 2$
ine Action 5	$T_{setpoint} = T_{in_predicted} - 1$
ine	

The interior temperature of a greenhouse based on external climatic conditions and historical can be obtained with several models, including linear regression, decision tree, gradient boost, random forest, and neural networks, can be compared to identify the most accurate and consistent predictive model.

The data set used consists of climatic data collected from a greenhouse over several months. The variables include:

- **TE:** Exterior Temperature
- **HRE:** Exterior Relative Humidity
- **RGE:** Exterior Global Radiation
- **VV:** Wind Speed
- **DV:** Wind Direction
- **LL:** Rainfall
- **TI:** Interior Temperature
- **HRI:** Interior Relative Humidity

The data is resampled to intervals that are configured according to the installation, and lag features are created to incorporate historical data into the prediction model.

The actions taken are shown in Table 7.

7.3. Reinforcement Learning deployment

Reinforcement Learning (RL) is a subfield of machine learning where an agent learns to make decisions by interacting with an environment to maximize a cumulative reward. In the context of controlling the temperature of a greenhouse, an RL agent can learn when to turn the heater on or off to maintain the desired temperature and minimize energy consumption.

The RL model consists of the following elements:

- **State (s):** Represents the current situation of the environment. In our case, the state can include the internal temperature of the greenhouse, the external temperature, and solar radiation.
- **Action (a):** The decision taken by the agent in each state. In our case, actions are turning the heater on ($a = 1$) or off ($a = 0$).
- **Reward (r):** The feedback received by the agent after taking an action in a state. The reward can be a function of the internal temperature and energy consumption.
- **Policy (π):** The strategy followed by the agent to make decisions. The policy maps states to actions.
- **Value ($V(s)$):** The expected value of the cumulative reward starting from state s following policy π .
- **Q-Value ($Q(s, a)$):** The expected value of the cumulative reward starting from state s taking action a and following policy π .

The goal of the RL agent is to learn an optimal policy π^* that maximizes the cumulative reward. This is formalized in the optimal control problem in a Markov Decision Process (MDP).

RL Model Formulas

State Value Function:

$$V^{\pi}(s) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s \right] \quad (5)$$

where $\gamma \in [0, 1)$ is the discount factor that weights the importance of future rewards.

Action Value Function (Q-Value):

$$Q^{\pi}(s, a) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s, a_0 = a \right] \quad (6)$$

Q-Value Update (Q-Learning):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (7)$$

where α is the learning rate.

8. Phase 5. Training

To start with prior knowledge and avoid the initial errors that RL introduces, actions, functions, and reward policies are analyzed first through simulation and then using real data obtained in the cultivation process. Firstly, an analysis is carried out using differential equations that model the climatic behavior in a greenhouse and then apply it with the data obtained in the greenhouse.

8.1. RL Algorithm application in greenhouse model based in differential equations model

The RL agent intervenes in the control process following the following steps:

1. **Observe Current State:** The agent observes the current state s_t , which includes the internal temperature (T_{in}), the external temperature (T_{out}), and solar radiation (Q_{solar}).
2. **Select Action:** Based on its policy π , the agent chooses an action a_t (turn the heater on or off).
3. **Apply Action:** The chosen action is applied to the environment.
4. **Observe Reward and Next State:** The agent receives a reward r_t and observes the next state s_{t+1} .
5. **Update Policy:** The agent updates its policy using the learning algorithm, such as Q-Learning.

By formulating temperature control as a sequential decision-making problem, RL algorithms can adaptively adjust based on real-time environmental data and set-point conditions. One crucial aspect in applying RL to temperature control is the choice of **state representation**, which captures relevant information about the environment for decision-making.

The process is based on the technician's previous configuration and a theoretical model of temperature behavior on which the RL algorithm performs the calculations. This model can be based on energy balances or be a model obtained by studying the behavior of greenhouse conditions.

Selecting an appropriate state representation is essential to allow the RL agent to effectively learn and adapt its control policies to achieve optimal performance under varying environmental conditions. Furthermore, careful selection of **control actions** is equally important, as it determines how the RL agent interacts with the environment. By choosing suitable control actions, such as adjusting the temperature set-point, the RL agent can effectively optimize energy usage. Finally, algorithms are developed to optimize the **reward functions** associated with states and actions, ensuring that the RL agent learns to make decisions that lead to the most favorable outcomes. This approach allows the development of intelligent control systems capable of dynamically responding to changing environmental conditions and ambient conditions needs.

The Q learning algorithm has **five actions**(Table 4):

The Table 6 describes examples of potential **reward functions** applicable in RL algorithms.

Table 6. Different Reward Functions used.

Strategy	Reward Function Proposal
Reward function penalizes the agent for using the heating system	$R1 = \begin{cases} -1 & \text{if } Q_{\text{heating}} > 0 \\ 0 & \text{if } Q_{\text{heating}} = 0 \end{cases}$
Reward function penalizes the use of the heating system taking into account the efficiency and the actual deviation from the desired temperature range	$R2 = -(\alpha \cdot E + \beta \cdot D1 + \gamma \cdot L1)$
Refined Reward Function: Energy consumption penalty. Temperature stability. Exceeding maximum and minimum temperatures. Penalty for frequent changes in the set point	$R3 = -(\alpha \cdot E + \beta \cdot D2 + \gamma \cdot L2 + \delta \cdot S)$

An appropriate RL algorithm, such as Q learning, deep Q network (DQN), or posterior policy optimization (PPO), should be chosen, based on the complexity of the task and the available computational resources, where:

- α , β , and γ are weight coefficients that can be tuned.
- E is the energy consumption penalty.
- T_{in} is the current inside temperature.
- T_{setpoint} is the setpoint temperature.
- $D1, D2$ is the penalty for temperature deviation.

$$D1 = |T_{\text{in}} - T_{\text{setpoint}}|$$

$$D2 = |T_{\text{in}} - T_{\text{setpoint}}| + (T_{\text{in}} - T_{\text{setpoint}})^2$$

- $L1, L2$ are the limit penalty for exceeding the maximum or minimum temperature. The limit penalty is applied when the inside temperature exceeds the maximum or minimum temperature limits:

$$L1 = \begin{cases} |T_{\text{in}} - T_{\text{max}}| & \text{if } T_{\text{in}} > T_{\text{max}} \\ |T_{\text{in}} - T_{\text{min}}| & \text{if } T_{\text{in}} < T_{\text{min}} \\ 0 & \text{if } T_{\text{min}} \leq T_{\text{in}} \leq T_{\text{max}} \end{cases}$$

$$L2 = \begin{cases} (T_{\text{in}} - T_{\text{max}})^2 & \text{if } T_{\text{in}} > T_{\text{max}} \\ (T_{\text{min}} - T_{\text{in}})^2 & \text{if } T_{\text{in}} < T_{\text{min}} \\ 0 & \text{otherwise} \end{cases}$$

where T_{max} and T_{min} are the maximum and minimum allowable temperatures, respectively.

- S is the set point change penalty.
- The set-point change penalty discourages frequent and large adjustments to the set-point temperature:

$$S = |T_{\text{set-point}} - T_{\text{previous_set-point}}|$$

where $T_{\text{previous_set-point}}$ is the set-point temperature from the previous time step.

The diagram illustrates a Reinforcement Learning (RL) control loop for a greenhouse temperature system. The main components and their interactions are as follows:

- Control Loop:** A central block containing a heater/cooler unit labeled "on/off". It receives a setpoint $T_{setpoint}$ and a feedback signal ϵ from the RL Agent. It outputs heat Q_t and energy E_t to the greenhouse.
- Greenhouse:** A block representing the greenhouse environment. It receives Q_t and E_t from the Control Loop. It contains a thermometer indicating the internal temperature T_{in} and a power meter showing energy consumption Kwh_t . It also receives external temperature T_{out} .
- RL Agent:** Represented by a brain icon, it receives the internal temperature T_{in} from the greenhouse and the error signal ϵ from the Control Loop. It outputs the $T_{setpoint}$ to the Control Loop and an **Action** $f(T_{setpoint})$ to the RL Control analysis block.
- RL (Q-Learning) Reward function:** A block that receives the internal temperature T_{in} and provides a reward signal to the RL Agent.
- RL Control analysis:** A block that receives the **Action** $f(T_{setpoint})$ and provides feedback to the RL Agent.
- RL feedback:** A block that receives the internal temperature T_{in} and provides a feedback signal to the RL Agent.

The governing equation for the temperature change in the greenhouse is:

$$\frac{dT_{in}}{dt} = \frac{Q_{solar} + Q_{heat} - U \cdot A \cdot (T_{in} - T_{out})}{C_{greenhouse}}$$

1a) Cooling temperature control analysis using the differential equation model (Figure 7)

In R1a-R1b the RL system manages to maintain the inside temperature within the desired range more effectively than the fixed setpoint system. Temperature variations with RL control are less pronounced and stay closer to the set-point compared to the fixed set-point system. In R2a-R2b Similar to R1a-R1b, the RL control maintains a more stable temperature, with fewer and smaller deviations from the set point. The fixed set point system shows greater fluctuations and occasionally falls outside the desired temperature range. In R3a-R3b the graph reinforces the previous observations, where the RL system exhibits better control over the temperature, keeping it within the desired limits more consistently than the fixed set-point system. The fixed set point control shows larger temperature changes and less precision in maintaining the set point. R1a and R3a appear to perform the best, with

higher and more consistent energy savings at different sampling times. The Top Left Subplot shows a slight edge in overall savings stability.

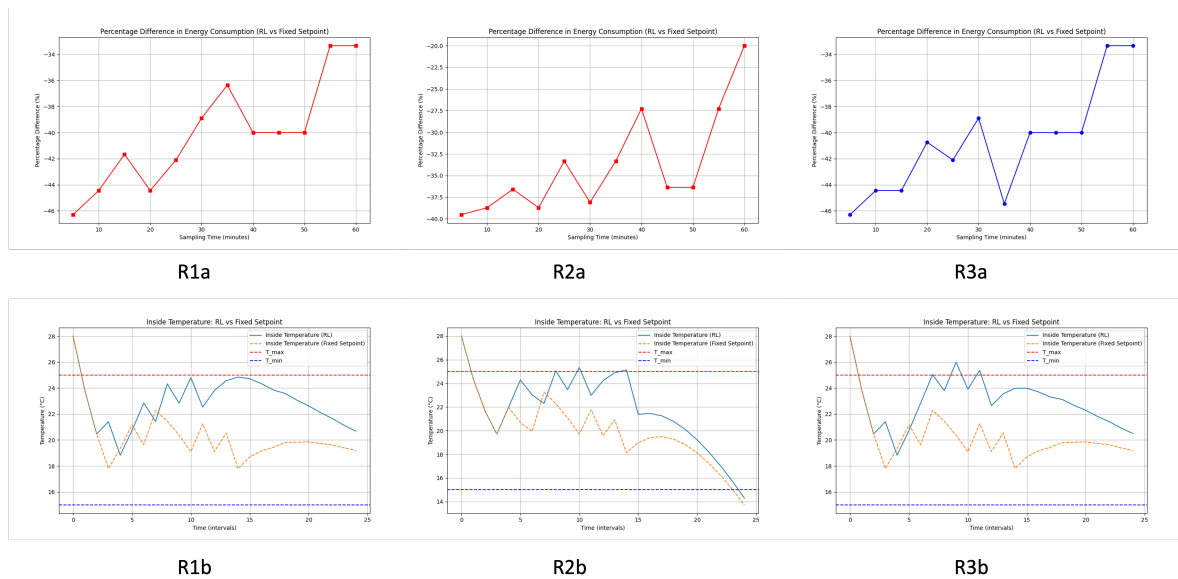


Figure 7. Comparative analysis in greenhouse cooling process

General Improvement with RL Control

Across all energy consumption graphs, the RL system consistently outperforms the fixed set point system in terms of energy savings. The improvement is particularly notable for shorter sampling times (5 to 20 minutes), where the savings are more substantial. Even at longer sampling intervals (up to 60 minutes), the RL system maintains a significant percentage of energy savings.

Temperature Control Effectiveness

The lower subplots demonstrate that the RL control system maintains the desired temperature range more effectively than the fixed set-point system. The RL system results in smaller deviations from the set-point, indicating better control and stability. The fixed set-point system shows larger temperature fluctuations and occasionally fails to keep the temperature within the desired range.

Conclusion in cooling process using the differential equation model

The analysis indicates that the RL control system is highly effective in reducing energy consumption compared to the fixed set-point system. Shorter sampling times yield the best results, with the RL system showing a 45% improvement in energy efficiency. Furthermore, the RL system demonstrates superior temperature control, maintaining the desired range more consistently and with fewer deviations. In general, the RL system is a preferable choice for optimizing energy use and temperature stability in greenhouse temperature control scenarios.

In the R1a-R1b subgraph shows a consistent improvement in energy consumption with the RL system as the sampling time increases. For sampling times between 5 and 10 minutes, the RL system shows significant energy savings of around -45%. As the sampling time increases, the savings decrease slightly but remain significant, ending around -35% at 60 minutes. In R2a-R2b exhibits a similar trend, with the RL system performing better than the fixed set-point system. The initial savings are around -40%, with fluctuations as the sampling time increases. The savings dip slightly around 30 to 40 minutes but show an overall improvement, ending around -20% for 60 minutes. In R3a-R3b The trend here is also consistent with the other graphs, showing that the RL system consistently saves energy compared to the fixed set-point system. The savings are initially around -45% for shorter sampling

times. Despite fluctuations, the savings remain significant throughout, ending around -35% for 60 minutes.

1b) Heating temperature control analysis using the differential equation model (Figure 8)

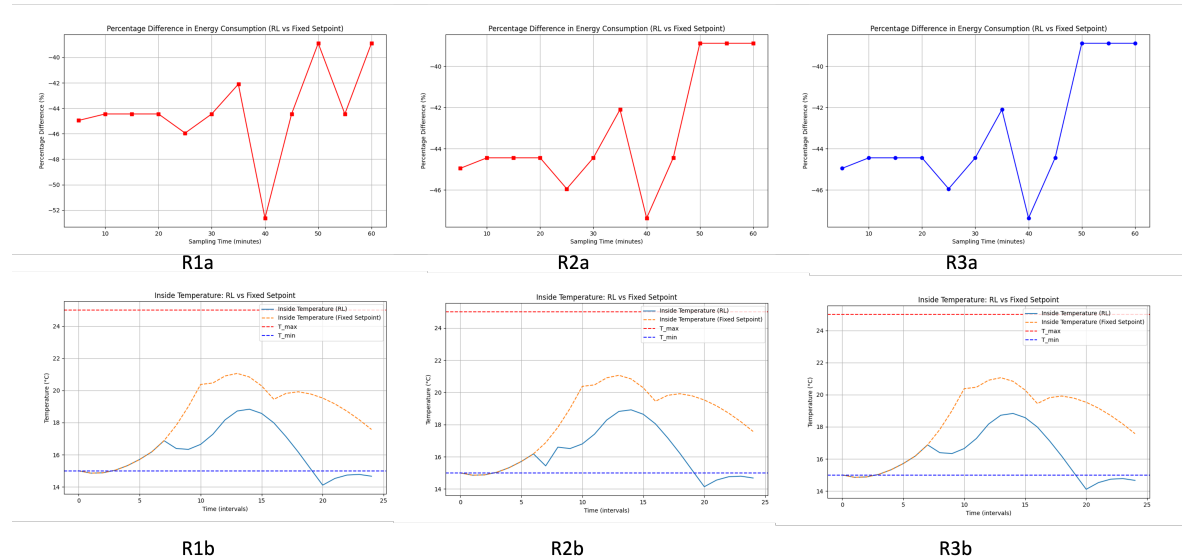


Figure 8. Comparative analysis in greenhouse heating process

General Improvement with RL Control

Across all energy consumption graphs, the RL system consistently outperforms the fixed set point system in terms of energy savings. The improvement is particularly notable for shorter sampling times (5 to 20 minutes), where the savings are more substantial. Even at longer sampling intervals (up to 60 minutes), the RL system maintains a significant percentage of energy savings.

Temperature Control Effectiveness

The bottom subplots demonstrate that the RL control system maintains the desired temperature range more effectively than the fixed set point system. The RL system results in smaller deviations from the set point, indicating better control and stability. The fixed set point system shows larger temperature fluctuations and occasionally fails to keep the temperature within the desired range.

Conclusion in heating process using the differential equation model

The analysis indicates that the RL control system is highly effective in reducing energy consumption compared to the fixed set-point system. Shorter sampling times yield the best results, with the RL system showing a 45% improvement in energy efficiency. Furthermore, the RL system demonstrates superior temperature control, maintaining the desired range more consistently and with fewer deviations. In general, the RL system is a preferable choice for optimizing energy use and temperature stability in greenhouse temperature control scenarios.

8.2. RL greenhouse control based in temperature prediction model

The first goal is to design and implement a model that predicts indoor temperature (TI) based on various environmental variables. The model takes advantage of past data and predicted future values to make accurate forecasts. Once there is an estimated value of the temperature evolution inside the greenhouse, the RL algorithm is applied acting on the set-point, depending on the said evolution. From the previous study using the differential equations model, it was concluded that the reward function R1 is viable to display the algorithm. For all these reasons, in this analysis with data obtained in the greenhouse we are going to use this reward function.

Prediction model

The model predicts the indoor temperature (*TI*) using the following variables:

- External Temperature (*TE*)
- External Relative Humidity (*HRE*)
- Wind Direction (*DV*)
- Wind Speed (*VV*)
- External Global Radiation (*RGE*)
- Internal Relative Humidity (*HRI*)

The prediction is based on the past values (lags) of these variables for the previous 60 minutes and the predicted external temperatures for the next 60 minutes.

We created lagged characteristics for the last 60 minutes and leading characteristics for the future external temperature (*TE*). This allows the model to capture temporal dependencies. The data set is divided into training and testing sets. The features are standardized to have a mean of 0 and a standard deviation of 1. We use a Linear Regression model to predict the indoor temperature.

For predicting the interior temperature of the greenhouse, Random Forest and Linear Regression models are the most reliable based on their lower RMSE and MAE values. The Neural Network model’s high error rates suggest that it may not be suitable for this particular prediction task without further tuning or perhaps a different architecture or feature set. Linear regression is chosen for its simplicity and interpretability. The model’s performance is evaluated using the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) on both the training and testing datasets.

The Figure 9 compares the performance of different machine learning models based on two metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Data were captured from November 2023 to May 2024.

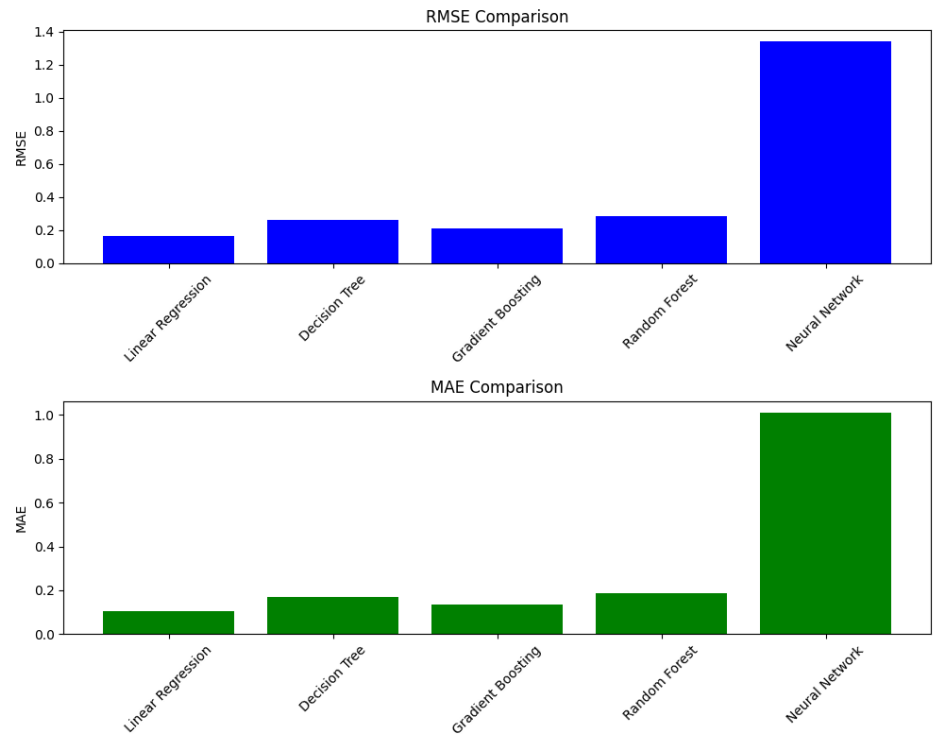


Figure 9. Machine learning methods to predict the temperature inside the greenhouse (*TI*)

2a) Cooling control analysis using the temperature prediction model (Figure 10)

The RL agent was trained to control the greenhouse temperature. The performance of the RL-based control was compared to that of a fixed set point control. The total energy consumption and the inside temperature regulation were evaluated.

The actions selected depend on the predicted temperature. In this way, it does not depend on a value set in the set point, but on variable values that depend on the behavior of the greenhouse.

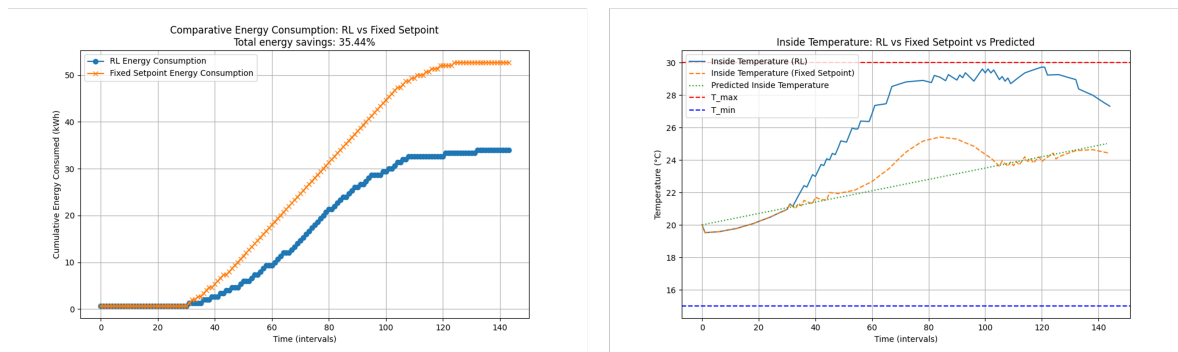


Figure 10. Energy Consumption: RL vs Fixed Setpoint (left) and Inside Temperature: RL vs Fixed Setpoint vs Predicted (right)

Figure 10 shows the regulation of inside temperature using RL control, fixed setpoint control, and predicted inside temperature. The RL control maintains the temperature closer to the upper limit (T_{max}) compared to the fixed set-point control. This dynamic adjustment helps minimize energy consumption while ensuring that the temperature stays within the desired range. The predicted temperature provides a reference for the RL control actions. The cumulative energy consumption graph (the graph on the right) demonstrates that the RL-based control performs better than the fixed set point control, achieving a total energy savings of 35.44%. This significant reduction in energy usage highlights the efficiency of the RL algorithm in optimizing the control strategy. The graph on the right reveals that the RL control effectively maintains the temperature within the desired range (between T_{min} and T_{max}). The RL control shows a more dynamic response compared to the fixed set-point control, which can lead to more efficient heating. The predicted temperature aligns well with the actual inside temperature regulated by the RL control, indicating accurate predictions and effective control actions. The RL-based temperature control (in the cooling phase) method shows significant advantages over fixed set point control in terms of energy savings and effective temperature regulation. The ability of RL to adaptively adjust the set point temperature based on predicted inside temperatures leads to optimized energy consumption and better maintenance of the desired temperature range.

2b) Heating control using the temperature prediction model (Figure 11)

Figure 11 compares the cumulative energy consumption between the RL-based control and the fixed set-point control. The RL control demonstrates significant energy savings, with a total reduction of 25.93% compared to the fixed set-point control. This indicates the effectiveness of the RL algorithm in optimizing energy usage while maintaining the desired temperature.

The second figure in the composite image shows the regulation of the inside temperature using RL control, fixed set-point control, and the predicted inside temperature. The RL control maintains the temperature closer to the upper limit (T_{max}) compared to the fixed setpoint control. This dynamic adjustment helps minimize energy consumption while ensuring that the temperature stays within the desired range. The predicted temperature provides a reference for the RL control actions.

The cumulative energy consumption graph demonstrates that the RL-based control performs better than the fixed set point control, achieving a total energy savings of 25.93%. This significant reduction in energy usage highlights the efficiency of the RL algorithm in optimizing the control strategy.

The inside temperature regulation graph reveals that the RL control effectively maintains the temperature within the desired range (between T_{min} and T_{max}). The RL control shows a more dynamic response compared to the fixed set point control, which can lead to more efficient heating. The predicted temperature aligns well with the actual inside temperature regulated by the RL control, indicating accurate predictions and effective control actions.

The RL-based temperature control method shows significant advantages over the fixed set point control in terms of energy savings and effective temperature regulation. The ability of RL to adaptively adjust the setpoint temperature based on predicted inside temperatures leads to optimized energy consumption and better maintenance of the desired temperature range.

Table 7. Actions taken by RL algorithm in heating case using temperature (TI) prediction. The control values are limited by T_{max} and T_{min}

ine Action 1	$T_{setpoint} = T_{in_predicted}$
ine Action 2	$T_{setpoint} = T_{in_predicted} - 1$
ine Action 3	$T_{setpoint} = T_{in_predicted} - 2$
ine Action 4	$T_{setpoint} = T_{in_predicted} - 3$
ine Action 5	$T_{setpoint} = T_{in_predicted} - 4$
ine	

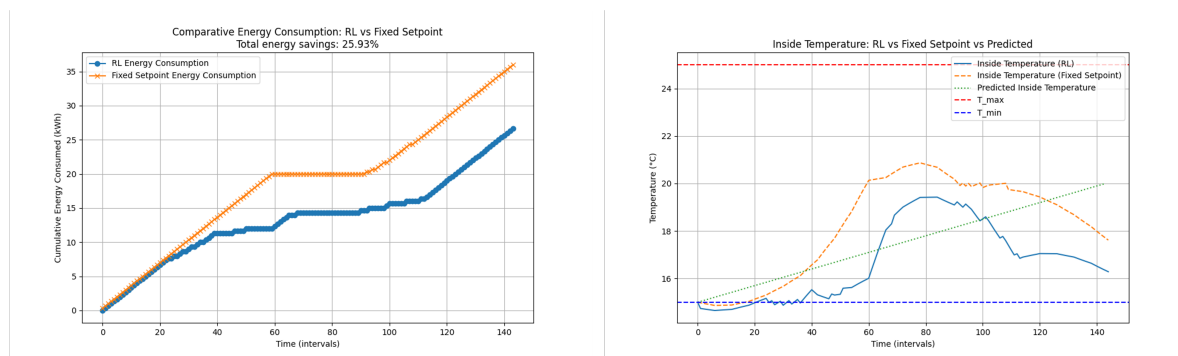


Figure 11. Comparative Energy Consumption: RL vs Fixed Setpoint and Inside Temperature: RL vs Fixed Setpoint vs Predicted

This Q-learning algorithm optimizes (in the heating phase) the greenhouse set-point temperature by selecting from five actions to adjust the set-point. The reward function penalizes heating usage, encouraging the algorithm to find a set-point that minimizes energy consumption while maintaining the desired temperature range.

Analysis of Energy Consumption and Savings for different $T_{set-point}$

The graph on the left of Figure 12 illustrates the cumulative energy consumption over time for different set point temperatures ($T_{setpoint}$). The graph compares the energy consumption between Reinforcement Learning (RL) based control and a fixed set-point control for various setpoints.

Key Observations:

- **Energy Consumption Patterns:**
 - The energy consumption increases over time for both RL and fixed set point controls.
 - For lower $T_{setpoint}$ values, the energy consumption is generally lower. As $T_{setpoint}$ increases, energy consumption increases.
- **Comparison between RL and Fixed Set point:**
 - RL control consistently consumes less energy compared to fixed set point control for all set points.

- The difference in energy consumption between the RL and the fixed set point control is more pronounced at higher set points.

Right Graph: Energy Savings for Different $T_{setpoint}$

The graph on the right of Figure 12 shows the percentage of energy savings achieved by using the RL control compared to the fixed set point control in different $T_{setpoint}$ values.

Key Observations:

- **Energy Savings Trend:**
 - The energy savings increase with higher $T_{setpoint}$ values.
 - The energy savings range from approximately 36% at $T_{setpoint} = 18.0^{\circ}\text{C}$ to nearly 50% at $T_{setpoint} = 22.0^{\circ}\text{C}$.
- **Efficiency of RL Control:**
 - The RL control becomes more efficient in terms of energy savings as the $T_{setpoint}$ increases.
 - This indicates that RL control is particularly beneficial in scenarios where higher set points are required, resulting in significant energy savings.

The analysis of the graphs reveals that the RL-based control offers substantial energy savings compared to the fixed set-point control across various set points. The energy savings are more pronounced at higher set points, indicating the effectiveness of RL control in optimizing energy consumption while maintaining desired temperature conditions in the greenhouse. This makes RL control a promising approach for energy-efficient temperature regulation.

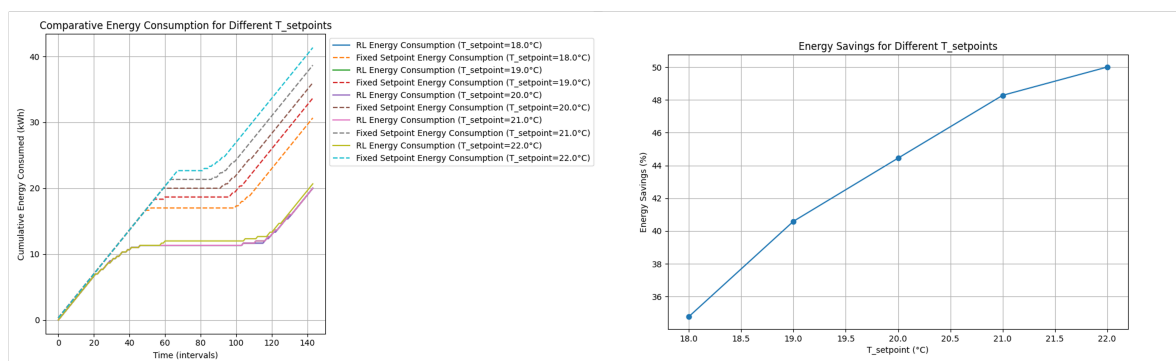


Figure 12. Comparative analysis taking different reference temperatures (setpoints)

9. Conclusion

The integration of Internet of Things (IoT) protocols and Reinforcement Learning (RL) methodologies has been shown to be effective in managing and optimizing greenhouse operations for industrial hemp cultivation. This combination not only enhances operational efficiency, but also maintains selected temperatures and optimizes energy consumption more effectively than classical control methods. By reducing the need for constant human intervention, this technological integration minimizes labor costs and increases scalability for larger agricultural enterprises.

The RL based control system shows significant energy savings while maintaining the desired temperature ranges, outperforming traditional fixed set-point control systems. Specifically, the study shows energy savings of up to 45% during cooling processes and 25.93% during heating processes. In addition, this new control approach simplifies the workload of technicians by eliminating the need for complex analyzes to achieve the same results, allowing them to focus on higher-level oversight and maintenance tasks.

Integration with IoT plays a crucial role in this setup, enabling real-time data acquisition and seamless communication between various greenhouse subsystems. IoT devices collect and transmit

environmental data that RL algorithms use to make informed dynamic adjustments to greenhouse conditions. This IoT integration ensures that the RL model can adapt to changing conditions promptly and accurately, thus optimizing resource use and improving overall system responsiveness.

RL algorithms are capable of adaptively adjusting set-point temperatures based on real-time data and predictions, leading to optimized energy consumption and better maintenance of desired environmental conditions. The study validates the practical implementation of RL models in automated greenhouses in the real world, showcasing their ability to scale and adapt to different types of crops and environmental conditions. These conclusions highlight the potential of combining IoT and RL technologies to improve the efficiency, scalability, and sustainability of greenhouse operations, particularly for industrial hemp cultivation.

Author Contributions: Conceptualisation, J.M.C.-Z., F.J.F.-P.; methodology, S.P.-P., J.M.C.-Z., F.J.F.-P., J.A.B.-C and M.P.-H; software, S.P.-P., J.A.B.-C and M.P.-H; validation, S.P.-P., J.M.C.-Z., F.J.F.-P. and J.A.B.-C.; formal analysis, S.P.-P., J.M.C.-Z., F.J.F.-P., M.P.-H and J.A.B.-C.; investigation, S.P.-P., J.M.C.-Z. and M.P.-H; resources, J.M.C.-Z. and F.J.F.-P.; data analysis, S.P.-P. and J.M.C.-Z.; writing—original draft preparation, S.P.-P. and M.P.-H; writing—review and editing, J.M.C.-Z., J.F.-P. and J.A.B.-C; visualization, J.M.C.-Z. and F.J.F.-P.; supervision, J.M.C.-Z. and F.J.F.-P.; project administration, J.M.C.-Z. and S.P.-P; funding acquisition, J.M.C.-Z. and F.J.F.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This study is part of the AGROALNEXT program (AGROALNEXT/2022/048) and was supported by MCIN with funding from the European Union NextGenerationEU (PRTR-C17.I1) and the Generalitat Valenciana, and the Valencian Innovation Agency (AVI) under scientific innovation unit (UCIE Ars Innovatio) of the University of Alicante

Institutional Review Board Statement: In this section, you should add the Institutional Review Board Statement and approval number, if relevant to your study. You might choose to exclude this statement if the study did not require ethical approval. Please note that the Editorial Office might ask you for further information. Please add “The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of NAME OF INSTITUTE (protocol code XXX and date of approval).” for studies involving humans. OR “The animal study protocol was approved by the Institutional Review Board (or Ethics Committee) of NAME OF INSTITUTE (protocol code XXX and date of approval).” for studies involving animals. OR “Ethical review and approval were waived for this study due to REASON (please provide a detailed justification).” OR “Not applicable” for studies not involving humans or animals.

Data Availability Statement: We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section “MDPI Research Data Policies” at <https://www.mdpi.com/ethics>.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: Declare conflicts of interest or state “The authors declare no conflicts of interest.” Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results must be declared in this section. If there is no role, please state “The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results”.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three letter acronym
LD	linear dichroism

References

1. Kiumarsi, B.; Vamvoudakis, K.; Modares, H.; Lewis, F. Optimal and Autonomous Control Using Reinforcement Learning: A Survey. *IEEE Transactions on Neural Networks and Learning Systems* **2018**. <https://doi.org/10.1109/TNNLS.2017.2773458>.
2. Perera, A.T.D.; Kamalaruban, P. Applications of reinforcement learning in energy systems. *Renewable & Sustainable Energy Reviews* **2021**. <https://doi.org/10.1016/J.RSER.2020.110618>.
3. Wang, Y.; Velswamy, K.; Huang, B. A Long-Short Term Memory Recurrent Neural Network Based Reinforcement Learning Controller for Office Heating Ventilation and Air Conditioning Systems **2017**. <https://doi.org/10.3390/PR5030046>.
4. Kazmi, H.; Mehmood, F.; Lodeweyckx, S.; Driesen, J. Deep Reinforcement Learning based Optimal Control of Hot Water Systems. *Energy* **2018**. <https://doi.org/10.1016/J.ENERGY.2017.12.019>.
5. Ruelens, F.; Claessens, B.; Quaiyum, S.; Schutter, B.D.; Babuska, R.; Belmans, R. Reinforcement Learning Applied to an Electric Water Heater: From Theory to Practice. *IEEE Transactions on Smart Grid* **2018**. <https://doi.org/10.1109/TSG.2016.2640184>.
6. Zhang, Z.; Chong, A.; Pan, Y.; Zhang, C.; Lam, K.P. Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. *Energy and Buildings* **2019**. <https://doi.org/10.1016/J.ENBUILD.2019.07.029>.
7. Mason, K.; Grijalva, S. A review of reinforcement learning for autonomous building energy management. *Computers & Electrical Engineering* **2019**. <https://doi.org/10.1016/J.COMPELECENG.2019.07.019>.
8. Hosseinloo, A.H.; Ryzhov, A.; Bisch, A.; Ouerdane, H.; Turitsyn, K.; Dahleh, M.A. Data-driven control of micro-climate in buildings: An event-triggered reinforcement learning approach. *Applied Energy* **2020**. <https://doi.org/10.1016/J.APENERGY.2020.115451>.
9. Sun, Q.; Wang, X.; Liu, Z.; Mirsaedi, S.; He, J.; Pei, W. Multi-agent energy management optimization for integrated energy systems under the energy and carbon co-trading market **2022**. <https://doi.org/10.1016/J.APENERGY.2022.119646>.
10. al Ani, O.; Das, S. Reinforcement Learning: Theory and Applications in HEMS. *Energies* **2022**. <https://doi.org/10.3390/EN15176392>.
11. Fu, Q.; Han, Z.; Chen, J.; Lu, Y.; Wu, H.; Wang, Y.; Fu, Q.; Han, Z.; Chen, J.; Lu, Y.; Wu, H.; Wang, Y. Applications of reinforcement learning for building energy efficiency control: A review **2022**. <https://doi.org/10.1016/J.JOBE.2022.104165>.
12. Mauree, D.; Naboni, E.; Coccolo, S.; Perera, A.T.D.; Nik, V.M.; Scartezzini, J.L. A review of assessment methods for the urban environment and its energy sustainability to guarantee climate adaptation of future cities. *Renewable & Sustainable Energy Reviews* **2019**. <https://doi.org/10.1016/J.RSER.2019.06.005>.
13. Kazmi, H.; Suykens, J.A.K.; Balint, A.; Driesen, J. Multi-agent reinforcement learning for modeling and control of thermostatically controlled loads. *Applied Energy* **2019**. <https://doi.org/10.1016/J.APENERGY.2019.01.140>.
14. Chen, L.; Xu, L.; Wei, R.; Chen, L.; Xu, L.; Wei, R. Energy-Saving Control Algorithm of Venlo Greenhouse Skylight and Wet Curtain Fan Based on Reinforcement Learning with Soft Action Mask **2023**. <https://doi.org/10.3390/AGRICULTURE13010141>.
15. Ruelens, F.; Iacovella, S.; Claessens, B.; Belmans, R. Learning Agent for a Heat-Pump Thermostat with a Set-Back Strategy Using Model-Free Reinforcement Learning. *Energies* **2015**. <https://doi.org/10.3390/EN8088300>.
16. Vázquez-Canteli, J.R.; Ulyanin, S.; Kämpf, J.H.; Nagy, Z. Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities. *Sustainable Cities and Society* **2019**. doi:10.1016/J.SCS.2018.11.021.
17. Sutton, R.; Barto, A. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks* **2005**. <https://doi.org/10.1109/TNN.1998.712192>.
18. Peters, M.; Ketter, W.; Saar-Tsechansky, M.; Collins, J. A reinforcement learning approach to autonomous decision-making in smart electricity markets. *Machine Learning* **2013**. <https://doi.org/10.1007/S10994-013-5340-0>.
19. Kar, S.; Moura, J.M.F.; Poor, H.V. QD -Learning: A Collaborative Distributed Strategy for Multi-Agent Reinforcement Learning Through Consensus + Innovations. *IEEE Transactions on Signal Processing* **2013**. <https://doi.org/10.1109/TSP.2013.2241057>.

20. Sierla, S.; Ihasalo, H.; Vyatkin, V. A Review of Reinforcement Learning Applications to Control of Heating, Ventilation and Air Conditioning Systems. *Energies* **2022**. <https://doi.org/10.3390/EN15103526>.
21. Vázquez-Canteli, J.R.; Nagy, Z. Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy* **2019**. <https://doi.org/10.1016/J.APENERGY.2018.11.002>.
22. Ferrández-Pastor, F.J.; García-Chamizo, J.M.; Nieto-Hidalgo, M.; Mora-Pascual, J.; Mora-Martínez, J. Developing Ubiquitous Sensor Network Platform Using Internet of Things: Application in Precision Agriculture. *Sensors* **2016**, *16*. <https://doi.org/10.3390/s16071141>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.