CHEMICAL ENGINEERING TRANSACTIONS
VOL. 103, 2023

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Petar S. Varbanov, Panos Seferlis, Yee Van Fan, Athanasios Papadopoulos

ISBN 979-12-81206-02-1; **ISSN** 2283-9216

DOI: 10.3303/CET23103014

Improving Resource Use Efficiency in Plant Factories Using Deep Reinforcement Learning for Sustainable Food Production

Benjamin Decardi-Nelson*, Fengqi You

Cornell University, Ithaca, New York, USA
bd368@cornell.edu

Plant factories with artificial lighting (PFALs) are heralded as a potential solution to enhance the resilience of the food production system by amplifying productivity per unit area of land. However, PFALs typically demand higher resource consumption than traditional greenhouses or open-field farming. To optimize resource use in PFALs, we developed a nonlinear, automatic control system utilizing deep reinforcement learning (DRL). The proposed system implicitly learns the intricate dynamics of crop behavior and environmental variables, aiming to curtail resource, particularly energy consumption in PFALs, while ensuring that these environmental variables stay within desired operating parameters. We demonstrate the efficacy of the proposed DRL-based automatic control system through a case study: a shipping container PFAL situated in Ithaca, New York. We evaluate the ability of the proposed DRL system to reduce energy consumption by comparing its energy consumption to that of the conventional control method. Our findings indicate that, in typical Ithaca summer and winter conditions, the DRL-based control system can potentially decrease energy consumption by about 31 % and 23 % compared to the conventional control method.

1. Introduction

The rapidly growing human population, dwindling resources, and shifts in climate patterns are increasingly straining the food production system (Vogel et al., 2019). To strengthen the resilience of the current food system, innovative methods that produce more food with fewer resources are needed (Pennisi et al., 2019). One promising approach is Controlled Environment Agriculture (CEA), which includes plant factories and greenhouses, known to enhance sustainable resilience in food production (Benke and Tomkins, 2017). Plant factories with artificial lighting (PFALs) are noteworthy as they could dissociate stable crop production from geographical climate constraints while maintaining high crop quality and efficiency (Beacham et al., 2019). Unlike traditional greenhouses, PFALs can be entirely isolated from the external environment, necessitating precise control over lighting and ventilation for optimal crop growth and resource conservation (Kozai, 2018). Operational costs, especially those related to energy use, present a significant barrier to PFAL adoption. For instance, artificial lighting in PFALs comprises about 80 % of total electricity consumption (Kozai et al., 2019). The development of PFAL automation systems that ensure both the PFAL's optimal operation and efficient resource use is of paramount importance.

Although extensive studies on optimal control techniques for greenhouses exist (Hu and You, 2022), their application to PFALs remains underexplored. Early efforts to integrate optimal control techniques into PFAL operations can be traced back to Morimoto et al. (1995). Their study utilized a data-driven autoregressive moving average (ARMA) model of plant physiological processes within a PFAL, leading to the development of a linear quadratic regulator (LQR) for its optimal operation. However, they overlooked uncertainties arising from the linear ARMA model and the effects of external disturbances. Deng et al. (2018) later addressed disturbance issues by developing a robust linear optimal control method for PFALs, but they only considered disturbances from bacteria and aphids and omitted modeling uncertainties. Due to the limited applicability of linear control methods, Xu et al. (2021) developed a nonlinear optimal control method to investigate energy consumption reduction and lettuce yield in a lettuce PFAL. The application of these optimal control techniques largely depends

Paper Received: 15 March 2023; Revised: 10 April 2023; Accepted: 25 May 2023
Please cite this article as: Decardi-Nelson B., You F., 2023, Improving Resource Use Efficiency in Plant Factories Using Deep Reinforcement Learning for Sustainable Food Production, Chemical Engineering Transactions, 103, 79-84 DOI:10.3303/CET23103014

on the availability of mathematical models (Hu et al., 2023). However, due to the challenges of developing accurate mathematical models of biological systems, alternatives such as deep reinforcement learning (DRL) have been proposed (Ajagekar et al., 2023a). DRL methods have shown significant performance in building climate control (Xie et al., 2023). Given the limitations of conventional control methods and the advancements in DRL techniques, there is a pressing need to investigate automated DRL-based control methods for PFALs (van Delden et al., 2021). This led Zheng et al. (2021) to develop a deep deterministic policy gradient (DDPG)-based DRL for PFAL automatic control. Their focus, however, remained limited to day-to-day agricultural management activities, overlooking the climate's hour-by-hour impact on crops. They did not account for energy consumption in the reward function and the presence of uncertainties in the control system. Consequently, there is an immediate need to develop better-automated control methods based on DRL, which not only regulate the environmental factors of PFALs but also curtail resource consumption.

In this work, we present a nonlinear automatic control framework based on DRL for regulating the environmental variables and streamlining resource use in a PFAL. The DRL-based automatic control framework uses only the plant weight information and environmental factors like indoor air temperature, humidity, and CO_2 concentration to streamline the PFAL operations. We use lettuce cultivation in a shipping container located in Ithaca, New York, as a case study to demonstrate how the DRL-based control framework can be employed to efficiently manage the growth environment in PFALs, optimizing resource use.

2. Deep reinforcement learning framework

In DRL, the agent (in this context controller) refines its decision-making capabilities through continuous interaction with an environment, utilizing a balance of exploration and exploitation (Sutton and Barto, 2018). The environment in this context is characterized as a Markov Decision Process, defined by the following parameters:

- a state space S
- an action space A
- a transition function $T: S \times A \rightarrow S$
- a reward function $R: S \times A \rightarrow R$
- a discount factor γ which takes a value between 0 and 1.

In this study, the state and action spaces of the PFAL are depicted in Tables 1 and 2. It is important to note that the only crop-specific information given to the DRL agent pertains to the crop's dry weight. The feasibility of this approach is largely due to recent advancements in estimating crop fresh weight using artificial intelligence techniques (Reyes-Yanes, 2020). We also introduce a photoperiod indicator variable, which assumes a value of 0 during the dark period and 1 during the light period, signaling whether the artificial light is off or on. Considering the varying temperature needs of the crops during light and dark periods, providing data on the optimal indoor temperature range helps the DRL controller adapt to these changes. Likewise, the DRL controller is given outdoor temperature and humidity values, enabling it to identify when the outdoor conditions are suitable for conserving energy (Ajagekar et al., 2023b). This practice of exchanging indoor air with outdoor air to regulate the PFAL environmental variables to save energy, known as a heating, ventilation, and air conditioning (HVAC) economizer, is widely recognized (Eaton et al., 2023).

Table 1: State-space for the DRL framework

Variable	Description	Units
w	Crop dry weight	kg/m^2
c	Indoor CO_2 concentration	kg/m^3
T	Indoor temperature	°C
h	Indoor humidity	kg/m^3
T_{lb}	Desired indoor lower bound temperature	°C
T_{ub}	Desired indoor upper bound temperature	°C
i_{photo}	Photoperiod indicator	-
T_o	Outdoor temperature	°C
h_o	Outdoor humidity	kg/m^3
t_{day}	Time of day	s

The control actions presented in Table 2 symbolize the operational actuators within the PFAL, which are adjusted to manage environmental variables such as lighting, temperature, humidity, and CO_2 levels within the PFAL. For instance, a dehumidifier is utilized to extract excess humidity generated by the crops through transpiration (Chen et al., 2021). Similarly, the artificial lighting system provides light energy, which is necessary for plant photosynthesis.

Table 2: Action space for the DRL framework

Variable	Description	Units
U_L	Power to the artificial lighting system	W/m^2
U_c	Supplemental CO_2 supply rate	kg/m^2s
U_q	Cooling or heating rate	W/m^2
U_h	Dehumidification rate	kg/m^2s
U_v	Ventilation rate	m^3/m^2s

The PFAL state variables are described by a system of ordinary differential equations (ODEs). The variables are the crop dry weight, indoor CO_2 concentration, indoor air temperature, and indoor relative humidity. The crop dry weight is described by the two-state crop model developed by Van Henten (1994) and is given in Eqs. (1) and (2).

$$\frac{dw_{ns}}{dt} = c_\alpha \phi_{photo} - r_{gr} w_s - \phi_{resp} - \frac{1-c_\beta}{c_\beta} r_{gr} w_s \tag{1}$$

$$\frac{dw_s}{dt} = r_{gr} w_s \tag{2}$$

where w_{ns} (kg/m^2) and w_s (kg/m^2) denote the non-structural and structural dry weights, ϕ_{photo} (kg/m^2s) denotes the canopy rate of photosynthesis, r_{gr} (1/s) is the specific growth rate, ϕ_{resp} (kg/m_2s) is the maintenance respiration rate, t (s) is the time, and c_α and c_β are crop growth parameters (Chen et al., 2022a). The total crop dry weight w (kg/m^2) is obtained from the structural and non-structural dry weight using the following relation: $w = w_s + w_{ns}$. The state space of the DRL framework is an extension of the three environmental variables and the crop's total dry weight. Additional information (as detailed in Table 1), which is crucial for enhancing computational efficiency during the training phase, is incorporated into the state space. Notably, the DRL controller operates without information on crop photosynthesis, transpiration, and respiration rates, as these are challenging to measure accurately. The anticipation is that the DRL controller will implicitly acquire this knowledge to make optimal sequential decisions. During the DRL training phase, the agent receives the observed state (s_t) and reward (r_t) at each time-step (t) from the environment and, in turn, chooses an action (a_t), which is applied to the environment. The goal is to obtain an optimal action for the current state that maximizes the cumulative reward. The reward function employed in this study is outlined in Eq.(3).

$$r = -\sum_{i=0}^{n_u} p_i + \frac{\Delta w}{\Delta t} \tag{3}$$

where p_i denotes the price of the i-th control activity within the PFAL, n_u is the number of inputs, and $\Delta w/\Delta t$ is the crop growth rate. The costs for electricity and CO_2 are 0.2051 US $/kWh and 0.2 US $/kg (Chen et al., 2022b). In this work, we used the soft actor-critic (SAC) DRL method (Haarnoja et al., 2018) to obtain the optimal policy.

3. Case study

In this study, we present a case study to demonstrate the effectiveness of the DRL control framework to improve the energy use efficiency of the PFAL compared to the conventional control method. We begin this section by first describing the simulation settings. Thereafter, we present the results of the trajectories of the PFAL system under the control of the DRL control system. Finally, we compare the energy use of the PFAL under the control of the DRL and the conventional control methods for typical summer and winter conditions in Ithaca, New York. This study employs a PFAL housed within an airtight 40-foot shipping container with dimensions (12.2 m x 2.5 m x 3.0 m) and located in Ithaca, New York. The PFAL relies entirely on electricity to fulfill its energy needs, including heating and cooling. We consider lettuce cultivation as a case study to exemplify the potential of DRL in efficiently managing the PFAL's growing environment while minimizing resource use. In this study, we use typical summer and winter conditions in Ithaca, New York, to demonstrate the effectiveness of the DRL-based automatic control system.

The sampling time *Δt* is chosen as 10 min. This means that the environmental factors in the PFAL are regulated every 10 min. The growing period was taken as 28 days. The photoperiod was selected as 16 h per day. The desired temperature is fixed at 22 – 25 °C during the light period and 18 – 20 °C during the dark period (Ahmed et al., 2020). The relative humidity and CO_2 levels within the PFAL are supposed to be in the range of 70 – 80 % and 800 – 1,200 ppm during the light period. The values of the state and the action variables were scaled to values between 0 and 1 using the upper bounds on the variables. To compare the performance of the DRL control system, we implemented a control system that is made up of on/off control for the lighting systems and proportional control for the environmental variables. This control setup represents the current control technology used in PFALs. The DRL framework was implemented in Python programming language using the Tianshou package (Weng et al., 2022). The training of the DRL agent was conducted on an Intel® Core™ i9-10920X CPU with 3.50 GHz 24-Core processor and 250 GB of random-access memory (RAM), and 2 x NVIDIA GeForce RTX 3080 10 GB video RAM. The neural network for the DRL policy is made up of 2 hidden layers with 256 units per layer and regularized linear units (ReLu) as the activation function.

Figures 1 and 2 show the profiles of the PFAL under the DRL control strategy for typical summer and winter months. Considering that the DRL control objective is to minimize energy consumption while maximizing the crop growth rate, both trajectories lead to higher final crop dry weight, albeit a slightly lower value in the winter case. Ensuring constraint satisfaction is a key challenge in DRL. It can be seen in both figures that the DRL control strategy was able to ensure that the constraints were respected. Although, in some instances during the dark period in the summer case, the upper bound of the CO_2 concentration was violated. In the case of the temperature, however, the constraints were always satisfied due to the prioritization of this variable and its importance in crop growth.

It can be seen from both figures that the same supplemental lighting strategy is employed. There is a gradual increase in light intensity as the crop matures to ensure effective utilization. Similarly, a strategy of low ventilation during the light period and high ventilation during the dark period is used by the DRL control strategy to reduce CO_2 wastage. This strategy is used because, during the light period, elevated CO_2 levels are desired in the PFAL to accelerate crop growth. It can also be seen that the DRL control system does more heating during the winter case and more cooling during the summer case, which is consistent with the outdoor weather conditions. This phenomenon happens because of the consideration of ventilation in the PFAL, which breaks the independence of the PFAL from the outdoor conditions.

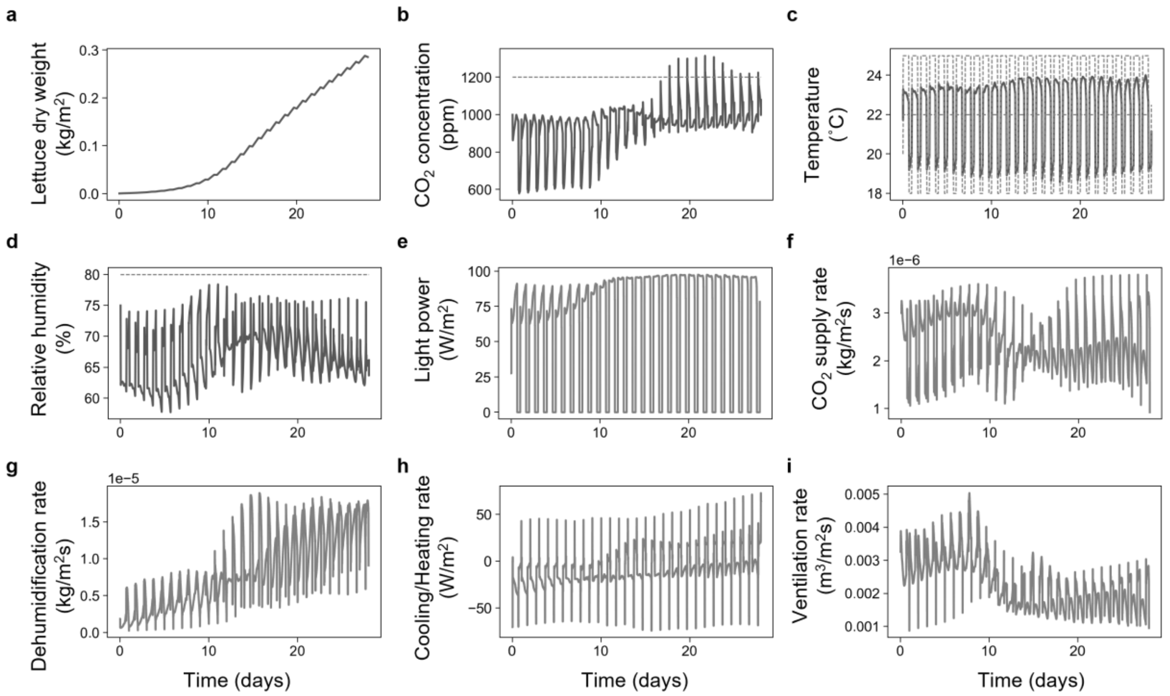

Figure 1: Profiles of the environmental variables (a – d) and the control actions (e – i) of the PFAL under the DRL control strategy for a typical summer month in Ithaca. The solid lines are the profiles of the states and the actions, while the dashed lines are the desired operating bounds

To quantify the ability of the DRL control strategy to streamline energy use in PFALs, we compared its performance to that of the conventional control strategy. As mentioned earlier, the conventional control strategy is made of up on/off control for the artificial lighting system and proportional control for the CO_2 supply, ventilation, dehumidification, and heating/cooling systems. The comparison of the energy use of the two strategies is presented in Table 3. For the two conditions, the conventional control strategy uses more energy per kilogram of crop produced compared to that of the DRL. This could primarily be due to the non-optimality of the conventional control strategy, which leads to excessive resource wastage (Liang et al., 2018). By streamlining operations in the PFAL, the DRL control system can reduce energy consumption resulting in 31 % and 23 % savings in the summer and winter cases. A possible reason for the low percentage of energy savings in the winter case could be due to the colder outdoor conditions, which required reheating of the indoor air mixed with the outdoor air because of ventilation.

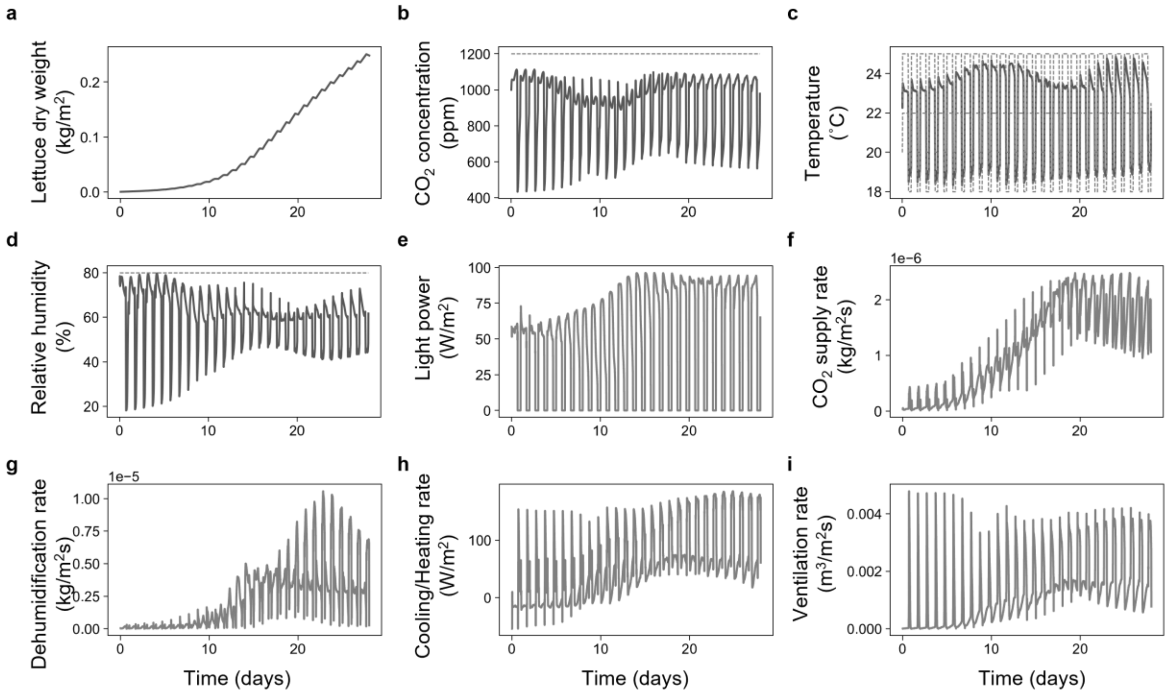

Figure 2: Profiles of the environmental variables (a – d) and the control actions (e – i) of the PFAL under the DRL control strategy for a typical winter month in Ithaca. The solid lines are the profiles of the states and the actions, while the dashed lines are the desired operating bounds

Table 3: Energy use per kilogram of fresh weight for the two control strategies

Control strategy	Conventional	DRL	% savings
Summer (kWh/kg)	9.08	6.27	30.9
Winter (kWh/kg)	9.93	7.68	22.7

4. Conclusions

In this work, we have presented a DRL control framework for energy utilization in PFALs. The framework minimizes the cost of the control activities and maximizes the crop growth in the PFAL while ensuring that the environmental variables remain within the desired operating bounds. We demonstrated the efficiency and performance of the DRL control system using a shipping container PFAL. The DRL system employs a gradual increase in artificial light intensity, and a low light period ventilation, high dark period ventilation strategy to streamline energy and CO_2 utilization in the PFAL. A comparison with the conventional control strategy typically employed in PFAL and made up of on/off control and proportional control shows that the DRL control strategy can save 31 % and 23 % of energy in the summer and winter cases.

References

Ahmed H.A., Yu-Xin T., Qi-Chang Y., 2020, Optimal control of environmental conditions affecting lettuce plant growth in a controlled environment with artificial lighting: A review. African Journal of Botany, 130, 75-89.

Ajagekar A., You F., 2023a, Deep Reinforcement Learning Based Unit Commitment Scheduling under Load and Wind Power Uncertainty. IEEE Transactions on Sustainable Energy, 14, 803-812.

Ajagekar A., Mattson N.S., You F., 2023b, Energy-efficient AI-based Control of Semi-closed Greenhouses Leveraging Robust Optimization in Deep Reinforcement Learning. Advances in Applied Energy, 9, 100119.

Beacham A. M., Vickers L.H., Monaghan J.M., 2019, Vertical farming: a summary of approaches to growing skywards. The Journal of Horticultural Science and Biotechnology, 94(3), 277-283.

Benke K., Tomkins B., 2017, Future food-production systems: vertical farming and controlled-environment agriculture. Sustainability: Science, Practice and Policy, 13(1), 13-26.

Chen W.-H., You F., 2021, 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, 107, 103-113.

Chen W.-H., Mattson N.S., You F., 2022a, Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. Applied Energy, 320, 119334.

Chen W.-H., You F., 2022b, Semiclosed Greenhouse Climate Control Under Uncertainty via Machine Learning and Data-Driven Robust Model Predictive Control. IEEE Transactions on Control Systems Technology, 30, 1186-1197.

Deng X., Dou Y., Hu D., 2018, Robust closed-loop control of vegetable production in plant factory. Computers and Electronics in Agriculture, 155, 244-250.

Eaton M., Shelford T., Cole M., Mattson N., 2023, Modeling resource consumption and carbon emissions associated with lettuce production in plant factories. Journal of Cleaner Production, 384, 135569.

Haarnoja T., Zhou A., Hartikainen K., Tucker G., Ha S., Tan J., Kumar V., Zhu H., Gupta A., Abbeel P., 2018, Soft actor-critic algorithms and applications. arXiv preprint arXiv:1812.05905.

Hu G., You F., 2022, 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, 168, 112790.

Hu G., You F., 2023, An AI framework integrating physics-informed neural network with predictive control for energy-efficient food production in the built environment. Applied Energy, 348, 121450.

Kozai T., 2018, Plant factories with artificial lighting (PFALs): Benefits, problems, and challenges, Chapter In: Smart Plant Factory: The Next Generation Indoor Vertical Farms, Springer, Berlin, Germany, 15-29.

Kozai T., Niu G., Takagaki M., 2019, Plant factory: an indoor vertical farming system for efficient quality food production, Academic Press, Cambridge, Massachusetts, USA.

Liang Y., Hui C.W., You F., 2018, Multi-objective economic-resource-production optimization of sustainable organic mixed farming systems with nutrient recycling. Journal of Cleaner Production, 196, 304-330.

Morimoto T., Torii T., Hashimoto Y., 1995, Optimal control of physiological processes of plants in a green plant factory. Control Engineering Practice, 3, 505-511.

Pennisi G., Orsini F., Blasioli S., Cellini A., Crepaldi A., 2019, Resource use efficiency of indoor lettuce (Lactuca sativa L.) cultivation as affected by red: blue ratio provided by LED lighting. Scientific Reports, 9(1), 1-11.

Reyes-Yanes A., Martinez P., Ahmad R., 2020, Real-time growth rate and fresh weight estimation for little gem romaine lettuce in aquaponic grow beds. Computers and Electronics in Agriculture, 179, 105827.

Sutton R.S., Barto A.G., 2018, Reinforcement learning: An introduction, MIT Press, Cambridge, MA, USA.

Van Delden S.H., Sharath Kumar M., Butturini M., Graamans L.J.A., Heuvelink E., Kacira M., Kaiser E., Klamer R.S., Klerkx L., Kootstra G., Loeber A., 2021, Current status and future challenges in implementing and upscaling vertical farming systems. Nature Food, 2(12), 944-956.

Van Henten E., 1994, Greenhouse climate management: an optimal control approach, PhD Thesis, Wageningen University and Research, The Netherlands.

Vogel E., Donat M.G., Alexander L.V., Meinshausen M., Ray D.K., Karoly D., Meinshausen N., Frieler K., 2019, The effects of climate extremes on global agricultural yields. Environmental Research Letters, 14(5), 054010.

Weng J., Chen H., Yan D., You K., Duburcq A., Zhang M., Su Y., Su H., Zhu J., 2022, Tianshou: A highly modularized deep reinforcement learning library. Journal of Machine Learning Research, 23(267), 1-6.

Xie J., Ajagekar A., You F., 2023, Multi-Agent attention-based deep reinforcement learning for demand response in grid-responsive buildings. Applied Energy, 342, 121162.

Xu D., Ahmed H.A., Tong Y., Yang Q., van Willigenburg L.G., 2021, Optimal control as a tool to investigate the profitability of a Chinese plant factory-lettuce production system. Biosystems Engineering, 208, 319-332.

Zheng Z., Yan P., Chen Y., Cai J., Zhu F., 2021, Increasing Crop Yield Using Agriculture Sensing Data in Smart Plant Factory, In: Wang G., (Eds) SpaCCS 2020. Lecture Notes in Computer Science, 12383, 345-356.