UC Davis

UC Davis Electronic Theses and Dissertations

Title

Closing the Loop in Hydroponic Systems: From Global Sustainability Challenges to AI-Driven Nutrient Management

Permalink

https://escholarship.org/uc/item/2x05k7r6

ISBN

9798290615431

Author

Karimzadeh, Saeed

Publication Date

2025-06-15

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Closing the Loop in Hydroponic Systems: From Global Sustainability Challenges to AI-Driven Nutrient Management

By

SAEED KARIMZADEH

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biological Systems Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Md Shamim Ahamed, Chair

Ali Moghimi

Andre Daccache

Committee in Charge

2025

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Md Shamim Ahamed, for his unwavering support, guidance, and trust throughout this research. His mentorship was instrumental in shaping the direction and execution of this work. I am also sincerely thankful to Dr. Ali Moghimi and Dr. Andre Daccache, members of my thesis committee, for their valuable insights and continued encouragement. Additionally, I would like to acknowledge the support of the qualifying exam committee members—Dr. Stavros Vougioukas, Dr. Majdi Abou Najm, Dr. Brian N. Bailey, and Dr. Paolo D'Odorico—for their thoughtful feedback and contributions to my academic development.

This research was supported by multiple funding sources and institutional partnerships. I gratefully acknowledge financial and technical support from Keysight Technologies and the California Department of Food and Agriculture (CDFA). Additional support was provided by the Department of Land, Air, and Water Resources and the Department of Biological and Agricultural Engineering at UC Davis, as well as the UC Davis Innovation Institute for Food and Health and the UC Davis College of Engineering. I also wish to express my sincere gratitude to Dr. Rulli at Politecnico di Milano, Italy, and Dr. McAllister at Delft University of Technology, Netherlands, for hosting me as an in-absentia PhD researcher during my international research stays. Their guidance and institutional support were instrumental in advancing the global scope of this work.

Dedication

To my beloved family and friends, whose unwavering love and quiet sacrifices have sustained me throughout this journey. To my parents, Habib and Parivash, whose selfless sacrifices and enduring support have shaped every step of my path. Their love and perseverance in the face of life's many challenges have been my foundation. To my family—Vahid, Amin, Neda, and Aeen—thank you for your constant encouragement, belief in me, and the strength of our bond. To my wife, Bita—my steadfast companion, source of strength, and greatest supporter. Her patience, resilience, and boundless encouragement through every challenge have made this achievement possible. This dissertation is as much hers as it is mine.

Abstract

Global population growth and economic development continue to intensify the demand for food production, placing immense pressure on irrigated agriculture, an essential source of sustenance for billions. Controlled environment agriculture (CEA) presents a promising alternative by enabling precise control over environmental conditions, thus improving crop yield and quality. Despite their promise, the practical feasibility and adaptability of closed-loop hydroponic systems remain constrained by various technical and operational challenges, particularly in achieving their full water and nutrient conservation potential. We hypothesize that the transition to closed-loop hydroponics in CEA could be strategically facilitated by aligning system design and precision management with conventional practices currently employed in open-field production. This dissertation explores CEA's potential to enhance the sustainability of food systems, addressing key pain points through a comprehensive engineering lens—from global-scale assessments to farm-level technological innovations.

The first study (Chapter 2) employs a water–nutrient–salinity–energy nexus framework to evaluate the sustainability of transitioning from open-field cultivation to closed-loop hydroponic greenhouse systems. Closed-loop hydroponic systems significantly reduce water use per kg of lettuce by 6% in the UK and over 90% in Bangladesh and India, but this is offset by higher energy demands, ranging from 48 MJ/kg (UK) to 66 MJ/kg (India). Focusing on global lettuce production, this analysis reveals significant improvements in resource efficiency and reductions in environmental impact with greenhouse hydroponic production compared with open fields, especially in water-scarce regions. However, the elevated energy demands of greenhouses remain

a critical trade-off, necessitating integrated solutions and policy considerations to support this transition.

Although hydroponic systems significantly improve resource use efficiency, especially water and nutrients, many commercial operations only partially recirculate nutrient solutions or periodically discharge them, undermining the system's sustainability potential. Achieving fully closed-loop hydroponics is obstructed by two primary barriers: i) sensor inaccuracy (Chapter 3) and ii) nutrient imbalance (Chapter 4). The second study (Chapter 3) investigates operational challenges at the farm scale, particularly the reliability of electrical conductivity (EC) and pH sensors used in hydroponic nutrient solution management. Faulty sensor readings can compromise the precision of fertigation strategies. Leveraging advancements in artificial intelligence, this chapter evaluates the performance of five machine learning models (ANN, RF, SVM, LSTM, and KNN) for fault detection and diagnosis in EC and pH sensors. The five types of synthetic faults (bias, precision, drift, stuck, and physical) were induced into the experimental data from an indoor hydroponic system. RF performs better with accuracy in fault detection, achieving 93.7% for EC and 96.5% for pH sensors. In fault diagnosis, RF also consistently performs better than other models for the EC sensor, with accuracies up to 100% for spike and 97.1% for bias. For the pH sensor, performance varies by fault type: SVM excelled in bias with 100% accuracy, while RF shows the highest accuracy for normal, drift, and precision degradation, reaching up to 92.1%. This chapter also focused on sensor-independent fault detection for EC and pH sensors using environmental data, water volume, and dosing events. The results underscore the effectiveness of sensor-based and sensor-independent approaches to enhance the sustainability of hydroponic production through precision nutrient monitoring and management.

The third study (Chapter 4) advances toward closed-loop nutrient management by shifting from traditional EC-based control to individualized nutrient dosing. This chapter introduces a novel Model Predictive Control (MPC) framework designed to regulate six essential macronutrients: Nitrogen (N), Phosphorus (P), Potassium (K), Calcium (Ca), Magnesium (Mg), and Sulfur (S). This novel implementation of MPC enables precise, dynamic regulation of six selected fertilizers and water inputs, ensuring that all macronutrient concentrations are consistently maintained within agronomically recommended thresholds. Simulations across five scenarios showed that individualized dosing improved water use efficiency (5.9 to 34 g/L) and fertilizer use efficiency (6.2 to 12.5 g/g) while reducing input costs—water (0.32 to 0.08 USD/kg) and fertilizer (1.78 to 0.88 USD/kg).

The outcomes from this research provide critical information on the water-food-energy-sanity nexus for sustainable lettuce production. The outcomes will offer critical insights to various stakeholders about sustainability dynamics and challenges globally. The developed algorithms for fault detection and diagnosis, as well as precision dosing, will provide new insight for improving the sustainability of hydroponic production with precision monitoring and management. Although this research is limited by the absence of an economic feasibility assessment, the findings provide a foundation for the future integration of AI-based decision support tools and adaptive control technologies in commercial hydroponic operations.

Table of Contents

ACKNOWLEDGMENTS ..... **II**

DEDICATION ..... **III**

ABSTRACT ..... **IV**

TABLE OF CONTENTS ..... **VII**

LIST OF TABLES ..... **X**

LIST OF FIGURES ..... **XII**

CHAPTER 1: INTRODUCTION **1**

1.1. BACKGROUND 1

1.1.1. Global challenges in food production *1*

1.1.2. Emergence of controlled environment agriculture (CEA) *3*

1.1.3. Environmental trade-offs and sustainability of CEA *5*

1.1.4. Integrating machine learning and mechanistic approaches in next-generation CEA *8*

1.2. PROBLEM STATEMENT 10

1.3. SIGNIFICANCE OF THE STUDY 11

1.4. RESEARCH OBJECTIVES, HYPOTHESES, AND DISSERTATION STRUCTURE 13

1.4.1. Chapter 2: Nexus frameworks for open-field and hydroponic production *13*

1.4.2. Chapter 3: Fault detection and diagnosis *14*

1.4.3. Chapter 4: Ion-based dosing *15*

1.6. REFERENCES 17

CHAPTER 2: GLOBAL WATER-NUTRIENT-SALINITY-ENERGY NEXUS IN LETTUCE PRODUCTION: FROM OPEN-FIELD IRRIGATION TO CLOSED-LOOP HYDROPONICS IN GREENHOUSES **26**

2.1. ABSTRACT 26

2.2. INTRODUCTION 27

2.3. MATERIALS AND METHODS 32
2.3.1. Lettuce production on a global scale 32
2.3.2. Open-field soil-based production 33
2.3.3. Closed-loop hydroponics in greenhouse production 37
2.3.4. Specific water and energy use per yield comparison 41
2.3.5. Global-scale implementation 42
2.4. RESULTS AND DISCUSSION 46
2.4.1. Water-nutrient-salinity interactions in open-field 46
2.4.2. Water use in hydroponics and potential for sustainable water management 57
2.4.3. Greenhouse energy dynamics 60
2.4.4. Estimation of specific water and energy use 64
2.4.5. Uncertainty analysis 69
2.4.6. Study limitations 72
2.4.7. Prospective avenues for future research 74
2.5. SUMMARY AND CONCLUSION 76
2.6. REFERENCES 79
CHAPTER 3: MACHINE LEARNING-BASED FAULT DETECTION AND DIAGNOSIS OF ELECTRICAL CONDUCTIVITY AND PH SENSORS IN HYDROPONIC SYSTEMS 96
3.1. ABSTRACT 96
3.2. INTRODUCTION 97
3.3. MATERIALS AND METHODS 101
3.4.1. Cultivation conditions and data collection 101
3.4.2. Normal and faulty operation in CEA systems 103
3.4.3. Feature analysis 106
3.4.4. Develop machine learning algorithms 107
3.4.5. Data preprocessing and model parametrization 112
3.4.6. Model performance evaluation metrics 112
3.4.7. Cross-validation 114

3.4. Results and discussion 115
3.4.1. Sensors' behavior 115
3.4.2. Hyperparameter optimization 117
3.4.3. Sensor-based fault detection for EC and pH 119
3.4.4. Sensor-independent fault detection for EC 122
3.4.5. Fault diagnosis for EC and pH 125
3.4.6. Comparative performance evaluation with cross-validation 130
3.4.7. Model selection trade-offs in hydroponic systems 132
3.4.8. Study limitations and future research 133
3.5. Conclusion 135
3.7. References 137
CHAPTER 4: OPTIDOSE: A MODEL PREDICTIVE CONTROL FRAMEWORK FOR PRECISE WATER AND FERTILIZER DOSING IN CLOSED-LOOP HYDROPONICS 144
4.1. Abstract 144
4.2. Introduction 145
4.3. Materials and methods 150
4.3.1. Experimental setup and conditions 150
4.3.2. Formulation of the dynamic model 153
4.3.3. Overview of the MPC framework 155
4.3.4. Scenario modeling 161
4.3.5. Performance metrics 164
4.4. Results and discussion 166
4.4.1. Temporal nutrient dynamics 166
4.4.2. Optimizing resource use under precision nutrient management scenarios 168
4.4.3. Water and fertilizer consumption and efficiency 172
4.4.4. Economic implications 173
4.4.5. Study limitations and future direction 173
4.5. Conclusion 176

4.6. REFERENCES 177

CHAPTER 5: CONCLUSIONS AND FUTURE RECOMMENDATIONS 186

5.1. CONCLUDING REMARKS 186

5.2. BROADER POTENTIAL IMPACTS 189

5.3. LIMITATIONS AND FUTURE DIRECTION 190

5.4. REFERENCES 193

APPENDICES 194

APPENDIX A. CHAPTER 2 SUPPLEMENTARY MATERIAL 194

List of acronyms, abbreviations, symbols, and subscripts 194

APPENDIX B. CHAPTER 3 SUPPLEMENTARY MATERIAL 216

APPENDIX C. CHAPTER 4 SUPPLEMENTARY MATERIAL 225

COPYRIGHT PERMISSIONS 226

List of Tables

Table 1. Dissertation chapters and related publications. ____ *16*

Table 2. Other publications resulting from this dissertation. ____ *17*

Table 3. Comparative analysis of water consumption and irrigation salt-score (ISS) in open field lettuce cultivation (furrow and drip systems) across the top 15 lettuce-producing countries based on harvested areas. ____ *54*

Table 4. Comparative analysis of country-level average specific water use (SWU) and specific energy use (SEU) for lettuce production under open-field cultivation systems (furrow and drip irrigation) vs. closed-loop hydroponics in greenhouse systems across the top 15 major lettuce-producing countries. ____ *68*

Table 5. Summary of mathematical formulations, fault parameters, and typical causes of EC and pH sensor faults in the hydroponic nutrient solution (Jan et al., 2021). ____ *104*

Table 6. Confusion matrix in the context of a binary classification task. ____ *113*

Table 7. Comparison of five machine learning classifiers for fault diagnosis of EC and pH sensors across six fault categories. RF emerged as the most reliable classifier for diagnosing faults in EC and pH sensors, as well as distinguishing normal operational states. Most models demonstrated high accuracy in diagnosing sensor stuck faults compared to other fault types. 127

Table 8. List of selected fertilizer compounds used in the nutrient solution formulation, including their chemical names, compositions, and market price (USD/kg). 158

Table 9. Suitable lower and upper concentration bounds (in parts per million, ppm) for essential macronutrients in nutrient solution formulations and uptake rate in lettuce production (Albornoz & Lieth, 2016; Mattson & Peters, 2014). 160

Table 10. Performance of different nutrient and water optimization strategies compared to a baseline scenario. Metrics include total water and fertilizer consumption, wastewater generation, nutrient deficiency and toxicity index, resource use efficiency, and associated costs. 173

Table 11. Lettuce parameters used in this study related to evapotranspiration, evaporation, and salinity. 198

Table 12. Comparison of irrigation and fertigation strategies for salinity simulation in open-field lettuce production. 199

Table 13. Lettuce cultivation areas across the diverse Köppen-Geiger climate classes worldwide. 200

Table 14. List of the variables used in order to perform the analysis in this study. 202

Table 15. Relative changes (%) in water use, nutrient leaching, and salinity stress comparing El Niño (2015–2016) to the 5-year average (2019–2023). 209

Table 16. Impact of narrowing greenhouse temperature setpoints on energy requirements across global climate zones. 214

Table 17. Hyperparameter tuning results for five machine learning models in classifying EC and pH sensors for fault detection and diagnosis. 216

Table 18. Performance metrics (precision, recall, f1-score, and overall accuracy) in predicting EC and pH across different sensor fault detection models (KNN, SVM, RF, ANN, LSTM). 220

Table 19. Performance metrics (precision, recall, and f1-score) in predicting EC and pH across different sensor fault diagnosis models: (a) KNN, (b) SVM, (c) RF, (d) ANN, and (e) LSTM. 220

Table 20. Pseudocode of the approach for sensor fault detection and diagnosis. 222

Table 21. Schedule of water and nutrient solution replenishment events in the nutrient solution tank during the experimental period. 225

List of Figures

Figure 1. Flowchart of methodology: Schematic representation of the research goal, methodological framework, and datasets used in this study. 45

Figure 2. Global water dynamics in lettuce cultivation (2015-2016), showing: a) green water consumption (rainwater absorbed by soil), b) blue water withdrawal (extraction of surface and groundwater), and c) differences between blue and green water with furrow and drip irrigations. 48

Figure 3. Comparison of green and blue water usage in lettuce production with furrow (a) and drip (b) irrigation and their differences (c) across 30 Köppen-Geiger climate classifications. Boxplots are color-coded to represent the five primary climate groups: A (tropical, green), B (arid, red), C (temperate, blue), D (continental, orange), and E (polar, violet). 49

Figure 4. The impact of applying various irrigation electrical conductivity (EC) of 0.5, 1.0, 1.5, and 2.0 dS/m, employing furrow (a) and drip (b) systems. 50

Figure 5. Lettuce salt-affected areas using (a) drip and (b) furrow strategies at irrigation EC levels of 0.5, 1.0, 1.5, and 2.0 dS/m across 30 Köppen-Geiger climate classes (c) and major lettuce-producing countries (d). This figure illustrates the adaptability of climate conditions and the impact of EC levels on an irrigation strategy for lettuce cultivation, providing concise insights into global agricultural practices and water management strategies. 52

Figure 6. A comparative analysis of nutrient leaching into groundwater resources across various irrigation techniques and salinity levels. Higher salinity reflects greater dissolution of inorganic salt fertilizers in irrigation water. The figure includes a) country-level and b) global maps illustrating nutrient leaching from furrow and drip irrigation methods at an EC of 2.0 dS/m. 56

Figure 7. Global mapping of greenhouse systems water sustainability: a) total annual water requirement, b) the ratio of water for the pad and fan system over hydroponic water requirement, c) annual water required for the pad and fan system to meet cooling demand, d) number of months with sustainable water use (the amount of harvested blue water exceeds the crop water requirement) for lettuce greenhouse production ("0" indicating no months exceeding

requirement), e) percent of total cultivated area with sustainable water consumption across various months, countries, and climate classification. 59

Figure 8. Global spatial distribution of heating and cooling energy requirements for lettuce greenhouse production (a), accompanied by boxplots detailing the load distribution for climate classes (b), average energy fluxes in countries, and climate classes (c). 64

Figure 9. Global spatial variation in specific water (SWU) and energy (SEU) use per kilogram of lettuce production under contrasting agricultural systems, a) specific water use (L/kg) and specific energy use (MJ/kg) in greenhouses using hydroponic systems, b) specific water use (L/kg) in open-field agriculture under drip and furrow irrigation systems. 67

Figure 10. Overview of the experimental setup and sensor installations (a and f), time series data depicting climate variables, temperature (b) and relative humidity (c), nutrient solution volume and dosing (d), and spectral composition analysis of the fluorescent growing lights (e). The highlighted yellow section indicates the 16-hour photoperiod during which plants were exposed to light. 103

Figure 11. EC and pH sensing in the lettuce hydroponic system, illustrating normal data acquisition and synthetic faults. Subfigures (a) through (e) depict simulated faults for EC: bias, drift, precision degradation, stuck, and spike. Subfigures (f) through (j) illustrate the corresponding faults for pH. 106

Figure 12. Comprehensive methodological workflow for sensor-based fault detection and diagnosis in hydroponic sensors, consisting of four steps: (1) data acquisition from sensors, (2) data preparation with synthetic fault generation and class balancing, (3) model training using five machine learning algorithms, and (4) validation for fault detection and diagnosis. 107

Figure 13. Violin plots comparing normal and faulty observations (including bias, drift, precision degradation, stuck, and spike) for EC (a) and pH (b) sensors in the hydroponic system. Each subplot represents a specific fault scenario, including bias (up and down), drift (up and down), precision degradation, spike, and stuck sensor behavior. The red and green horizontal lines indicate the mean and median values, respectively, while the dashed black lines mark the 25th and 75th percentiles. 116

Figure 14. Confusion matrices of sensor-based fault detection in hydroponic solution monitoring systems using five selected machine learning models. The figure illustrates sensor faults compared to normal readings in two commonly used sensors: EC (a) and pH (b). Each confusion matrix provides a detailed comparison between the predicted

observations (x-axis) and the actual observations (y-axis). RF provided the most robust sensor-based fault detection performance in hydroponic systems. 121

Figure 15. Time-series comparison of observed and predicted EC in a hydroponic lettuce system using (a) SVR and (b) ANN models, based on climate variables and dosing signals. Predicted values (red) closely follow observed values (blue), with ANN showing higher accuracy and better sensor-independent fault detection. Performance metrics include MAE, RMSE, and R^2. 124

Figure 16. Confusion matrices of sensor-based fault diagnosis for EC (a) and pH (b) sensors in hydroponic solution monitoring systems using five selected machine learning models. Each confusion matrix compares predicted fault categories (x-axis) against actual fault categories (y-axis), including normal, bias, drift, precision degradation, spike, and stuck conditions. RF is the most reliable choice for multi-class fault diagnosis in hydroponic systems. 129

Figure 17. Cross-validation performance of machine learning models for EC and pH sensor's fault detection (a), diagnosis (b), and EC prediction (c). RF demonstrates the highest robustness and reliability for both fault detection and multi-class fault diagnosis in EC and pH sensors. For EC prediction, SVR shows lower uncertainty compared to ANN. 131

Figure 18. (a) Lettuce growth progression from transplanting (Feb 22, 2024) to near-harvest (Mar 18, 2024), (b) Schematic of the closed-loop NFT hydroponic system with two vertical tiers, (c) Time series of measured parameters: air temperature, humidity, EC, and nutrient solution volume, illustrating system dynamics and nutrient replenishment events. 152

Figure 19. Monitoring and modeling of plant growth and evapotranspiration. (a) Images of the measured individual plants across three rows of NFT in the lower tier levels. (b) Biomass accumulation trends during the growing season. Left: Individual plant biomass growth curves. Right: Comparison between experimental average biomass (blue dots) and model-predicted values (red line). (c) Temporal dynamics of evapotranspiration. Left: Daily evapotranspiration rate ($ET_{c,daily}$), and right: evapotranspiration per kg of lettuce ($ET_{c,\ per\ kg}$). 154

Figure 20. Overview of the proposed control framework for optimal macronutrient dosing for hydroponic production. 156

Figure 21. Temporal dynamics of nutrient concentrations in the tank under different fertigation scenarios. Simulated daily concentrations (mg/L) of six essential nutrients—nitrogen (N), phosphorus (P), potassium (K), calcium (Ca),

magnesium (Mg), and sulfur (S)—throughout a 30-day growing cycle (days after transplanting, DAT) of lettuce under five fertigation strategies: Baseline, Scenario 1, Scenario 2, Scenario 3, and Scenario 4. 168

Figure 22. Dynamic water management under various nutrient delivery scenarios. Water dynamics are shown across five strategies over a 30-day cultivation. Top panels: Temporal evolution of total water volume within the system. Bottom panels: Daily volume of water added to maintain nutrient solution levels 170

Figure 23. Temporal fertilizer dosing profiles across various nutrient delivery scenarios. Fertilizer dosing amounts (grams) are shown for six key macronutrient sources: Calcium Nitrate, Potassium Phosphate, Magnesium Nitrate, Magnesium Sulfate, Potassium Nitrate, and Potassium Sulfate—over a 30-day growing cycle. 171

Figure 24. Cultivated area distribution (a) (Monfreda et al., 2008) and trends in lettuce yield, harvested area (b), and global nitrogen, phosphorous, and potassium fertilizer application (c). 197

Figure 25. The global distribution of Köppen-Geiger climate classifications indicates 30 distinct climate zones worldwide (Beck et al., 2018). 200

Figure 26. Comparing sustainable regions for lettuce greenhouse production by utilizing harvested blue water for hydroponics. 202

Figure 27. Monthly energy fluxes fluctuation in China (largest lettuce producer country), and in Cfa (temperate, without dry season, hot summer, the largest lettuce cultivated climate class). 203

Figure 28. Global relative differences (%) in blue and green water use between the El Niño period (2015–2016) and the 5-year average (2019–2023) under different irrigation systems. 210

Figure 29. Relative differences (%) in nutrient leaching and salinity stress coefficient between the El Niño period (2015–2016) and the 5-year average (2019–2023) under different irrigation systems and water salinities. 211

Figure 30. Relative deviation in precipitation (%) between the El Niño period (2015–2016) and the 5-year average (2019–2023). 212

Figure 31. Spatial distribution of annual heating energy demand across Iran. (a) Estimated annual heating load (MJ/m2) based on our high-resolution thermal energy balance model; (b) corresponding annual heating demand (GJ) estimated by Nasrollahi et al. (2021). 213

Figure 32. Global impact of narrowing greenhouse temperature setpoints on energy demand. Relative difference (%) in (a) cooling, (b) heating energy requirements, (c) longwave radiation, (d) conduction and convection, (e) greenhouse floor, and (f) infiltration. 214

Figure 33. Pearson correlation coefficients between climate variables and dosing pulses with EC (a) and between EC and pH with three previous observations (b, c). 219

Figure 34. Schematic representation of k-fold cross-validation for evaluating model performance using EC and pH datasets. 224

Chapter 1: Introduction

1.1. Background

1.1.1. Global challenges in food production

The dual pressures of population growth and economic advancement necessitate a significant increase in global food production. The global population is anticipated to reach 9.6 billion by 2050 and could potentially climb to 10.9 billion by 2100 (Gerland et al., 2014). As the global population continues to expand and societal demands for diverse food options rise, the ability to meet the nutritional needs of this discerning world hinges substantially on our capacity to uphold and sustain productive agricultural soils. Urbanization is growing, with city dwellers expected to increase from 3.6 billion in 2011 to 6.3 billion by 2050 (United Nations, 2014). As cities expand, increasingly complex supply chains will exert profound impacts on both urban and rural systems (Kennedy et al., 2007). In response to these challenges, urban horticulture has emerged as a promising strategy to localize food production, reduce supply chain vulnerabilities, and enhance urban food security (Benke & Tomkins, 2017).

Open-field agriculture remains the dominant form of food production globally, supplying over 90% of the world's vegetables and staple crops (FAO, 2021). At the same time, climate change and human-induced environmental disruptions have placed unprecedented stress on freshwater resources. Irrigation, which accounts for approximately 70% of global surface and groundwater withdrawals (Foley et al., 2011; Siebert et al., 2010), is central to global food supply systems. As irrigated agriculture remains the cornerstone of food production for billions, achieving sustainable food security is fundamentally tied to the sustainability of irrigation systems (Schmitt et al., 2022).

Moreover, the continued expansion and intensification of conventional agricultural practices to satisfy rising global food demand carry significant environmental risks. These practices are often linked to widespread ecological degradation, including deforestation, nutrient runoff leading to eutrophication, loss of biodiversity, soil salinity intensification, and deterioration of air and water quality (Karimzadeh et al., 2024; Malone & Newton, 2020; Wade et al., 2020).

Soil processes play a pivotal role in influencing the quality of the food supply, potable water sources, and the air we breathe. The significance of soil cannot be overstated in relation to human well-being and sustenance. Current estimates indicate that approximately 15% of the world's total land area has undergone degradation (Wild, 2003). On the other hand, approximately one-fifth of the cultivated soil is subjected to irrigation, playing a role in around 40% of the world's food production (Rosa & D'Odorico, 2019). Concurrently, the growing populace is placing additional strain on freshwater resources, intensifying the need for water in both domestic and industrial sectors. While escalating demands for maintaining healthy freshwater environments and ecosystems are diverting increasing volumes of water, many Mediterranean and semi-arid regions are witnessing restrictions on water availability for irrigated agriculture. For example, the Sustainable Groundwater Management Act (SGMA) in California limits agricultural water use, which also negatively impacts agricultural production. Public Policy Institute of California (PPIC) reports that up to 900,000 acres of California farmland could be forced out of production (or fallowed) due to water scarcity (California Farm Water Coalition, 2024). In light of the global constraints on high-quality freshwater availability, enhancing the water use efficiency of irrigated agriculture has become imperative (Hopmans et al., 2021). This form of agricultural intensification aims to achieve more with less, minimizing environmental impact and addressing contributions to,

or adaptations to, climate change. In this context, rethinking traditional agricultural paradigms is essential to ensure resilient food systems for future generations.

1.1.2. Emergence of controlled environment agriculture (CEA)

Controlled Environment Agriculture (CEA) presents a promising substitute for the existing food supply chain and a possible alternative for local food production in regions with unsustainable open field production. CEA encompasses a spectrum, starting from basic shade structures and hoop houses, progressing to greenhouses, and extending to complete indoor vertical farms. The most sophisticated setups involve fully automated, closed-loop systems that regulate temperature, humidity, CO_2, lighting, water, and fertilizer. These systems are capable of resource-efficient and optimal crop production with supplemental lighting for greenhouses and artificial lighting for vertical farming (SharathKumar et al., 2020).

The economic outlook for greenhouse and indoor vertical farming reflects a rapidly expanding market. The global indoor farming (greenhouses and vertical farming) market was valued at USD 45.97 billion in 2024 and is projected to grow to USD 138.09 billion by 2033 at a compound annual growth rate (CAGR) of 13% (Bothare, 2025). This growth is fueled by factors such as year-round production, resource efficiency, local production, and consumer demand for pesticide-free, nutrient-rich produce. Compared to traditional farming, indoor systems—particularly hydroponics and vertical farming—offer significantly higher yields per square meter and reduced land and water use. Greenhouses remain dominant in several regions, especially in Asia and North America, contributing substantially to fresh produce markets. However, the economic viability of these systems remains challenging due to high upfront capital costs, energy demands (especially for lighting and climate control), and skilled labor requirements. For instance, energy expenditures can account for 12–25% of an indoor farm's operational budget, with breakeven electricity prices

ranging from USD 3–5/kWh in the U.S., making profitability highly sensitive to location and infrastructure efficiency (Bothare, 2025). Nonetheless, technological innovations and strong investor interest create opportunities to lower costs and improve scalability, making both greenhouse and indoor farming central to the future of climate-adaptive, urban-centered agriculture.

In CEA, crops may be grown hydroponically, where roots are bathed in nutrient-dense water, or aeroponically, where roots are regularly misted with water and nutrients. Water and nutrient management challenges are crucial to address in hydroponic systems for enhanced crop yield and reduced environmental impact. CEA systems often outperform conventional field agriculture regarding productivity and crop quality by enabling precise control over environmental parameters. However, CEA is not a replacement for open-field agriculture since these systems alone are unlikely to produce sufficient crop yields to fully satisfy the caloric and nutritional demands of urban populations (Dsouza et al., 2023) due to scalability constraints, crop type limitations, and its high energy demand, particularly in vertical farming systems (Al-Chalabi, 2015; Barbosa et al., 2015). Over the past four decades, yields from CEA systems—particularly for crops such as tomato and sweet pepper—have more than doubled (Heuvelink et al., 2024). This improvement is attributed to technological advancements in greenhouse design (e.g., enhanced light transmissivity, where a 1% increase in light can lead to a 0.5–1% rise in yield; Marcelis et al., 2006), the adoption of modern soilless cultivation methods, climate control systems, and significant progress in crop genetics (Higashide & Heuvelink, 2009). A key environmental advantage of certain CEA systems lies in their closed-loop design, which offers greater insulation from external climatic variability, including fluctuations in precipitation and temperature. This enhanced control increases system resilience to environmental stressors and significantly reduces

exposure to pests and pathogens. As a result, pesticide application—and the associated environmental contamination—is substantially lower than conventional open-field agriculture. Moreover, CEA systems, particularly hydroponic and aeroponic setups, enable precise nutrient management and recirculation (Goddek et al., 2019; Rufí-Salís et al., 2020). This capacity enhances nutrient use efficiency while eliminating nutrient leaching and mitigating phosphorus and nitrogen pollution (Cowan et al., 2022).

1.1.3. Environmental trade-offs and sustainability of CEA

Despite its benefits, the widespread adoption of CEA is constrained by its technological complexity, high operational costs, and social acceptance barriers, all of which pose challenges to its scalability and integration into mainstream food systems (Cowan et al., 2022). The resource-intensive nature of CEA also raises important questions about its long-term sustainability. Several studies have quantified the environmental benefits of CEA, including reductions in transportation, water usage, and pesticide application. However, uncertainty persists regarding whether CEA systems are ultimately more environmentally sustainable than conventional open-field agricultural practices (Iddio et al., 2020). For instance, a life cycle assessment (LCA) that included a comprehensive evaluation of inputs such as chemicals and fertilizers, irrigation systems, cultivation machinery, harvesting operations, and the transportation and refrigeration of produce to distribution centers found that field-grown vegetables have a lower global warming potential (GWP) of approximately 0.37 kg CO_2-eq/kg. In comparison, vegetables grown in passive greenhouses exhibited a GWP of 1.10 kg CO_2-eq/kg, while those cultivated in heated greenhouses reached 2.13 kg CO_2-eq/kg (Zarei et al., 2019). Khoshnevisan et al. (2014) reported that greenhouse strawberry production in Gilan, Iran, consumes vastly more energy and emits significantly higher greenhouse gases (35,083.5 kg CO_2-eq/ha) compared to open-field cultivation

(803.4 kg CO_2-eq/ha). However, it is essential to interpret LCA findings with caution, as results are highly sensitive to the scope, boundaries, and methodological assumptions of the analysis (Hendrickson et al., 2010) as strategies such as improving heating efficiency, adopting renewable energy sources, and optimizing construction material selection could substantially reduce GWP in greenhouse systems. For example, in the Netherlands, natural gas consumption for greenhouse tomato heating has significantly declined from 55 m^3/m^2 in the early 1980s to approximately 40 m^3/m^2 today. As a result, energy use efficiency has improved by a factor of 3.5, reaching an output of around 2 kg of tomatoes per cubic meter of gas (Heuvelink et al., 2024). In experimental settings, even greater efficiency has been achieved: tomato yields of 75 kg/m^2 have been produced using only 12 m^3 of natural gas and 10 kWh of electricity, within greenhouses equipped with multiple thermal screens and active dehumidification systems (Heuvelink et al., 2024).

In lettuce production, CEA has the potential to be one of the least sustainable food production methods if implemented without careful planning, as it presents several environmental hotspots. However, when thoughtfully designed and appropriately scaled, particularly in regions characterized by high food demand and limited agro-climatic suitability for horticultural production, CEA can contribute meaningfully to the transformation toward more sustainable food systems (Casey et al., 2022). In Scotland, Sandison et al. (2023) found that electricity use accounts for 91% of vertical farming's carbon footprint, with emissions initially at 1.49 $kgCO_2$-eq/kg under the 2019 Scottish energy mix—higher than conventional open field farming—but these dropped to 0.42 $kgCO_2$-eq/kg by 2020 with more renewables, and could fall further to 0.33 $kgCO_2$-eq/kg under 100% renewable energy, making vertical farming a low-carbon solution (lower than open-field production). Martin et al. (2023) conducted a comprehensive LCA to evaluate the environmental performance of a large-scale commercial vertical farm in Sweden producing

packaged lettuce. Their analysis adopted a cradle-to-grave perspective and compared vertically farmed lettuce with domestically produced and imported conventionally grown lettuce. The vertical farm demonstrated lower greenhouse gas emissions relative to conventional systems; however, it exhibited higher environmental impacts in other categories, primarily due to its significant electricity consumption. Regarding energy use efficiency, Pennisi et al. (2019) reported that vertical farming systems can produce approximately 90 grams of fresh lettuce per kilowatt-hour when using a red:blue light ratio of 3, translating to around 11 kWh required to produce 1 kg of lettuce. This estimate aligns closely with the 10 kWh/kg figure reported by Martin et al. (2023). Nonetheless, energy efficiency values vary widely in the literature; for instance, Graamans et al. (2018) reported a considerably higher energy requirement of 17 kWh per kg of fresh lettuce. Such discrepancies can be attributed to differences in data sources, such as commercial-scale operations versus controlled experiments, as well as underlying methodological assumptions (Heuvelink et al., 2024).

The sustainability of agricultural production systems varies widely, prompting the United Nations to establish the 2030 Agenda for Sustainable Development in 2015, which outlines 17 Sustainable Development Goals (SDGs) aimed at guiding global efforts toward a more sustainable future (UN, 2015). Within this framework, Zhou et al. (2021) a comparative analysis of tomato production across four greenhouse systems was conducted, contrasting high-tech greenhouses in the Netherlands with low-tech greenhouses in Spain and evaluating both conventional and organic cultivation approaches. Their findings highlight that high-tech, soilless greenhouses with nutrient solution recirculation make a meaningful contribution to several SDGs by reducing water and nutrient losses. However, these systems rely heavily on fossil energy, which remains a major environmental drawback. Conversely, soil-based systems, particularly those without recirculation,

such as free-drainage setups, are associated with high water consumption and nitrogen leaching, affecting soilless systems when recirculation is not implemented (Massa et al., 2010).

The *Nexus* approach provides a comprehensive framework for understanding and managing the interconnections among water, energy, and food systems. By emphasizing the interdependence of these sectors, the approach aims to promote integrated decision-making that maximizes co-benefits and minimizes trade-offs (Irabien & Darton, 2016). Neglecting these interlinkages can result in unintended consequences that undermine sustainability and resource security. Ensuring access to water, energy, and food is essential for human development; however, pursuing security in one domain without regard for the others can disrupt the overall balance. El-Gafy et al. (2017) demonstrated that the nexus approach, in contrast to conventional non-nexus strategies, effectively integrates economic and environmental considerations, particularly within food and energy systems.

1.1.4. Integrating machine learning and mechanistic approaches in next-generation CEA

The next generation of CEA calls for dynamic control systems enhanced by machine learning (ML). Practitioners have increasingly advocated for ML-driven solutions that respond in real-time to plant needs. However, the widespread adoption of this approach has been limited by high data requirements and the slow pace of plant observation, both critical for training accurate predictive models (Cohen et al., 2022). While ML excels at uncovering complex patterns without needing explicit cause-and-effect relationships, mechanistic models offer foundational biological insights that can help structure and constrain ML models. Integrating both reduces the required training

data, lowers validation costs, and improves model interpretability, effectively opening the "black box" of ML predictions (Cohen et al., 2022).

Previous modeling efforts have revealed significant knowledge gaps in how multiple control variables interact in plant growth systems (Fiorani & Schurr, 2013; Griffiths & York, 2020). ML is well-suited to handle this complexity, yet it remains resource-intensive, especially due to the high replication demands of purely statistical models and the throughput limitations of current phenotyping technologies. Hybrid models that combine ML with mechanistic plant models are emerging as a promising solution, although their application in CEA remains limited (Cohen et al., 2022).

Monitoring nutrient concentrations is more feasible in hydroponic systems, which offer a more controlled context than soil-based agriculture (Tomasi et al., 2015). Conventional commercial-scale hydroponic production involves monitoring nutrient levels using electrical conductivity (EC) and potential of hydrogen (pH) measurements (Bacci et al., 2005; Bamsey et al., 2012). Accurate and reliable pH, EC, and nutrient concentration measurements are crucial in closed-loop hydroponic systems. These systems are vital for sustainable soilless cultures, as they enable resource savings and minimize the discharge of nutrients into the environment (Bouchaaba et al., 2015; Gruda, 2019). Meanwhile, the rising concern over environmental pollution from excessive fertilizer use has increased the emphasis on sustainable fertilization practices. This involves aligning nutrient supply with plant demand (Mattson & Van Iersel, 2011). While EC remains a widely used and cost-effective indicator of overall nutrient concentration, it does not provide information about individual ion levels, which can pose limitations in recirculating (Bamsey et al., 2012; Lee et al., 2017; Savvas et al., 2023). Farmers often rely on their experience to adjust EC dynamically, which has proven effective in many cases. However, EC readings may be affected

by sensor inaccuracies or shifts in ion composition, potentially leading to suboptimal decisions regarding nutrient replenishment. On the other hand, there is potential to enhance nutrient management through adaptive strategies that monitor and control individual nutrient concentrations. Achieving this requires a more comprehensive and precise framework for fertilizer and water dosing tailored to specific crop demands.

To overcome this, root-zone sensors, time-series monitoring of supply and drainage, and integrating empirical and theoretical models are employed to capture nutrient dynamics. The use of ion-selective electrodes (ISEs) with digital data management systems can enhance this process (Cho et al., 2017; Kim et al., 2023). However, in situ ion detection can suffer from signal interference. Artificial neural networks (ANNs) have shown promise in correcting these issues, particularly when combined with two-point normalization techniques (Cho et al., 2019). This approach has demonstrated accuracy in monitoring ions such as NO_3^-, K^+, and Ca^{2+}, providing a more refined and adaptive method for nutrient control.

Ultimately, the dynamic and nonlinear nature of plant growth, nutrient uptake efficiency, and genotypic-environmental interactions presents challenges for traditional mechanistic modeling. By leveraging the strengths of ML, future CEA systems can achieve a more responsive, efficient, and sustainable form of agricultural management (Cohen et al., 2022).

1.2. Problem statement

While CEA cannot entirely replace conventional open-field farming, it offers a viable complementary approach, particularly for leafy crops like lettuce. However, the sustainability and scalability of hydroponic systems hinge on addressing several persistent operational and technological challenges. A global-scale analysis provides insights into where hydroponic systems

offer a strategic sustainability advantage in expanding. It also identifies regions where conventional systems remain viable and where policy interventions such as targeted subsidies or infrastructure investment may be necessary to support a transition. This dual-layered approach—combining technological innovation with strategic, location-specific analysis—offers actionable pathways for both growers and policymakers, which is not well developed yet.

While hydroponic systems are often praised for their water and nutrient efficiency, their true sustainability is limited by critical operational challenges—foremost among them is suboptimal nutrient solution management. Currently, most commercial hydroponic systems rely on EC as a proxy for managing nutrient concentrations. EC is undeniably a valuable metric, offering a fast and accessible means to monitor total dissolved ions. However, relying solely on EC can result in unbalanced nutrient solutions, reduced crop performance, and frequent discharge of excess solution. On the other hand, inaccurate readings from EC and pH sensors—critical parameters for nutrient management—can also lead to imbalanced nutrient formulations and suboptimal solution preparation, resulting in nutrient waste and impaired plant performance. This discharge undermines the environmental benefits of closed-loop systems by wasting fertilizers and water and contributing to nutrient pollution. Therefore, developing nutrient-specific optimization algorithms that move beyond EC alone enables ion-targeted dosing that aligns with the physiological needs of crops to supplement nutrient management in addition to EC-based.

1.3. Significance of the study

This dissertation addresses pressing questions at the intersection of food security, sustainability, and technological innovation in agricultural systems. As global challenges related to climate change, resource scarcity, and population growth intensify, it is critical to reassess conventional farming practices and explore the potential of hydroponic production as a viable alternative. The

significance of this research lies in its comprehensive evaluation of open-field agriculture versus hydroponic greenhouses, its investigation into the sustainability of current hydroponics operations in greenhouses and vertical farms, and its contributions to optimizing resource use through advanced technological and agronomic strategies.

Specifically, this dissertation aims to shed light on three key questions:

1. **How do open-field and hydroponic systems compare globally, and is sustainable greenhouse production achievable under current practices?**

   This study provides a global-scale assessment of open-field versus closed-loop hydroponic systems for lettuce production, evaluating their performance across diverse regions, countries, and climatic zones. Quantifying water use as a key sustainability indicator, it examines whether current hydroponics in greenhouse practices can meet resource demands, which is crucial for water-scarce areas such as California. The findings offer critical insights for shaping policy and guiding future investments in hydroponic greenhouse infrastructure and management.

2. **How can machine learning models enhance hydroponic solution management and enable closed-loop hydroponic systems?**

   The study explores integrating machine learning models for fault detection and diagnosis to enable AI-driven decision support systems to improve the precision, efficiency, and automation of resource management in hydroponics, focusing on reducing waste and achieving fully closed-loop nutrient and water cycles.

3. **What is the impact of rethinking nutrient management by developing algorithms that precisely dose fertilizers and water to achieve target nutrient concentrations?** Moving beyond the traditional reliance on EC as a proxy for nutrient concentration, this

work investigates a paradigm shift toward ion-specific optimization of nutrient solutions. This approach allows for more accurate control of nutrient composition and enhanced efficiency in closed-loop hydroponic systems.

This dissertation addresses these questions through three core research objectives, contributing to advancing the scientific foundation for sustainable and technologically empowered hydroponic agriculture. The findings offer actionable insights for growers, researchers, policymakers, and technology developers seeking to accelerate the transition toward closed-loop, resource-efficient hydroponic production.

1.4. Research objectives, hypotheses, and dissertation structure

In response to the abovementioned challenges, this dissertation presents a comprehensive evaluation of hydroponic production, assessing its performance in contrast to conventional agriculture across multiple scales. The analysis spans from global-scale comparisons to farm-level operational insights in three main research objectives. The objectives of each chapter are outlined as follows:

1.4.1. Chapter 2: Nexus frameworks for open-field and hydroponic production

Hypothesis:

The choice of irrigation techniques significantly affects crop water demand, soil salinity, and nutrient leaching in open-field lettuce systems, while energy requirements for heating and cooling of greenhouse facilities vary by geography and season. By spatially integrating salinity risk, irrigation practices, and water quality data, it is possible to identify environmentally suitable zones for open-field production and regions where greenhouses present greater agronomic or environmental advantages.

Objectives:

This chapter introduces a nexus-based framework to assess the benefits and trade-offs of transitioning from conventional open-field agriculture to advanced hydroponic production in controlled environment greenhouses. The analysis is conducted at a global scale, encompassing diverse climatic zones and environmental conditions, with a particular focus on the world's leading lettuce-producing regions. The findings offer strategic insights for policymakers, helping to determine where conventional open-field systems remain sustainable and where targeted subsidies and support mechanisms may be necessary to facilitate a transition toward a soil-less hydroponic system.

Chapters 3 and 4 focus on advancing closed-loop hydroponic operations, a critical component for the long-term sustainability of hydroponic systems as highlighted in Chapter 1. These chapters address two complementary approaches toward achieving fully closed-loop hydroponic nutrient management: (1) employing machine learning models for accurate fault detection and diagnosis of sensors used in nutrient solution monitoring, and (2) moving beyond reliance on EC alone by developing optimization algorithms for precise fertilizer dosing, ensuring the key macronutrient concentrations remain within optimal ranges.

1.4.2. Chapter 3: Fault detection and diagnosis

Hypothesis:

Due to potential inaccuracies in nutrient solution sensors, machine learning models trained on simulated fault data, when combined with environmental and operational variables, can enhance fault detection and diagnosis accuracy while enabling reliable, sensor-independent prediction of nutrient solution parameters.

Objective:

Achieving a fully closed-loop hydroponic system, where nutrient solutions are efficiently recirculated, remains both a critical goal and a significant challenge in modern CEA facilities. This chapter investigates using artificial intelligence to enhance nutrient solution management by addressing common issues and inaccuracies in EC and pH sensor performance. Given the central role of sensors in monitoring and adjusting nutrient composition, their reliability is essential for system sustainability. Data-driven approaches are therefore employed to detect and diagnose sensor faults, enabling automated alerts to growers and facilitating more accurate and efficient nutrient management in hydroponic systems.

1.4.3. Chapter 4: Ion-based dosing

Hypothesis:

In hydroponic systems, nutrient imbalances arise from differential ion uptake by plants; however, optimizing nutrient tank size and employing advanced dosing strategies, such as model predictive control (MPC), can enhance water and fertilizer use efficiency while minimizing the risks of nutrient deficiency and toxicity through precise regulation of individual macronutrient concentrations.

Objective:

Building upon the foundation established in Chapter 3, Chapter 4 introduces a more granular, nutrient-specific approach to managing hydroponic solutions in pursuit of fully closed-loop circulation. This chapter proposes a novel perspective that enables the formulation of customized nutrient solutions tailored to the specific requirements of various crops. Advanced control algorithms are employed to precisely regulate water volume and selected fertilizers' dosing,

ensuring that nutrient concentrations in the hydroponic solution remain within target ranges. This advanced methodology enhances the accuracy, efficiency, and adaptability of nutrient management in hydroponic systems.

Tables 1 and 2 present a comprehensive summary of the primary outcomes of this dissertation and the subsequent research directions it has catalyzed. Table 1 outlines the three core objectives addressed across Chapters 2 to 4, while Table 2 highlights the derivative studies and peer-reviewed publications that have emerged based on the knowledge from this work.

Table 1. Dissertation chapters and related publications.

Chapter	**Publication**
2	**Karimzadeh, S.**, Daccache, A., Rulli, M. C., & Ahamed, M. S. (2025). Global water-nutrient-salinity-energy nexus in lettuce production: from open-field irrigation to closed-loop hydroponics in greenhouses. *Journal of Agriculture and Food Research*, 101935.
3	**Karimzadeh, S.**, Li, Z., & Ahamed, M. S. (2025). Machine learning-based fault detection and diagnosis of electrical conductivity and pH sensors in hydroponic systems. *Computers and Electronics in Agriculture, 237*, 110544.
4	**Karimzadeh, S.**, McAllister, R. D., & Ahamed, M. S. (2025). A Model Predictive Control Framework for Precise Water and Fertilizer Dosing in Closed-Loop Hydroponics. *Computers and Electronics in Agriculture* (ready to submit).

Table 2. Other publications resulting from this dissertation.

Karimzadeh, S., Abou Najm, M., Ahamed, M. S., Bou-Zeid, E., Camporese, M., Daccache, A., Hernandez, R., Katul, G. (2025). "Global Assessment of Agrivoltaics Systems Potential for Major Cropping Systems" *One Earth* (under review).
Li, Z., **Karimzadeh, S.**, Chavanapanit, A., Moghimi, A., & Ahmed, M. S (2025). Detection of Calcium Deficiency in Indoor-Grown Lettuce Under Led Lighting Using Computer Vision. *Smart Agricultural Technology*, Available at SSRN 5239218 (under review).
Ahamed M. S., Chowdhury M., Inam, A. S, Aindrila K. K., Islam M. N., **Karimzadeh, S.**, Tabassum, S., Kabir, S., Akter, N., Momin, A. "A Comprehensive Review of Advances in Sensing and Monitoring Technologies for Precision Hydroponic Cultivation", *Computers and Electronics in Agriculture, 110601*

1.6. References

Al-Chalabi, M. (2015). Vertical farming: Skyscraper sustainability? *Sustainable Cities and Society*, *18*, 74–77. https://doi.org/10.1016/j.scs.2015.06.003

Bacci, L., Battista, P., Rapi, B., Pardossi, A., Incrocci, L., & Carmassi, G. (2005). A System for Fertigation Management in Closed-Loop Soilless Culture of Tomato. *Acta Horticulturae*, *674*, 263–268. https://doi.org/10.17660/ActaHortic.2005.674.30

Bamsey, M., Graham, T., Thompson, C., Berinstain, A., Scott, A., & Dixon, M. (2012). Ion-Specific Nutrient Management in Closed Systems: The Necessity for Ion-Selective Sensors in Terrestrial and Space-Based Agriculture and Water Management Systems. *Sensors*, *12*(10), 13349–13392. https://doi.org/10.3390/s121013349

Barbosa, G., Gadelha, F., Kublik, N., Proctor, A., Reichelm, L., Weissinger, E., Wohlleb, G., & Halden, R. (2015). Comparison of Land, Water, and Energy Requirements of Lettuce Grown Using Hydroponic vs. Conventional Agricultural Methods. *International Journal of Environmental Research and Public Health*, *12*(6), 6879–6891. https://doi.org/10.3390/ijerph120606879

Benke, K., & Tomkins, B. (2017). Future food-production systems: Vertical farming and controlled-environment agriculture. *Sustainability: Science, Practice and Policy*, *13*(1), 13–26. https://doi.org/10.1080/15487733.2017.1394054

Bothare, V. (2025). *Indoor farming market size, share & trends analysis report by growing system, facility type, crop type, and region (2025 – 2033)* [Market report]. Straits Research. Retrieved June 9, 2025, from https://straitsresearch.com/report/indoor-farming-market

Bouchaaba, Z., Santamaria, P., Choukr-Allah, R., Lamaddalena, N., & Montesano, F. F. (2015). Open-cycle drip vs closed-cycle subirrigation: Effects on growth and yield of greenhouse soilless green bean. *Scientia Horticulturae*, *182*, 77–85. https://doi.org/10.1016/j.scienta.2014.11.007

California Farm Water Coalition. (2024, April 4). *Water scarcity's economic toll on California agriculture: Why we must act now*. Farm Water. https://farmwater.org/farm-water-news/water-scarcitys-economic-toll-on-california-agriculture-why-we-must-act-now/

Casey, L., Freeman, B., Francis, K., Brychkova, G., McKeown, P., Spillane, C., Bezrukov, A., Zaworotko, M., & Styles, D. (2022). Comparative environmental footprints of lettuce supplied by hydroponic controlled-environment agriculture and field-based supply chains. *Journal of Cleaner Production*, *369*, 133214. https://doi.org/10.1016/j.jclepro.2022.133214

Cho, W. J., Kim, H.-J., Jung, D. H., Kang, C. I., Choi, G.-L., & Son, J.-E. (2017). An Embedded System for Automated Hydroponic Nutrient Solution Management. *Transactions of the ASABE*, *60*(4), 1083–1096. https://doi.org/10.13031/trans.12163

Cho, W.-J., Kim, H.-J., Jung, D.-H., Han, H.-J., & Cho, Y.-Y. (2019). Hybrid Signal-Processing Method Based on Neural Network for Prediction of NO3, K, Ca, and Mg Ions in Hydroponic Solutions Using an Array of Ion-Selective Electrodes. *Sensors*, *19*(24), 5508. https://doi.org/10.3390/s19245508

Cohen, A. R., Chen, G., Berger, E. M., Warrier, S., Lan, G., Grubert, E., Dellaert, F., & Chen, Y. (2022). Dynamically Controlled Environment Agriculture: Integrating Machine Learning and Mechanistic and Physiological Models for Sustainable Food Cultivation. *ACS ES&T Engineering*, *2*(1), 3–19. https://doi.org/10.1021/acsestengg.1c00269

Cowan, N., Ferrier, L., Spears, B., Drewer, J., Reay, D., & Skiba, U. (2022). CEA Systems: The Means to Achieve Future Food Security and Environmental Sustainability? *Frontiers in Sustainable Food Systems*, *6*, 891256. https://doi.org/10.3389/fsufs.2022.891256

Dsouza, A., Newman, L., Graham, T., & Fraser, E. D. G. (2023). Exploring the landscape of controlled environment agriculture research: A systematic scoping review of trends and topics. *Agricultural Systems*, *209*, 103673. https://doi.org/10.1016/j.agsy.2023.103673

El-Gafy, I., Grigg, N., & Waskom, R. (2017). Water-food-energy: nexus and non-nexus approaches for optimal cropping pattern. *Water Resources Management, 31*, 4971-4980.

Fiorani, F., & Schurr, U. (2013). Future Scenarios for Plant Phenotyping. *Annual Review of Plant Biology*, *64*(1), 267–291. https://doi.org/10.1146/annurev-arplant-050312-120137

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., ... Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342. https://doi.org/10.1038/nature10452

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., & Wilmoth, J. (2014). World population stabilization unlikely this century. *Science*, *346*(6206), 234–237. https://doi.org/10.1126/science.1257469

Goddek, S., Joyce, A., Kotzen, B., & Burnell, G. M. (Eds.). (2019). *Aquaponics Food Production Systems: Combined Aquaculture and Hydroponic Production Technologies for the Future*. Springer International Publishing. https://doi.org/10.1007/978-3-030-15943-6

Graamans, L., Baeza, E., Van Den Dobbelsteen, A., Tsafaras, I., & Stanghellini, C. (2018). Plant factories versus greenhouses: Comparison of resource use efficiency. *Agricultural Systems*, *160*, 31–43. https://doi.org/10.1016/j.agsy.2017.11.003

Griffiths, M., & York, L. M. (2020). Targeting Root Ion Uptake Kinetics to Increase Plant Productivity and Nutrient Use Efficiency. *Plant Physiology*, *182*(4), 1854–1868. https://doi.org/10.1104/pp.19.01496

Gruda, N. (2019). Increasing Sustainability of Growing Media Constituents and Stand-Alone Substrates in Soilless Culture Systems. *Agronomy*, *9*(6), 298. https://doi.org/10.3390/agronomy9060298

Hendrickson, C. T., Lave, L. B., & Matthews, H. S. (2010). *Environmental Life Cycle Assessment of Goods and Services* (0 ed.). Routledge. https://doi.org/10.4324/9781936331383

Heuvelink, E., Hemming, S., & Marcelis, L. F. M. (2024). Some recent developments in controlled-environment agriculture: On plant physiology, sustainability, and autonomous control. *The Journal of Horticultural Science and Biotechnology*, 1–11. https://doi.org/10.1080/14620316.2024.2440592

Higashide, T., & Heuvelink, E. (2009). Physiological and Morphological Changes Over the Past 50 Years in Yield Components in Tomato. *Journal of the American Society for Horticultural Science*, *134*(4), 460–465. https://doi.org/10.21273/JASHS.134.4.460

Hopmans, J. W., Qureshi, A. S., Kisekka, I., Munns, R., Grattan, S. R., Rengasamy, P., Ben-Gal, A., Assouline, S., Javaux, M., Minhas, P. S., Raats, P. A. C., Skaggs, T. H., Wang, G., De Jong Van Lier, Q., Jiao, H., Lavado, R. S., Lazarovitch, N., Li, B., & Taleisnik, E. (2021). Critical knowledge gaps and research priorities in global soil salinity. In *Advances in Agronomy* (Vol. 169, pp. 1–191). Elsevier. https://doi.org/10.1016/bs.agron.2021.03.001

Iddio, E., Wang, L., Thomas, Y., McMorrow, G., & Denzer, A. (2020). Energy efficient operation and modeling for greenhouses: A literature review. *Renewable and Sustainable Energy Reviews*, *117*, 109480. https://doi.org/10.1016/j.rser.2019.109480

Irabien, A., & Darton, R. C. (2016). Energy–water–food nexus in the Spanish greenhouse tomato production. *Clean Technologies and Environmental Policy, 18,* 1307-1316.

Karimzadeh, S., Hartman, S., Chiarelli, D. D., Rulli, M. C., & D'Odorico, P. (2024). The tradeoff between water savings and salinization prevention in dryland irrigation. *Advances in Water Resources*, *183*, 104604. https://doi.org/10.1016/j.advwatres.2023.104604

Kennedy, C., Cuddihy, J., & Engel-Yan, J. (2007). The Changing Metabolism of Cities. *Journal of Industrial Ecology*, *0*(0), 070322093406001. https://doi.org/10.1162/jiec.0.1107

Khoshnevisan, B., Shariati, H. M., Rafiee, S., & Mousazadeh, H. (2014). Comparison of energy consumption and GHG emissions of open field and greenhouse strawberry production. *Renewable and Sustainable Energy Reviews, 29*, 316-324.

Kim, J., Kim, H.-J., Gang, M.-S., Kim, D.-W., Cho, W.-J., & Jang, J. K. (2023). Closed Hydroponic Nutrient Solution Management Using Multiple Water Sources. *Journal of Biosystems Engineering*, *48*(2), 215–224. https://doi.org/10.1007/s42853-023-00182-0

Lee, J. Y., Rahman, A., Azam, H., Kim, H. S., & Kwon, M. J. (2017). Characterizing nutrient uptake kinetics for efficient crop production during Solanum lycopersicum var. Cerasiforme Alef. Growth in a closed indoor hydroponic system. *PLOS ONE*, *12*(5), e0177041. https://doi.org/10.1371/journal.pone.0177041

Malone, T. C., & Newton, A. (2020). The Globalization of Cultural Eutrophication in the Coastal Ocean: Causes and Consequences. *Frontiers in Marine Science*, *7*, 670. https://doi.org/10.3389/fmars.2020.00670

Marcelis, L. F. M., Broekhuijsen, A. G. M., Meinen, E., Nijs, E. M. F. M., & Raaphorst, M. G. M. (2006). Quantification of the Growth Response to Light Quantity of Greenhouse Grown Crops. *Acta Horticulturae*, *711*, 97–104. https://doi.org/10.17660/ActaHortic.2006.711.9

Martin, M., Elnour, M., & Siñol, A. C. (2023). Environmental life cycle assessment of a large-scale commercial vertical farm. *Sustainable Production and Consumption*, *40*, 182–193. https://doi.org/10.1016/j.spc.2023.06.020

Massa, D., Incrocci, L., Maggini, R., Carmassi, G., Campiotti, C. A., & Pardossi, A. (2010). Strategies to decrease water drainage and nitrate emission from soilless cultures of greenhouse tomato. *Agricultural Water Management*, *97*(7), 971–980. https://doi.org/10.1016/j.agwat.2010.01.029

Mattson, N. S., & Van Iersel, M. W. (2011). Application of the "4R" Nutrient Stewardship Concept to Horticultural Crops: Applying Nutrients at the "Right Time." *HortTechnology*, *21*(6), 667–673. https://doi.org/10.21273/HORTTECH.21.6.667

Pennisi, G., Orsini, F., Blasioli, S., Cellini, A., Crepaldi, A., Braschi, I., Spinelli, F., Nicola, S., Fernandez, J. A., Stanghellini, C., Gianquinto, G., & Marcelis, L. F. M. (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), 14127. https://doi.org/10.1038/s41598-019-50783-z

Rosa, L., & D'Odorico, P. (2019). The water-energy-food nexus of unconventional oil and gas extraction in the Vaca Muerta Play, Argentina. *Journal of Cleaner Production*, *207*, 743–750. https://doi.org/10.1016/j.jclepro.2018.10.039

Rufí-Salís, M., Petit-Boix, A., Villalba, G., Sanjuan-Delmás, D., Parada, F., Ercilla-Montserrat, M., Arcas-Pilz, V., Muñoz-Liesa, J., Rieradevall, J., & Gabarrell, X. (2020). Recirculating water and nutrients in urban agriculture: An opportunity towards environmental sustainability and water

use efficiency? *Journal of Cleaner Production*, *261*, 121213. https://doi.org/10.1016/j.jclepro.2020.121213

Sandison, F., Yeluripati, J., & Stewart, D. (2023). Does green vertical farming offer a sustainable alternative to conventional methods of production?: A case study from Scotland. *Food and Energy Security*, *12*(2), e438. https://doi.org/10.1002/fes3.438

Savvas, D., Giannothanasis, E., Ntanasi, T., Karavidas, I., Drakatos, S., Panagiotakis, I., Neocleous, D., & Ntatsi, G. (2023). Improvement and validation of a decision support system to maintain optimal nutrient levels in crops grown in closed-loop soilless systems. *Agricultural Water Management*, *285*, 108373. https://doi.org/10.1016/j.agwat.2023.108373

Schmitt, R. J. P., Rosa, L., & Daily, G. C. (2022). Global expansion of sustainable irrigation limited by water storage. *Proceedings of the National Academy of Sciences*, *119*(47), e2214291119. https://doi.org/10.1073/pnas.2214291119

SharathKumar, M., Heuvelink, E., & Marcelis, L. F. M. (2020). Vertical Farming: Moving from Genetic to Environmental Modification. *Trends in Plant Science*, *25*(8), 724–727. https://doi.org/10.1016/j.tplants.2020.05.012

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation – a global inventory. *Hydrology and Earth System Sciences*, *14*(10), 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

The State of the World's Land and Water Resources for Food and Agriculture – Systems at breaking point (SOLAW 2021). (2021). FAO. https://doi.org/10.4060/cb7654en

Tomasi, N., Pinton, R., Dalla Costa, L., Cortella, G., Terzano, R., Mimmo, T., Scampicchio, M., & Cesco, S. (2015). New 'solutions' for floating cultivation system of ready-to-eat salad: A review. *Trends in Food Science & Technology*, *46*(2), 267–276. https://doi.org/10.1016/j.tifs.2015.08.004

United Nations. (2014). *World urbanization prospects: The 2014 revision - highlights*. UN. https://doi.org/10.18356/527e5125-en

Wade, C. M., Austin, K. G., Cajka, J., Lapidus, D., Everett, K. H., Galperin, D., Maynard, R., & Sobel, A. (2020). What Is Threatening Forests in Protected Areas? A Global Assessment of Deforestation in Protected Areas, 2001–2018. *Forests*, *11*(5), 539. https://doi.org/10.3390/f11050539

Wild, A. (2003). *Wild, A. (2003). Soils, land and food: Managing the land during the twenty-first century*. Cambridge University Press.

Zarei, M. J., Kazemi, N., & Marzban, A. (2019). Life cycle environmental impacts of cucumber and tomato production in open-field and greenhouse. *Journal of the Saudi Society of Agricultural Sciences*, *18*(3), 249–255. https://doi.org/10.1016/j.jssas.2017.07.001

Zhou, D., Meinke, H., Wilson, M., Marcelis, L. F. M., & Heuvelink, E. (2021). Towards delivering on the sustainable development goals in greenhouse production systems. *Resources, Conservation and Recycling*, *169*, 105379. https://doi.org/10.1016/j.resconrec.2020.105379

Chapter 2: Global water-nutrient-salinity-energy nexus in lettuce production: from open-field irrigation to closed-loop hydroponics in greenhouses

Reference: Karimzadeh, S., Daccache, A., Rulli, M. C., & Ahamed, M. S. (2025). Global water-nutrient-salinity-energy nexus in lettuce production: from open-field irrigation to closed-loop hydroponics in greenhouses. *Journal of Agriculture and Food Research*, 101935.

2.1. Abstract

Hydroponic greenhouse production is a proven approach to enhancing resource efficiency and overcoming limitations related to land availability, soil degradation, and climate variability. However, assessing the benefits and challenges of transitioning to greenhouses remains underexplored. This study provides a global assessment of the water-nutrient-salinity-energy nexus in lettuce production, spanning from soil-based open-field systems to closed-loop hydroponic greenhouse technologies. The study examines soil-water interactions in conventional open-field agriculture and highlights critical challenges, including soil salinization and nutrient leaching inherent to open-field farming. While transitioning to closed-loop hydroponics in greenhouses offers a potential solution to these issues, it comes with a considerable trade-off in the form of elevated energy demands for greenhouse operations. This study quantifies both the water demand for hydroponics and the energy required for heating and cooling, while also evaluating the contribution of harvested rooftop rainfall. Adopting closed-loop hydroponic systems in greenhouses can reduce water consumption for lettuce production by 6% to over 90%, while eliminating risks of soil salinization and groundwater contamination. Harvested rooftop rainfall can meet 100% of the hydroponics water demand in some locations, such as Bangladesh,

during peak monsoon months (June–August). However, globally, the total energy demands for heating and cooling range from 39 to 123 MJ/kg, depending on climate. Furthermore, the minimum annual energy requirement for greenhouse operation was found to be 1854 MJ/m2 in Cwc climates (temperate, dry winters, and cold summers), with energy distribution favoring heating (62%) over cooling (38%). In India, specific water use (evapotranspiration and water for evaporative cooling) decreased by 93%, from 1146 L/kg in open-field furrow systems to 85 L/kg in closed-loop hydroponics, with an energy demand of 66 MJ/kg. These findings offer comprehensive, global-scale insights across 30 Köppen–Geiger climate zones and the 15 leading lettuce-producing countries, providing a foundation for policymakers to plan future agricultural development.

Keywords: Agricultural water management; Energy demand; Global scale nexus; Greenhouses; Nutrient leaching; Open-field agriculture

2.2. Introduction

Conventional open-field farming faces challenges in producing food sustainably and ensuring fair distribution to meet the needs of a growing global population. Currently, 3.3 billion people depend on unsustainable agricultural practices for their food supply (Schmitt et al., 2022). With the global population expected to reach 10.2 billion by 2050 (Kummu et al., 2017), there is an urgent need for more sustainable food production methods. Conventional agricultural methods not only deplete water resources and compromise water quality (Elliott et al., 2014; Postel et al., 1996) but also contribute to one-fourth of the world's greenhouse gas emissions (Crippa et al., 2021; Rosa et al., 2021; Rosa & Gabrielli, 2023).

The cultivated areas devoted to lettuce (*Lactuca sativa*) cultivation have witnessed a remarkable surge, registering a substantial 180% increase over the past six decades (FAOSTAT, 2023, Figure

24b). China, the United States, India, Spain, and Italy are the main lettuce producers. Nevertheless, the trajectory of lettuce yield has undergone a linear ascent, reaching a plateau commencing in 1992 (Figure 24b). Notably, despite an intensification in applied fertilizers—especially nitrogen, phosphorus, and potassium—no substantial increase in yield has been observed (Figure 24c). In recent decades, there has been a surge in interest in controlled environment agriculture (CEA) in the forms of greenhouses and vertical farms, which is characterized by precise control over essential factors such as light, temperature, humidity, water, and nutrients (Goodman & Minner, 2019). Meanwhile, soil-based greenhouse cultivation systems are often associated with substantial nitrate leaching losses (Incrocci et al., 2020). Excessive or inefficient irrigation practices are a primary driver of this phenomenon, exacerbating nutrient displacement through deep percolation beyond the root zone and contributing to the accumulation of salts within the rhizosphere, thereby intensifying soil salinization. Alternatively, hydroponic systems (i.e., soilless plant growth using recirculated nutrient-rich water) have the potential to address water stress and soil degradation in arid and semi-arid regions dedicated to horticulture, as they require 70-95% less direct water and fertilizer compared to open-field soil-based agriculture (Araya et al., 2024). Barbosa et al. (2015) observed a water consumption of 20 L/kg/yr in lettuce hydroponic cultivation, significantly lower than the 250 L/kg/yr in conventional soil-based production (more than 90% reduction in water consumption). Similarly, Majid et al. (2021) reported a 64% reduction in water consumption of lettuce with nutrient film technique (NFT) systems compared to soil-based cultivation. The closed-loop system demonstrates greater economic efficiency than open-concept hydroponics, particularly in fertilizer savings (85% vs. 68%). While water savings are slightly higher in the closed-loop system (90% vs. 85%)(Al Shrouf, 2017; Castillo et al., 2014). Additional factors, such

as long-term sustainability and resource recovery, further highlight the benefits of closed-loop hydroponics.

Despite the irrigation water and nutrient-saving benefits of hydroponic systems in greenhouses, it is imperative to acknowledge that their high energy consumption may contribute to higher greenhouse gas emissions (Fathidarehnijeh et al., 2023). Ahamed et al. (2019) reported the average annual heating demand for greenhouse vegetable production in Saskatchewan, Canada, is around 1,632 MJ/m^2. However, the annual heating load could range between 1,000 MJ/m^2 to 3,350 MJ/m^2 in heating-dominated regions in North America based on greenhouse types, temperature set points, and location (Lavigueur, 2022). The cooling cost varies depending on active (50-300 MJ/m^2) and passive cooling (2-10 MJ/m^2), and most greenhouses are cooled passively with ventilation or with active evaporative cooling (Choab et al., 2021; Soussi et al., 2022). The annual energy demand of greenhouse production is 220-320 MJ/m^2 in Southern Europe (including Italy, Southern France, and Greece), while this value is up to 3,600 MJ/m^2 in North European countries (Poland, Germany, and the Netherlands; Dimitropoulou et al., 2023). Also, Barbosa et al. (2015) reported an average annual energy demand of 3,690 MJ/m^2 for lettuce production in greenhouses in Arizona, the United States, with values ranging from 2,757 to 4,757 MJ/m^2. Therefore, the sustainability of the greenhouse production method hinges on proper implementation, given the environmental hotspots associated with its intense heating, cooling, and lighting energy requirements. Nevertheless, with careful design and scaling, deploying greenhouses could be vital in transforming food systems towards sustainability, especially in regions with extreme climates (Casey et al., 2022).

Evaporative cooling systems are widely implemented as heat sinks for greenhouse climate control (Alnaimat, 2022). These systems utilize mist/fog, cooling pads, or sprinklers to convert sensible

heat into latent heat via water evaporation (Çaylı et al., 2021). Fan and pad evaporative cooling systems are widely used for greenhouse cooling in hot climates and remain effective even in humid subtropical regions, as demonstrated in southeastern China (Xu et al., 2015). They use exhaust fans, evaporative pads, and circulation (Tawalbeh et al., 2023). However, evaporative cooling in greenhouses requires substantial water inputs, particularly in arid regions. Fuchs et al. (2006) demonstrated that in Israeli rose greenhouses, water consumption for evaporative cooling often surpasses irrigation demand when ambient relative humidity falls below 50%. Similarly, Sabeh et al. (2011) reported that in Arizona, USA, water use for evaporative cooling (62 L/kg of tomatoes) was four times higher than irrigation (15 L/kg). In extreme cases, such as in Saudi Arabia's arid climate, evaporative cooling constitutes the dominant water use in greenhouse operations (Tsafaras et al., 2021).

Several studies evaluated the resources used for greenhouses in various regions. Despite being helpful, most of them performed the analysis on a local and city level (Barbosa et al., 2015; Graamans et al., 2017, 2018; Soufi et al., 2023; Van Ginkel et al., 2017; Weidner et al., 2021, 2022). Paucek et al. (2023) conducted a country-level feasibility analysis of African agriculture and explored the potential of indoor vertical farming, identifying countries such as South Africa, Seychelles, and Egypt as promising candidates for sustainable implementation based on a thorough analysis of development indicators. Meanwhile, the sophisticated interdependencies and competitive dynamics between water, energy, land, and food have acquired significant attention in recent years, typically known as the 'nexus' framework (Cristiano et al., 2021; D'Odorico et al., 2018). Within the CEA context, Weidner et al. (2022) employed a nexus approach, interconnecting food, energy, and land, delving into evaluating greenhouses and vertical farms. However, their

analysis was constrained to nine cities spanning Iceland, Sweden, the United States, Japan, Chile, South Africa, Singapore, and the United Arab Emirates.

A critical research gap exists regarding the need for a global-scale overview of the geospatial variability of suitable climates and country-specific conditions, which would underscore the benefits and costs of transitioning from open-field to greenhouse-based agricultural systems. Here, we present a novel application of the water–nutrient–salinity–energy nexus framework to evaluate the opportunities and challenges associated with both open-field and greenhouse lettuce production systems at a global scale. This approach provides critical insights into the potential trade-offs and benefits of such a transition, informing more sustainable and resource-efficient production strategies.

We evaluate water consumption in lettuce production, comparing open-field systems—comprising furrow irrigation (low-cost but inefficient) and drip irrigation (high efficiency)—with closed-loop hydroponics in high-tech greenhouse systems. Meanwhile, water quality plays an important role in shaping the challenges faced by open-field cultivation, including soil salinity and nutrient leaching. By examining water quantity and quality in open fields, we highlight the potential of hydroponic greenhouses to mitigate these limitations. Furthermore, we analyze the variability in energy demands for greenhouse operations across climatic zones and countries.

This study compares traditional soil-based open-field lettuce production with advanced climate-controlled, closed-loop hydroponic greenhouse systems incorporating supplementary lighting. The analysis focuses on the interrelated factors of water, nutrients, salinity, and energy—referred to as the water-nutrient-salinity-energy nexus.

The first part of the study examines soil-water dynamics in open-field agriculture, driven by both the quantity and quality of irrigation. Soil salt and water balance models are applied to:

(1) Quantify blue water (irrigation sourced from surface or groundwater) and green water (precipitation stored in the soil and absorbed by plants); and

(2) Evaluate soil salinity and nutrient leaching under different irrigation methods—furrow and drip—across four levels of irrigation water salinity: 0.5, 1.0, 1.5, and 2.0 dS/m.

The second part focuses on hydroponic greenhouse systems by:

(3) Assessing the spatial and temporal sustainability of water management through rainwater harvesting from greenhouse roofs for use in a closed-loop hydroponic system; and

(4) Quantifying the energy (heating and cooling) and water requirements (for both irrigation and the fan-and-pad cooling system).

By providing a global-scale assessment of these interconnected factors, the study offers strategic insights for governments, policymakers, companies, and investors considering the transition from open-field to greenhouse lettuce production..

2.3. Materials and methods

2.3.1. Lettuce production on a global scale

Although wheat, maize, rice, and soybean dominate 50% of global harvested areas (Grogan et al., 2022), lettuce was selected due to its current potential for indoor hydroponic cultivation and the limited economic viability of growing staple crops in soilless CEA production systems (Asseng et al., 2020).

2.3.2. Open-field soil-based production

2.3.2.1. Crop water requirement in open field

The crop water requirement (CWR) was calculated using the FAO56 crop coefficient method (Allen et al., 1998). Reference evapotranspiration (ET_o) was calculated with the Penman-Monteith equation standardized for short (12 cm), disease-free and well-irrigated grass (Eq. 2.1). The potential crop evapotranspiration (ET_c) was calculated by multiplying ET_o by the crop coefficient (K_c). K_c values varied based on the growth stages, which are in the supplementary material (Table 11).

$$ET_o = \frac{0.408\,\Delta\,(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \qquad (2.1)$$

where R_n is the net radiation (MJ/m^2/day), and G is the soil heat flux density (MJ/m^2/day), assumed to be zero at the daily timescale (Ahmadi et al., 2022; Allen et al., 1998). T is the mean daily air temperature (°C); u_2 is the wind speed at 2 m height (m/s); e_s is the saturation vapor pressure (i.e., 100% relative humidity, in kPa); e_a is the actual vapor pressure (kPa); Δ is the slope of the saturation vapor pressure curve (kPa/°C) at T (alterations in the e_s that occur in response to temperature and humidity variations); and γ is the psychrometric constant (kPa/°C, which relates the partial pressure of water in the air to the air temperature).

2.3.2.2. Soil water, salinity, and nutrient balances

Soil moisture dynamics within the root zone were determined through a daily water balance (Eq. 2.2) per each grid cell (Karimzadeh et al., 2024):

$$\frac{d}{dt}(\theta . Z_r) = P + I - ET_c - DP \qquad (2.2)$$

where θ is the volume-based soil moisture, Z_r is the depth of roots (mm), P is precipitation (mm/day), I is irrigation (mm/day), and DP is the rate of deep percolation (mm/day). Capillary rise (upward movement of water from shallow groundwater) was considered negligible in this study. For modeling water and soil salinity in the root zone, nutrients within this zone are assumed to mix uniformly and instantly (i.e., complete mix reactor). The nutrient balance is expressed as Eq. 2.3, as proposed by Karimzadeh et al. (2024):

$$\frac{d}{dt}(\theta . Z_r . C_{N,r}) = C_{N,I} . I - C_{N,r} . DP \quad (2.3)$$

where $C_{N,r}$ denotes the total dissolved solids (TDS) of nutrients present in the root zone water volume (mg/L), and $C_{N,I}$ is the nutrient concentration of the applied water via irrigation (mg/L). The $C_{N,r} . DP$ quantifies the mass of nutrients leaching into groundwater (mg). To maintain uniformity in nutrient measurements in both soil and water sources, TDS is converted to soil electrical conductivity (EC) of the saturated paste based on Eq. 2.4 given by Rhoades (2018):

$$TDS = k_e \times EC \quad (2.4)$$

The factor k_e can vary between 550 to 800 (Atekwana et al., 2004). The values of 640 and 800 were employed for EC below and above 5 dS/m, respectively (UC, 2023).

2.3.2.3. Lettuce salt tolerance

The impact of soil salinity on the ET_c rate can be assessed using the salinity stress coefficient (K_s), which accounted for the EC of the soil solution. In the absence of water stress, the negative effect of salinity on plant growth was quantified using the yield response factor in Eq. 2.5 (Allen et al., 1998):

$$K_s = 1 - \frac{b}{K_y \times 100}\left(EC_e - EC_{e,thr}\right) \quad (2.5)$$

EC_e is the electrical conductivity of the saturated extract and $EC_{e,thr}$ at the threshold level for lettuce tolerance (measured in dS/m). b is the yield reduction per percentage increase in EC_e (expressed as %/(dS/m)), and K_y is a factor indicating the response of yield to salinity stress (Table 11). Subsequently, the average K_s during the growing cycle is used to categorize salinity severity. The classification is based on K_s values of 1, 1-0.9, 0.9-0.7, 0.7-0.5, and below 0.5, denoting non-saline, slight saline, moderate saline, strong saline, and extremely saline conditions, respectively (Karimzadeh et al., 2024).

2.3.2.4. Irrigation practices

Irrigation was simulated under furrow and drip systems. Despite its lower efficiency (~60%), furrow irrigation remains widely used in many countries. Furrow irrigation was considered when soil moisture declined to 70% of the total available water (TAW), corresponding to the depletion fraction (p) threshold. This threshold maintained non-stress conditions for lettuce, as further depletion would impair transpiration due to limited water availability (Allen et al., 1998; Fu et al., 2024). The applied water replenished soil moisture close to field capacity, the maximum water-holding capacity of the soil after gravitational drainage.

Drip irrigation was considered a more efficient alternative to furrow irrigation, with an efficiency of around 90% (Brouwer et al., 1992; Yang et al., 2023). In this system, irrigation was considered daily to restore soil moisture to field capacity, except when precipitation exceeded crop evapotranspiration. Four water salinity levels 0.5, 1.0, 1.5, and 2.0 dS/m—were evaluated, representing a range from fresh surface water irrigation (0.5 dS/m) to moderately saline groundwater (2.0 dS/m). Table 12 summarizes the irrigation strategies investigated in this study.

2.3.2.5. Irrigation salt-score: Integrated water quality and quantity metric

A comprehensive assessment of irrigation impacts requires consideration of factors such as irrigation volume, water quality, and their interactions with plant-soil-atmosphere dynamics. To support water management decisions, a novel metric, the irrigation salt-score (ISS), was introduced. Unlike conventional salinity indices that primarily assess soil salinity or irrigation water quality in isolation, the ISS explicitly differentiated between blue water (irrigation) and green water (precipitation) to account for both anthropogenic and natural influences on soil salinity. By incorporating the irrigation water source (e.g., groundwater, surface water) and its role in salinity dynamics, ISS captured human-induced contributions to soil salinization. Simultaneously, it recognized green water as a natural salinity regulator, delivering high-quality (low salinity) water that facilitated salt leaching and contributed to the long-term sustainability of agricultural soils. This dual consideration positioned ISS as a robust framework for informing water management policies. ISS provided a comprehensive assessment of countries' susceptibility to secondary salinization by integrating precipitation patterns and climatic conditions, offering critical insights for sustainable land and water resource management. Here, ISS integrated the irrigation technique (e.g., drip or furrow), GW, BW, and crop-specific salt-affected areas within a country. The formulation of ISS is presented in Eq. 2.6:

$$\mathrm{ISS} = \left(\frac{\text{Blue Water}}{\text{Green Water} + \text{Blue Water}} \times \frac{\text{Salt affected areas}}{\text{Total harvested areas}}\right) \times 100 \qquad (2.6)$$

where salt-affected areas indicate the sum of areas impacted by all levels of salinity (slight to extreme, as defined in Section 2.3.2.3). The ISS ranged from 0 to 100, with higher values indicating a greater susceptibility to secondary salinization (human-induced salt buildup) resulting from improper irrigation quality or strategy as described in Section 2.3.2.4 and summarized in Table 12.

2.3.3. Closed-loop hydroponics in greenhouse production

To transition current open-field lettuce cultivation areas, we assumed that the designed greenhouse structure in Section 2.3.3.1 would be replicated in these regions. Afterward, we analyzed global-scale resource consumption in advanced, climate-controlled, closed-loop hydroponic greenhouse systems with supplementary lighting, which were designed to optimize plant growth under diverse climatic conditions. Salinity levels in hydroponic systems were assumed to continuously monitored by EC sensors, ensuring precise control of the nutrient solution composition. As a result, salinity-related impacts and nutrient leaching were effectively mitigated, rendering these concerns negligible and therefore excluded from analysis.

2.3.3.1. Greenhouse design and attributes

In this study, we considered a conceptually designed high-tech closed concept greenhouse under varying climatic conditions, covering an area of ~831 m^2. The structure was consisted of three spans, each measuring 43.3 m in length and 6.4 m in width, with the floor designed to minimize perimeter heat loss through a thermal conductivity of 1.4 $W/m^2/K$. The ridge height was 5.3 m, while the gutter height was around 3.7 m. The roof was angled at 26° and covered with double-layer polyethylene films, providing a solar transmissivity of 0.75 (Ahamed et al., 2018).

Advanced environmental control systems maintained set-point temperatures of 28°C during the day and 20°C at night, while ensuring optimal humidity (75%) and air velocity (0.2 m/s). A network of recirculating fans, powered by high-efficiency motors (0.9), enhanced air circulation and reduced thermal stratification. The greenhouse design also accounted for the physiological needs of crops, incorporating a leaf area index (LAI) of 3.0 (Graamans et al., 2017). Additional design specifications followed those outlined by Ahamed et al. (2018a).

The study employed a closed-loop NFT hydroponic system, where a thin film of nutrient solution continuously flowed over plant roots. This system ensures efficient nutrient uptake and aeration while preventing nutrient leaching and soil salinity through precise nutrient management and solution recirculation.

2.3.3.2. Crop water requirement in greenhouses

In accordance with the methodology used for the open field (Section 2.3.2.1), crop evapotranspiration was estimated as the product of ET_o and K_c. The FAO56 Penman-Monteith equation is the benchmark for estimating ET_o in agricultural water management (Ahmadi et al., 2024). However, its direct application to greenhouse environments was problematic and resulted in ET_o underestimation (Incrocci et al., 2020; Libardi et al., 2019). The primary challenge lays in accurately quantifying net radiation within greenhouses, as 60-80% of ET was driven by radiation (Montero et al., 2001; Villarreal-Guerrero et al., 2012). Unlike open field, where net radiation was well-characterized (see supplementary Eqs. S2.2–S2.9 for detailed formulations) using the methodology of Allen et al. (1998), greenhouse environments introduced additional complexities. Multiple reflections and re-emissions of longwave radiation altered the thermal balance, influencing air temperature and subsequently affecting ET_o estimates. Given these limitations, the modified Hargreaves-Samani model (Fernández et al., 2010; Hargreaves & Samani, 1985) was employed to calculate ET_o. This temperature-based approach accounted for global variations in extraterrestrial solar radiation while incorporating greenhouse transmissivity characteristics, offering a more robust estimation:

$$ET_{o,g} = 0.0023\left(T_{mean,g} + 17.8\right)\left(T_{max,g} - T_{min,g}\right)^{0.5} R_a \,.\, \tau \quad (2.7)$$

where $T_{mean,g}$, $T_{max,g}$, and $T_{min,g}$ are the mean, maximum, and minimum air temperature (°C) inside the greenhouse, respectively. τ is the transmissivity of greenhouse (unitless). The Hargreaves–Samani equation relies solely on measured temperature data and requires extraterrestrial radiation (R_a), which was calculated based on latitude, date, and time (see supplementary Eqs. S2.6–S2.9; Allen et al., 1998). This approach is particularly well-suited for scenarios characterized by limited climate data availability while ensuring accuracy in estimating ET_o. To accommodate diverse lettuce varieties with varying growth durations, an average greenhouse crop coefficient ($K_{c,g}$) value of 1.61 was adapted for lettuce, following Zhangzhong et al. (2023). Besides, a leaching fraction of 30% was adopted to account for the utilization of saline water sources, such as groundwater, and to mitigate excessive salt accumulation resulting from irrigation (Carmassi et al., 2007; Incrocci et al., 2020; Magán et al., 2019).

2.3.3.3. Greenhouse rainwater harvesting potential

Rainwater harvesting entails collecting and storing rainwater from impervious surfaces, such as rooftops, conserving, storing, and distributing fresh water, demonstrated in conventional, greenhouse, and urban agricultural use (Londra et al., 2021; Stout et al., 2017). Sustainable water usage was assessed by subtracting the initial flush—representing runoff containing potential pollutants—from total precipitation events. A standardized first flush value of 1 mm was considered (Almeida et al., 2021; Silva et al., 2017), and the remaining water, referred as harvested blue water (HBW) (Hoekstra, 2019), was used to evaluate the sustainability of hydroponic lettuce water consumption in greenhouses. Sustainability was achieved when the monthly HBW exceeded the CWR in closed-loop NFT greenhouse production.

2.3.3.4. Heating and cooling energy demand

To model the energy requirements, the GREENHEAT model was adapted, originally designed for predicting heating needs in greenhouses with fixed temperature and relative humidity set points (Ahamed et al., 2018). The model was based on a lumped-system analysis for closed greenhouses, accounting for various heat sources and sinks within the system. The heating/cooling energy requirement were calculated using the energy balance of a closed greenhouse, as expressed in Eq. 2.8:

$$Q_{c,h} = (Q_s + Q_{sl} + Q_{CO_2} + Q_m) - (Q_t + Q_i + Q_g + Q_p + Q_r + Q_{ET}) \quad (2.8)$$

Where Q_s is the net heat gain from solar radiation. Q_{sl}, Q_{CO_2}, and Q_m are heat gain from supplemental lighting, CO_2 generators, and motors (i.e., recirculating fans), respectively. Q_t is the heat transfer by conduction and convection. Q_i is the heat transfer due to air exchange in infiltration. Q_g signifies the heat transfer through the greenhouse floor. Q_p denotes the heat transfer along the perimeter. Q_r is the heat loss through longwave radiation. Q_{ET} represents the heat used in the process of plant evapotranspiration. A negative value (denoted as $Q_{c,h}$) indicated the need for heating, while a positive value signified a requirement for sensible cooling. A detailed explanation of the sub-modeling principles governing the terms in Eq. 2.8 was provided in the supplementary material (Eqs. S2.17 to S2.26). The greenhouse was equipped with cooling systems to regulate daytime temperatures below the set point and heating systems to maintain nighttime temperatures above the set point. The heating and cooling set points for lettuce were 20°C and 28°C, respectively. While the optimal daytime temperature for lettuce production was 24°C (Carotti et al., 2021), a slight yield reduction at 28°C (Carotti et al., 2021) was considered an acceptable trade-off to reduce water consumption for cooling (see Section 2.3.3.5).

2.3.3.5. Water consumption for evaporative cooling

Evaporative cooling systems are widely implemented as heat sinks for greenhouse climate control (Alnaimat, 2022). The water required for greenhouse cooling was calculated by converting the sensible cooling load estimated by the GREENHEAT model (Section 2.3.3.4) into evaporative demand. This was achieved by dividing the cooling load by the latent heat of water (2.45 MJ/kg) and adjusting for a conservative fan and pad cooling efficiency of 77.5% (Tawalbeh et al., 2023). This analysis did not account for the potential benefits of free cooling when the outside temperature is below the set point. Also, the actual cooling loads under evaporated cooling ventilation could be slightly lower than the load estimated for closed greenhouse operations. This study excluded the complexity of modeling evaporatively cooled, ventilated greenhouses to enable a feasible global-scale analysis.

2.3.4. Specific water and energy use per yield comparison

Specific water use (SWU) was determined by the ratio of water consumption—encompassing hydroponic and cooling needs—to the reported average yield of 41 kg/m^2/yr (Barbosa et al., 2015). For open-field systems, SWU was calculated by dividing total water use, including GW and BW through furrow and drip irrigation methods, over the open-field yield global map (Monfreda et al., 2008), which revealed remarkable geographic variability, with yields ranging from less than 4 kg/m²/yr in Bangladesh to more than 40 kg/m²/yr in the United States.

Specific energy use (SEU) for greenhouse lettuce production was assessed by quantifying the cumulative active sensible energy inputs ($Q_{c,h}$) required to maintain greenhouse operations. However, this study did not consider the energy used for moisture control (latent cooling) through ventilation. In contrast, direct energy use in open-field production was considered negligible, given

its the reported narrow range of 0.7 to 1.1 MJ/kg (Barbosa et al., 2015), which accounted for less than 2% of the energy demand for climate control in greenhouses (see Results and discussion). The indirect energy consumption associated with lettuce production—including energy embedded in the production, transportation, and maintenance of inputs and infrastructure—was beyond the scope of this analysis.

2.3.5. Global-scale implementation

The analysis compared lettuce production in open-field and greenhouse systems, exploring the potential and rationale for switching between the two systems and identifying suitable regions. The water-nutrient-salinity-energy nexus was assessed through soil salinity and nutrient leaching in open fields, energy requirements for greenhouses, and water consumption of both. The study incorporated global spatial data on lettuce-cultivated areas (Monfreda et al., 2008) and a country- and climate-class analysis to provide actionable insights for sustainable production.

2.3.5.1. Datasets and spatial analysis

The analysis integrated a comprehensive dataset encompassing climate, soil, and crop-specific variables to quantify the interactions within the lettuce production system under varying environmental conditions. Key parameters included extraterrestrial radiation and daily air temperature variables (Allen et al., 2001; Copernicus Climate Change Service, 2019a). Soil characteristics and salinity metrics were sourced from Köchy et al. (2015) and Shangguan et al. (2014). Hourly meteorological data, including air temperature, wind components, cloud cover, relative humidity, and soil temperature were from the Copernicus Climate Change Service (2019b). Elevations were derived from JISAO (2024). Given the predominant influence of climate conditions, the Köppen–Geiger climate classification was incorporated to account for climatic variability. This classification delineates global climates into five principal categories—tropical

(A), arid (B), temperate (C), continental (D), and polar (E)—and further subdivides them into 30 distinct subclasses based on characteristic patterns of seasonal temperature and precipitation (Beck et al., 2018). Figure 25 shows climate class regions, while Table 13 summarizes lettuce harvest areas by climate class. The methodologies detailed in Sections 2.3.2 to 2.3.4 were systematically applied at the pixel level across the global scale. Inputs, comprising both 2D and 3D datasets, were sourced in NetCDF format with varying native spatial resolutions. These datasets were subsequently resampled to a uniform resolution of 0.1° (~11.1 km at the equator) using the nearest-neighbor interpolation method in MATLAB to ensure spatial consistency across analyses. Among all tested interpolation methods (nearest neighbor, bilinear, bicubic, lanczos-2, lanczos-3, box filter, and triangle), the nearest neighbor approach yielded the smallest error in cultivated area, limiting the decrease to 30%. The initial soil EC dataset was provided across eight discrete soil depth layers (Shangguan et al., 2014). To approximate the average EC within the topsoil most relevant to lettuce root development, the mean EC across the upper five layers was computed, representing approximately the top 50 cm of the soil profile—an active zone for root-soil interactions. The resulting global datasets were spatially masked to delineate areas under lettuce cultivation (Monfreda et al., 2008). A comprehensive inventory of the variables employed, along with their sources and attributes, is provided in Table 14, as referenced in Sections 2.3.1 to 2.3.4. Figure 1. summarizes the research question, methodology, and used datasets.

2.3.5.2. Climate variability and results robustness

Our primary analysis was initially based on data from the 2015–2016 period, which coincided with a strong El Niño event. This period was deliberately chosen to capture extreme climatic scenarios, a common practice in controlled environment agriculture where greenhouse systems are often designed to withstand worst-case environmental stressors. To assess the robustness of our results

under varying climatic conditions, an uncertainty analysis was conducted using average climate data from 2019 to 2023. This period was selected to represent a more typical range of interannual variability and to reflect expected climatic conditions. Eq. 2.9 was applied to quantify the relative percentage change in key results, including BW, GW, nutrient leaching, K_s, HBW, and precipitation, under varying climatic scenarios.

$$\text{Relative Difference} = \frac{\text{Results}_{5-\text{year average}} - \text{Results}_{\text{El Niño year}}}{\text{Results}_{\text{El Niño year}}} \times 100 \quad (2.9)$$

By complementing this with a multi-year average dataset from a climatologically moderate period, we were able to quantify the degree of uncertainty in system performance metrics and evaluate the sensitivity of key outcomes to year-to-year climate variability. This dual-scenario approach ensured both resilience in design and relevance to real-world operational conditions.

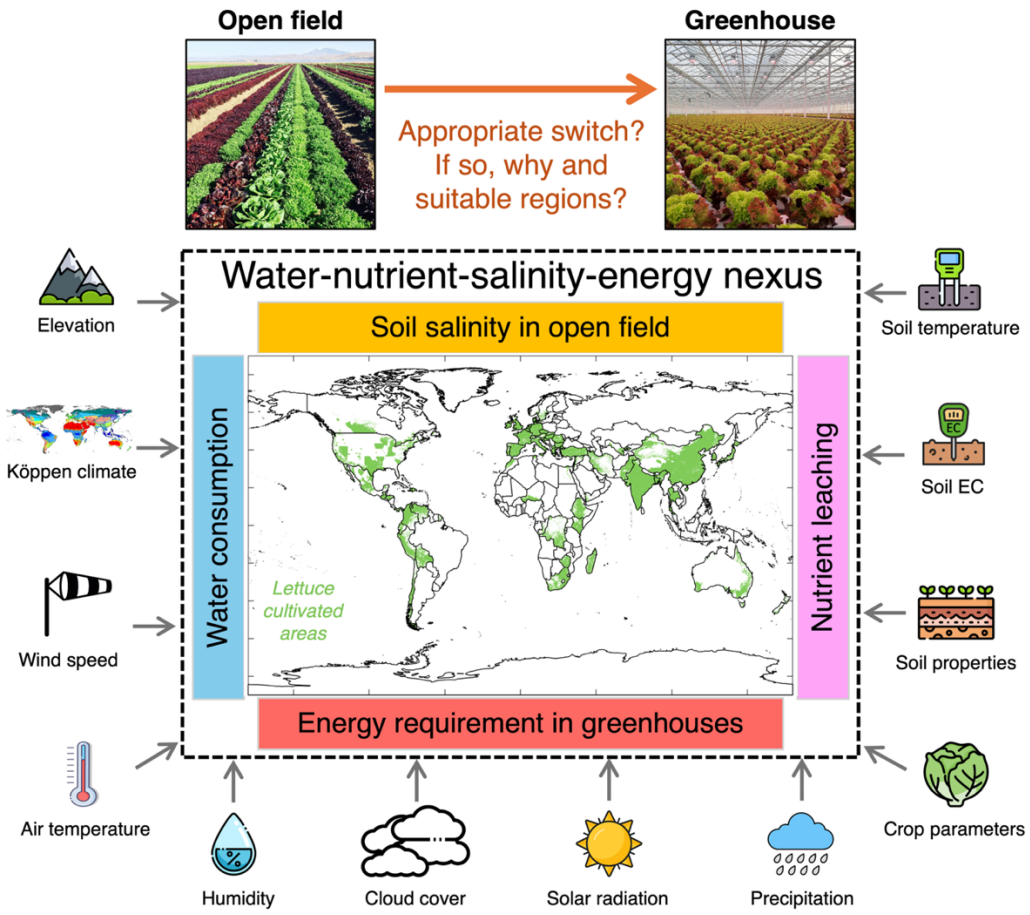


Figure 1. Flowchart of methodology: Schematic representation of the research goal, methodological framework, and datasets used in this study.

2.3.5.3. Sensitivity analysis of greenhouse energy demand

Cooling and heating set points are critical parameters that growers can dynamically adjust to optimize greenhouse climate control. To evaluate their influence on energy demand, we conducted a sensitivity analysis in which the cooling set point was reduced from 28°C to 26°C, and simultaneously the heating set point was raised from 20°C to 22°C. Thus, the temperature set point range was modified from 8°C (20–28°C) to 4°C (22–26 °C). Both configurations preserved environmental conditions suitable for optimal lettuce cultivation while inducing measurable

changes in heating and cooling demands, underscoring the importance of strategic set point management in balancing crop productivity with energy efficiency. The uncertainty and variation in heating and cooling loads were evaluated using the following equation:

$$\text{Relative Difference} = \frac{\text{Results}_{(22°\text{C}-26°\text{C})} - \text{Results}_{(20°\text{C}-28°\text{C})}}{\text{Results}_{(20°\text{C}-28°\text{C})}} \times 100 \quad (2.10)$$

2.4. Results and discussion

2.4.1. Water-nutrient-salinity interactions in open-field

2.4.1.1. Water

As expected, the blue water (BW) of furrow irrigation was notably higher than drip irrigation (Figure 2), primarily due to its lower efficiency (60%), which was driven by substantial losses through soil evaporation and deep percolation in contrast to the higher efficiency of drip irrigation (90%). However, in regions with high precipitation, furrow BW was lower than drip BW. Green water consumption (GW) was constantly larger for furrows than for drips (Figure 2). This was due to the dynamic differences in soil moisture between irrigation strategies, since maintaining soil conditions close to field capacity in drip irrigation (wetter soil) resulted in less rainfall retention in the root zone, as excess water drained away and became inaccessible to the roots. Thus, irrigation scheduling noticeably impacted GW, with the difference between furrow and drip irrigation ranging from 3.5% to 44.9%. This effect was particularly noticeable in specific climate categories. For instance, the mean GW of furrow was 44.9% (19.5 mm) and 31.9% (6.0 mm) higher than in the case of drip irrigation in Csc (temperate, dry summer, cold summer) and BWk (arid, desert, cold), respectively (Figure 3). The highest and lowest mean GW were 187.1 mm in Af (tropical, rainforest) for furrow irrigation and 19.0 mm in BWk for drip irrigation. Lettuce production was

found most in the Cfa (temperate, without dry season, hot summer) and Cwa (temperate, dry winter, hot summer) climates, where GW in the furrow system was found to be, on average, 16.5 and 12.2 mm higher than with drip irrigation, respectively. The highest mean BW withdrawal was 715.9 mm in furrow irrigation, which occurred in Dsa (cold, dry summer, hot summer). Interestingly, in certain climate classes (i.e., Af, Cwc, Cfb, Cfc, Dfa, ET, and EF), furrow irrigation exhibited a mean BW withdrawal of 2.9 to 22.1 mm lower than drip irrigation (Figure 3). This highlighted the impact of irrigation management strategies, where furrow irrigation, despite its lower efficiency, was applied only when soil moisture depleted to 70% of the TAW (see Section 2.3.2.4). In contrast, drip irrigation was scheduled to maintain soil moisture near field capacity through daily applications. Notably, precipitation events played a critical role in keeping soil moisture above the 70% TAW threshold, effectively reducing the activation of the less efficient furrow irrigation system. Thus, this BW reduction was linked to variations in soil moisture dynamics because of the drier soil condition in furrow irrigation, which led to a higher capacity to retain rainfall (resulting in higher GW), eliminating the necessity for additional BW application. Conversely, the mean BW withdrawal in other climate classes was 8.0 to 224.5 mm higher with furrow irrigation.

More importantly, in Cfa (temperate, without dry season, hot summer), where most lettuce was produced globally, the difference in mean BW was only limited to 0.2 mm (furrow: 189.6 and drip: 189.4 mm) while GW was 64.6 and 81.2 mm in drip and furrow systems, respectively (Figure 3). Therefore, in regions with relatively low BW withdrawal in the furrow system, rainfed lettuce production could be proposed, but a more detailed analysis of the combined water and salinity stress simulation was required to evaluate the benefit of using more GW. On the other hand, parts of the regions such as the United States, India, and China with high BW are located in severe

unsustainable BW regions identified by Mekonnen & Hoekstra (2020) for other crops (e.g., wheat, maize, rice, and soybean) led to the competition for water and importance of switching to greenhouses.

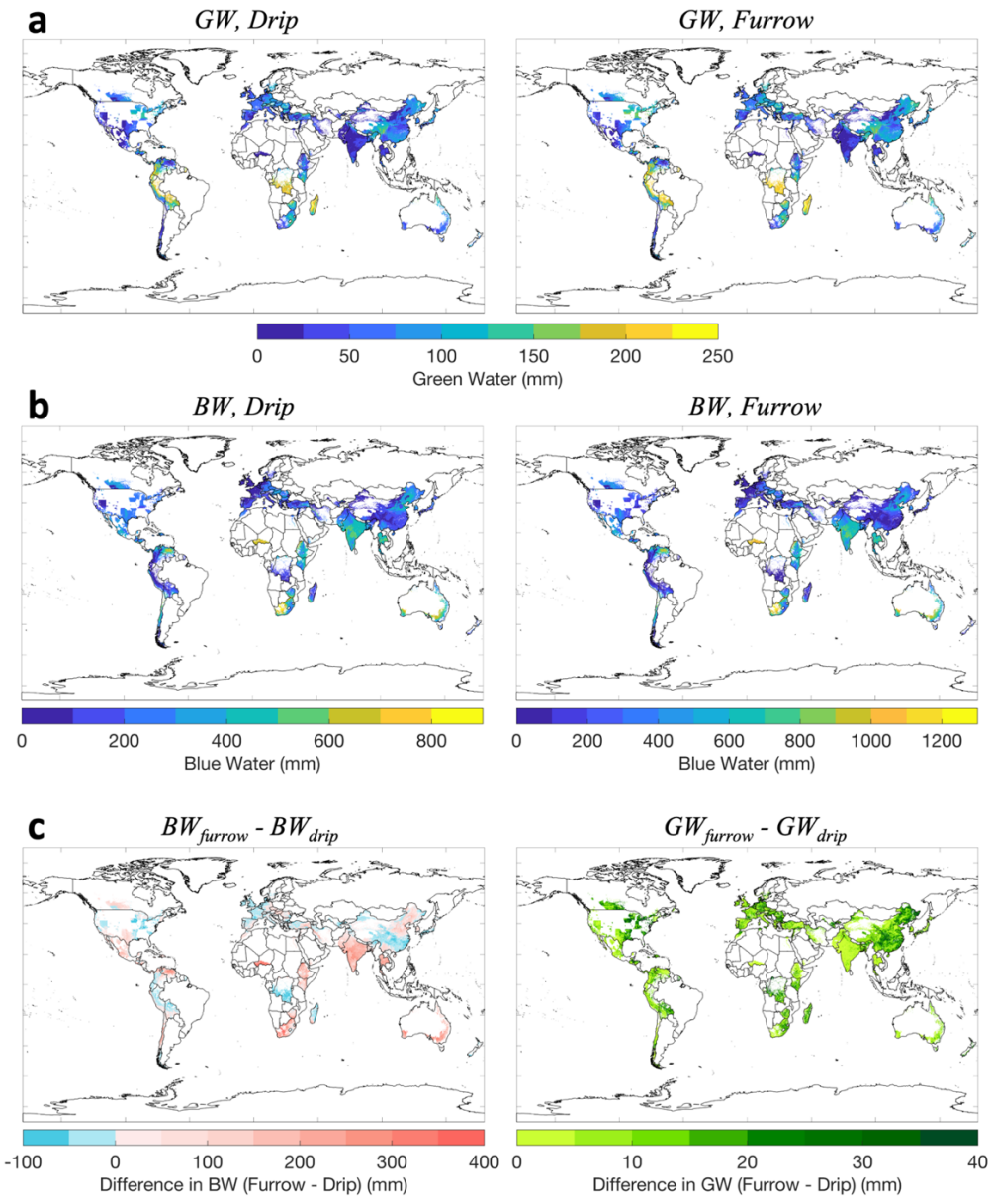


Figure 2. Global water dynamics in lettuce cultivation (2015-2016), showing: a) green water consumption (rainwater absorbed by soil), b) blue water withdrawal (extraction of surface and groundwater), and c) differences between blue and green water with furrow and drip irrigations.

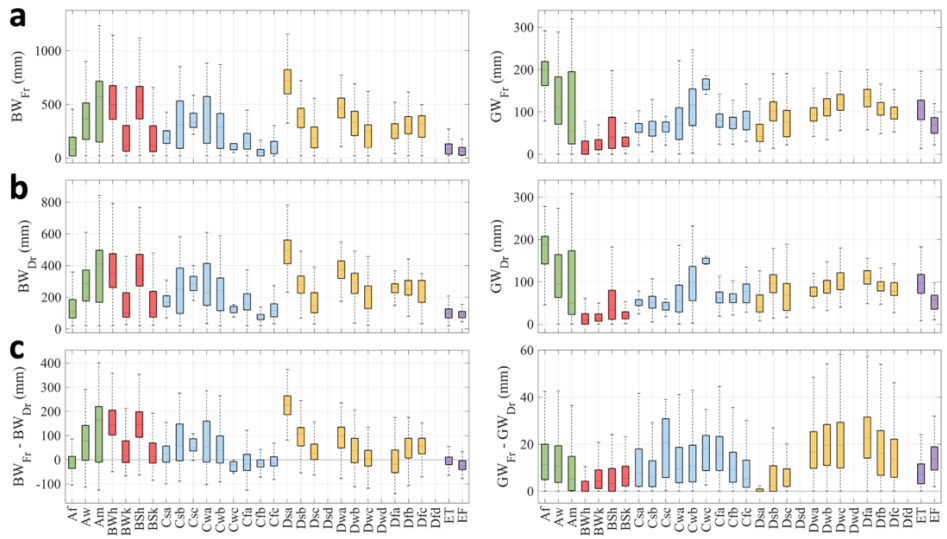


Figure 3. Comparison of green and blue water usage in lettuce production with furrow (a) and drip (b) irrigation and their differences (c) across 30 Köppen-Geiger climate classifications. Boxplots are color-coded to represent the five primary climate groups: A (tropical, green), B (arid, red), C (temperate, blue), D (continental, orange), and E (polar, violet).

2.4.1.2. Salinity

To evaluate the salinity outcomes of furrow vs. drip irrigation systems, we defined the "efficacy" of irrigation as the ratio between the applied water with non-saline soil, which differed from the irrigation efficiency (the ratio of irrigation water consumption and irrigation water withdrawals/supply). Efficacy hinged upon multiple variables. Salinization dynamics were linked to irrigation EC, with amplification or mitigation contingent on this parameter. Notably, the impact of irrigation system selection appeared less important for specific regions, primarily attributed to low ET_c or high precipitation in those climatic categories. Figure 4 shows that furrow irrigation was a more effective salinity management strategy. However, such efficacy comes at the cost of higher BW withdrawal, a potentially limiting factor in certain regions dealing with water scarcity (Liu et al., 2022). On the contrary, drip irrigation exemplified remarkable efficacy—distinct from

efficiency—in managing salinity, also reported by Hopmans et al. (2021). Soil characteristics, ranging from clay to sandy composition, further modulated the salinity mitigation or amplification dynamics. The effectiveness of sandy soil in managing high irrigation salinity was reported for dryland tomato production (Karimzadeh et al., 2024). However, in lettuce production, the influence of soil type was less significant due to the considerably shallower maximum root depth of lettuce (0.4 m) compared to tomatoes (1.1 m).

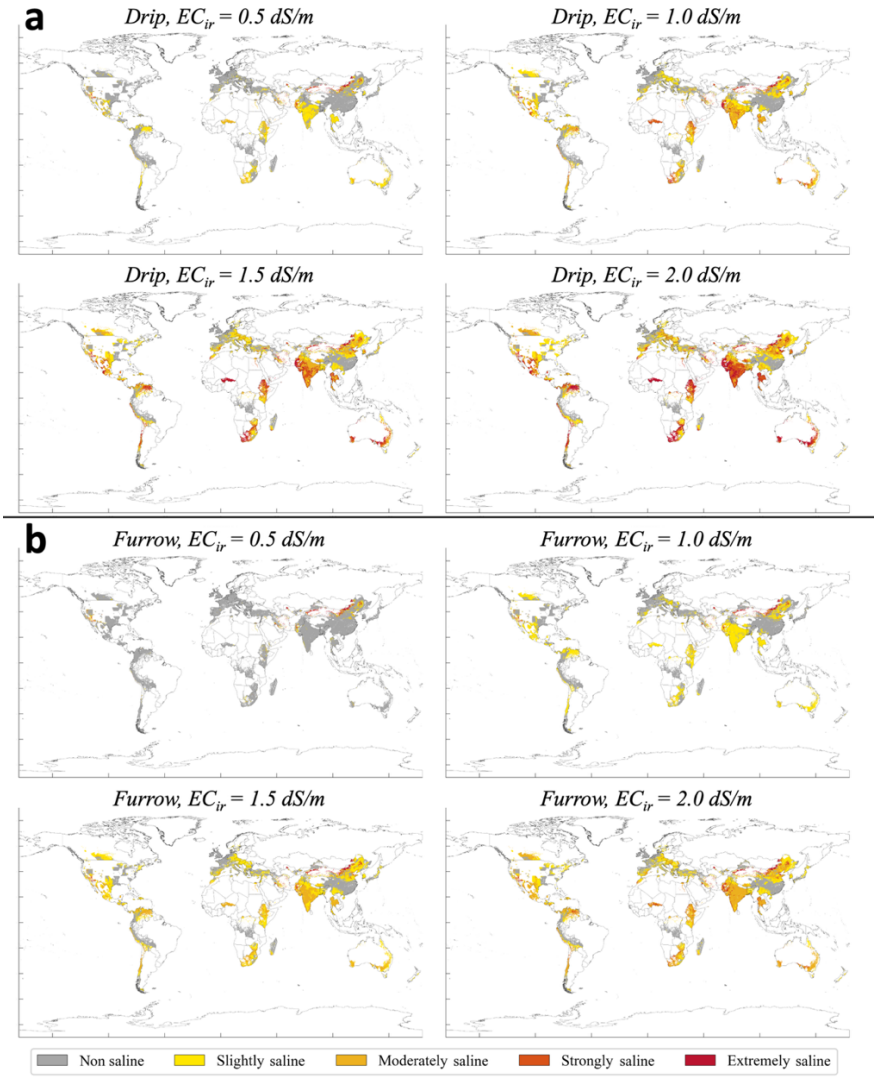


Figure 4. The impact of applying various irrigation electrical conductivity (EC) of 0.5, 1.0, 1.5, and 2.0 dS/m, employing furrow (a) and drip (b) systems.

The reflection of increasing water irrigation salinity and the effectiveness of irrigation systems varied among the climate classes. Figure 5 shows the quantity of salt-affected areas and the level of severity of salinity across climate classes and countries. Cfb (temperate, without dry season, warm summer) showed great resilience to increase applied water salinity. As furrow and drip irrigation salinity rose from 0.5 to 2.0 dS/m, the non-saline areas decreased by only 20.6% (from 42,561 to 33,801 ha) and 23.5% (from 42,379 to 31,280 ha), respectively. This was primarily attributed to the synergy between considerable average precipitation (3.3 mm) and the practice of letting the soil retain rainfall, which had a better water quality ($EC_{rainfall}$ of around 0 dS/m). This approach contrasted with the alternative of initiating irrigation with a higher salinity level of 2.0 dS/m. Similar behavior was observed in furrow irrigation in the Cfa (temperate, without dry season, hot summer) class with an average precipitation of 4.3 mm. This trend appeared to be influenced by the aridity index (AI), calculated as the ratio of mean annual precipitation to mean annual potential evapotranspiration (Trabucco & Zomer, 2022). The AI values for regions classified as Cfb and Cfa were 1.69 and 1.57, respectively. Notably, despite having a high AI of 2.0, regions categorized as Am (tropical, monsoon) were still impacted by high irrigation EC. This was largely due to the timing of planting (January 1st, 2016) and harvesting (April 15th, 2016), which did not align with the peak monsoon period in July and August, when 60 to 70% of the total annual rainfall typically occurred (Mahendra et al., 2023). This highlights the significance of synchronizing crop calendars with climatic patterns for effective agricultural salt management.

In the case of drip irrigation, due to more frequent irrigation (i.e., irrigation initiation when ET_c>P), the rate of adding salt by irrigation (EC_{ir} = 2.0 dS/m) was higher than the washing caused by precipitation, which led to a decrease in global non-saline areas from 308,648 ha (furrow) to 235,344 ha (drip) (Figure 5). Thus, when irrigation salinity was higher than the salinity tolerated

threshold of lettuce (EC_{ir}>EC_{thr}), decreasing the number of irrigation events could manage the secondary salinization. The impact of high EC irrigation was activated instantly in BWh (arid, desert, hot) with an AI of 0.15 (lowest among the classes). In such regions, the salinity was mainly governed by high atmospheric water demand and very low precipitation, which was consistent with Perri et al. (2022).

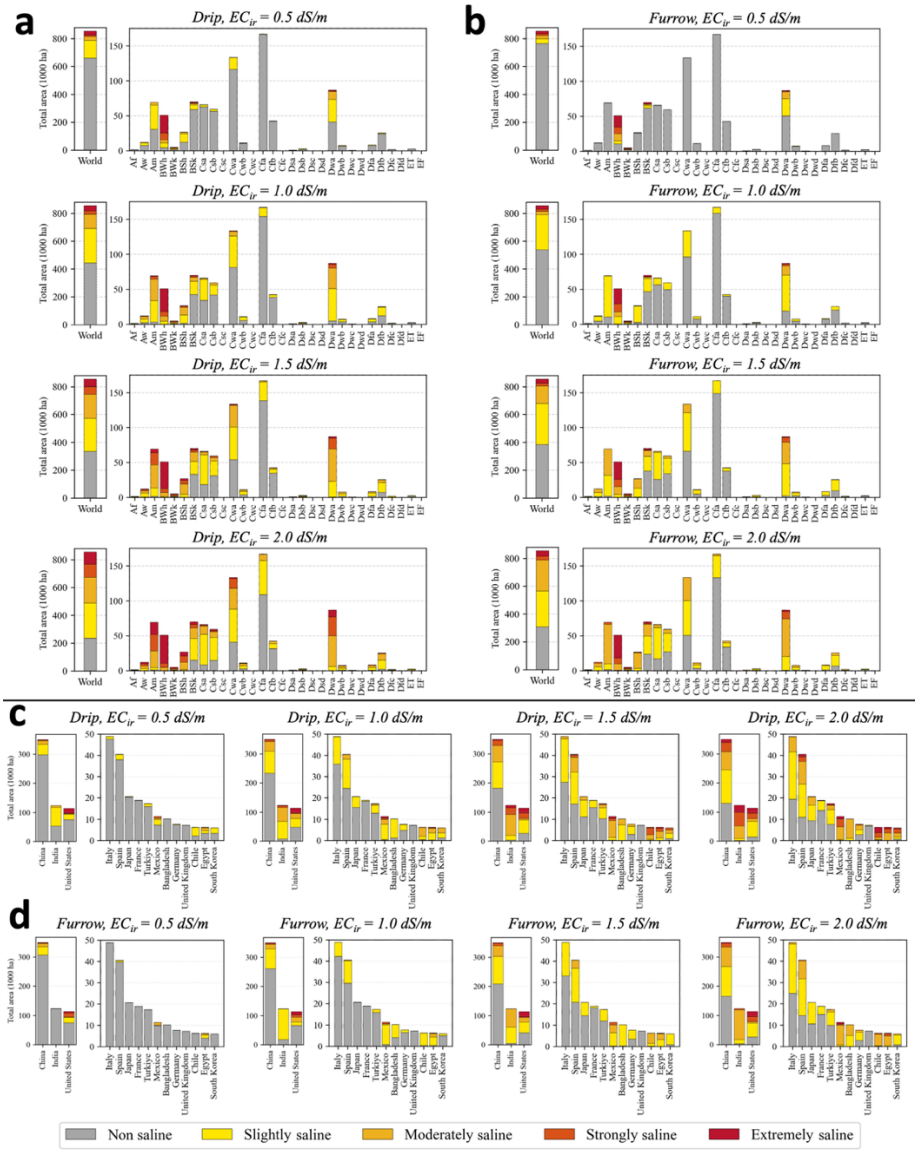


Figure 5. Lettuce salt-affected areas using (a) drip and (b) furrow strategies at irrigation EC levels of 0.5, 1.0, 1.5, and 2.0 dS/m across 30 Köppen-Geiger climate classes (c) and major lettuce-

producing countries (d). This figure illustrates the adaptability of climate conditions and the impact of EC levels on an irrigation strategy for lettuce cultivation, providing concise insights into global agricultural practices and water management strategies.

The country-level analysis (Figure 5c-d) also showed various countries' compatibility withstanding high salinity irrigation. For example, increasing water salinity from 0.5 to 2.0 dS/m reduced the non-saline areas in the United Kingdom by only 1.2% (from 7,225 to 7,212 ha) and 4.2% (from 7,225 to 6,924 ha) for furrow and drip irrigation, respectively. Conversely, the resiliency in tolerating high salinity irrigation was substantially low in India, Mexico, Bangladesh, Chile, and Egypt. The mentioned results were properly reflected by the introduced ISS, which quantified the importance of irrigation water quality and strategy in the resiliency of country's soil salinity stress. For example, by switching from furrow with a low irrigation EC of 0.5 dS/m to drip with a high EC of 2.0, the ISS abruptly increased from 0.3 to 90.8 in India, 11.4 to 91.2 in Mexico, 0 to 82.7 in Bangladesh, 0.9 to 86.1 in Chile, and 35.1 to 92.4 in Egypt whereas remained low for the United Kingdom (ISS = 2.1) and France (ISS = 14.6) (Table 3). Thus, countries with high ISS and water consumption could switch to hydroponic greenhouse systems, which offered a sustainable strategy to conserve water resources and mitigate the impacts of soil salinity. The AI was 1.8, 2.1, 1.0, 3.0, 0.6, and 0.0 for the United Kingdom, Bangladesh, India, Chile, Mexico, and Egypt (Table 3). The reason for high salt-affected areas in Bangladesh despite high AI was the high initial soil EC caused by saltwater intrusion in the coastal areas and the off-season monsoon months (described for Am climate class). Similar results occurred in Chile, where most of the lettuce planting date was either November 1st, 2015, or January 16th, 2016, while a severe El Niño-induced anomaly occurred in 2015 (Cai et al., 2020; Garreaud et al., 2020). when irrigation was applied with water EC of 2.0 dS/m, the most resilient countries were Ireland, the Netherlands, Ecuador, and Belgium, by having

non-saline lettuce areas of 100.0, 99.6, 99.0, and 96.6 % with AI of 2.1, 1.3, 4.8, and 1.3, respectively occurred with furrow system. In such countries, allowing water stress (e.g., rainfed cultivation) instead of applying high-salinity irrigation to mitigate water stress may be more beneficial to adopt compared to the salinity stress induced by high irrigation EC. Table 3 comprehensively summarizes the interaction of water and ISS in furrow and drip systems for the top fifteen countries in lettuce production.

Table 3. Comparative analysis of water consumption and irrigation salt-score (ISS) in open field lettuce cultivation (furrow and drip systems) across the top 15 lettuce-producing countries based on harvested areas.

			Furrow				Drip			
Country	**Harvested areas (1000 ha)**	**Annual aridity index**	**Total water consumption (Mm^3)**	**Average water consumption (mm/m^2)**	**ISS with EC_{ir} = 0.5 dS/m**	**ISS with EC_{ir} = 2.0 dS/m**	**Total water consumption (Mm^3)**	**Average water consumption (mm/m^2)**	**ISS with EC_{ir} = 0.5 dS/m**	**ISS with EC_{ir} = 2.0 dS/m**
China	350.5	1.2	1,137	292	9	38	992	252	11	46
India	123.2	1.0	748	592	0	92	542	424	53	91
United States	113.4	0.8	343	335	26	58	284	278	26	68
Italy	48.8	1.3	88	195	0	32	83	178	2	40
Spain	40.5	0.6	70	172	1	46	62	155	5	54
Japan	20.7	2.5	51	267	0	31	51	246	1	37
France	18.9	1.1	25	157	0	10	27	151	0	15
Turkiye	17.3	0.8	43	287	0	32	38	250	6	41
Mexico	11.3	0.6	54	467	11	92	40	343	32	91
Bangladesh	10.2	2.1	50	535	0	84	39	390	0	83
Germany	7.7	1.2	22	272	0	43	19	247	3	45
United Kingdom	7.2	1.8	7	100	0	0	8	105	0	2
Chile	6.3	3.0	38	377	1	87	27	284	58	86
Egypt	6.3	0.0	23	389	35	93	17	283	44	92
South Korea	5.9	1.2	24	398	0	70	20	349	34	70

2.4.1.3. Nutrient leaching

Although furrow irrigation was more effective in managing salinity due to the inherent low irrigation efficiency (60%), it substantially increased salt leaching (i.e., nutrients provided by

fertilizer) as deep percolation to groundwater. In other words, furrow irrigation washed out the rootzone, reducing lettuce's susceptibility to salinity in the topsoil while conveying the burden into ground resources. In contrast, drip irrigation performed considerably better in nutrient leaching due to lower blue water withdrawal and higher irrigation efficiency (90%). This was attributable to reduced irrigation water usage, which inherently delivered a lower salt/nutrient application mass.

The benefit of selecting irrigation systems to mitigate nutrient leaching varied substantially among countries. For instance, in the case of irrigation with water EC of 2.0 dS/m in China, the drip system resulted in 903 g/m^2 of salt leaching compared to 1,316 g/m^2 in furrow irrigation (a 45.7% reduction in nutrient leaching achieved by 14.8% savings in total water consumption). This effect became even more pronounced in countries with higher BW, such as India, Bangladesh, and Egypt, as irrigation was more frequent in the furrow system due to the synergy of high atmospheric water demand and limited precipitation during the growing cycle. In particular, this effect was most significant in Egypt, which relies heavily on irrigation (i.e., BW). Interestingly, Figure 6a shows that even using high-EC irrigation water (EC of 2.0 dS/m) in a drip system resulted in 1,506 g/m^2 of nutrient leaching, which is 24% less than the nutrient leaching from using low-EC irrigation water (0.5 dS/m) applied through a furrow system (1,989 g/m^2 of nutrient leaching). Figure 6 presents the global map and country-level nutrient leaching for lettuce production in the selected countries.

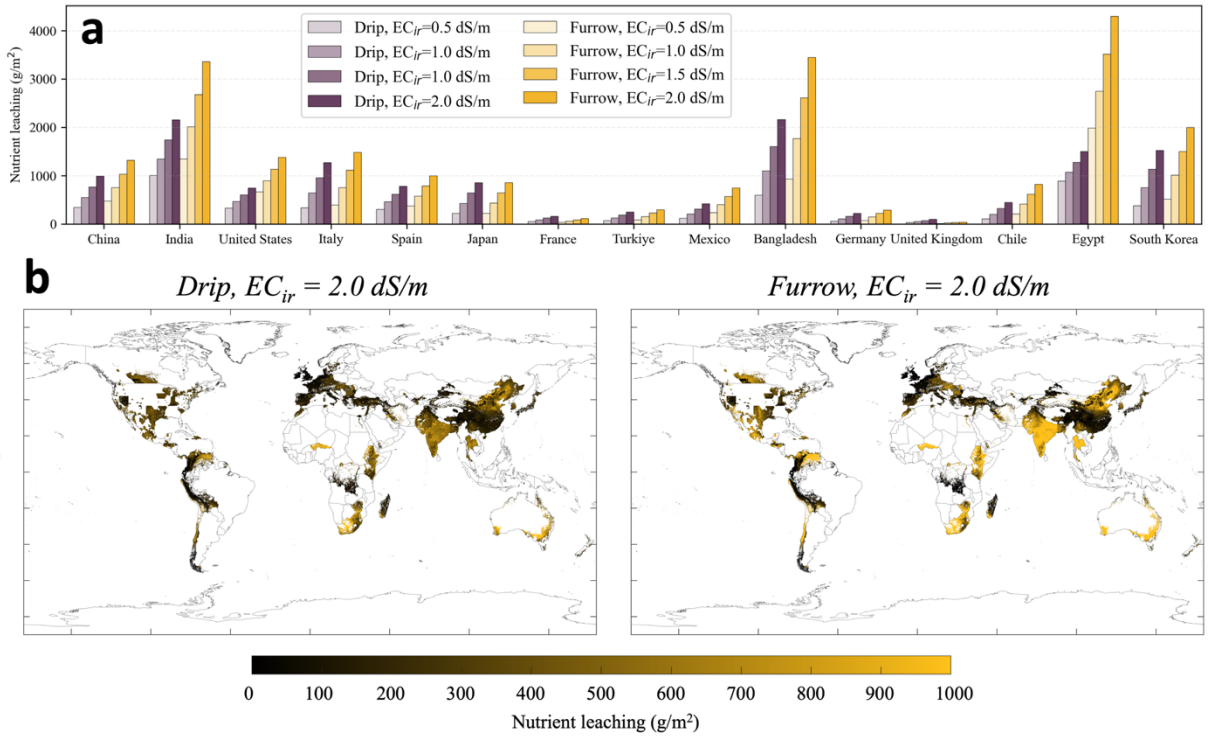


Figure 6. A comparative analysis of nutrient leaching into groundwater resources across various irrigation techniques and salinity levels. Higher salinity reflects greater dissolution of inorganic salt fertilizers in irrigation water. The figure includes a) country-level and b) global maps illustrating nutrient leaching from furrow and drip irrigation methods at an EC of 2.0 dS/m.

Nutrient leaching worsens groundwater salinity, elevating EC levels and degrading irrigation water quality. This, in turn, increases soil salinity during subsequent cultivation cycles, particularly in arid and semi-arid regions where precipitation (very low-EC water) is limited and insufficient to dilute soil salinity and groundwater. To break this cascading effect, transitioning to closed-loop hydroponic systems within controlled greenhouse environments offers a sustainable solution and is strongly recommended for regions with considerable ISS and high nutrient leaching risks, such as Egypt, India, and Bangladesh.

2.4.2. Water use in hydroponics and potential for sustainable water management

The sustainability of lettuce in greenhouses using hydroponic systems exhibited variability across diverse climate conditions and seasons. Figure 8d shows the duration, in months, where HBW from the roof was higher than CWR in greenhouses, with detailed monthly sustainability data provided in Figure 26. This analysis underscored the potential of adopting greenhouse-based lettuce production as a strategic response to water scarcity and mitigating soil salinity and nutrient leaching issues prevalent in open-field agriculture. In Bangladesh, a substantial portion of cultivated areas demonstrated sufficient water from HBW availability to sustain NFT greenhouse systems for three months annually from June to August and ~50% during April and May (Figure 8e). Furthermore, this approach effectively controlled the considerable nutrient leaching concerns, as elucidated in Section 2.4.1.3. Noteworthy regions, such as Japan, the southeast of China, the south of Chile, the north of Italy, the north of Iran, Colombia, and specific areas in Oregon, Washington, and the eastern part of Texas, exhibited the potential for sustaining over 50% of lettuce hydroponics water demand through HBW.

Although the benefit from greenhouses was evident in reducing lettuce water requirements (i.e., water consumed for closed-loop hydroponic solution), the water requirement for suitable climate conditions using a fan and pad system seemed to be comparably high. The ratio of evaporative cooling water requirements to hydroponic water use, as shown in Figure 8b, highlighted areas where cooling systems imposed a substantial additional water demand relative to the hydroponic requirements. In arid and semi-arid regions, particularly in South Asia, the Middle East, and portions of Australia, this ratio exceeded 80%, indicating that cooling systems may consume nearly as much or more water (i.e., coast of Queensland, Australia) than required for crop

evapotranspiration. These findings suggested that in such regions, the sustainability of greenhouse systems may be limited by their dependence on substantial cooling water inputs.

Conversely, temperate and tropical regions with moderate climates, including parts of Europe, the United States, and Southeast Asia, exhibited lower ratios, generally below 30%. These regions benefited from naturally favorable climatic conditions for lettuce production, reducing the reliance on water-intensive cooling systems. Interestingly, some high-latitude regions with cooler climates, such as northern Europe and Canada, demonstrated ratios near zero, reflecting minimal cooling water needs. Dry and cold areas could also benefit from free cooling with natural ventilation, which was ignored in this study.

The observed spatial trends emphasized the critical role of local climate conditions in determining the feasibility and efficiency of greenhouse systems. While hydroponic systems offered substantial water savings in crop cultivation, integrating fan-and-pad cooling systems could offset these advantages in water-scarce regions. These results underscored the need for tailored strategies that optimize cooling efficiency, such as hybrid cooling systems or alternative climate control methods, particularly in areas where water scarcity already limited agricultural productivity.

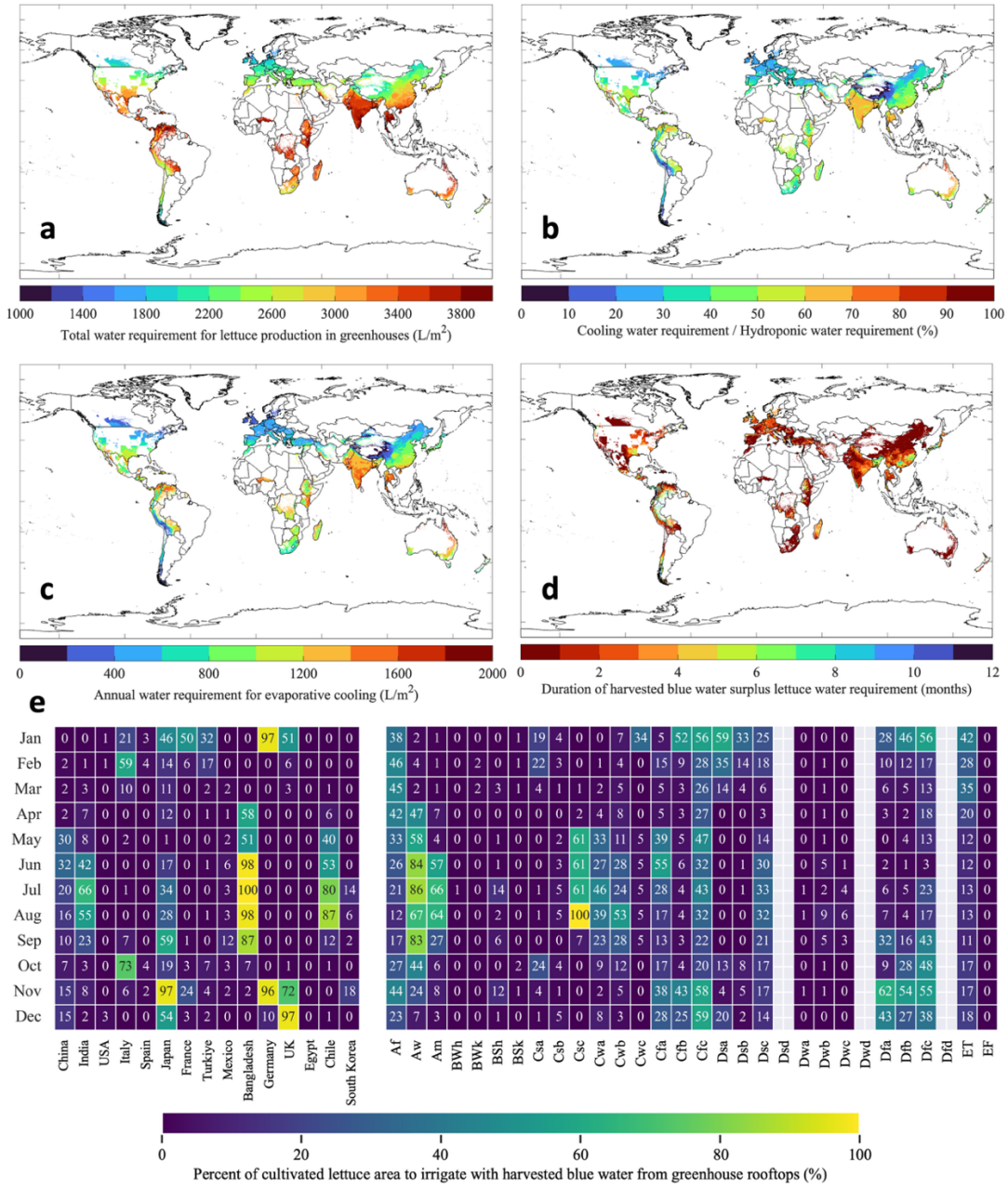


Figure 7. Global mapping of greenhouse systems water sustainability: a) total annual water requirement, b) the ratio of water for the pad and fan system over hydroponic water requirement, c) annual water required for the pad and fan system to meet cooling demand, d) number of months with sustainable water use (the amount of harvested blue water exceeds the crop water requirement) for lettuce greenhouse production ("0" indicating no months exceeding requirement),

e) percent of total cultivated area with sustainable water consumption across various months, countries, and climate classification.

2.4.3. Greenhouse energy dynamics

Heating and cooling energy requirements varied remarkably across the globe, primarily following the climate classification depicted in Figure 8a–c. The highest annual cooling requirement was 3,867 MJ/m^2, observed in the northern part of Queensland, Australia, classified as BSh (arid, steppe, hot), while the highest heating requirement was 5,018 MJ/m^2 in northern India, in the Jammu and Kashmir region, classified as EF (polar, frost). At the country-level analysis, the maximum average heating and cooling were 2,375 and 2,954 MJ/m^2 in Nepal and Sudan, respectively. In terms of climate analysis, the maximum average heating and cooling were 3,867 and 2,505 MJ/m^2 in EF (polar, frost) and BWh (arid, desert, hot). For Cfa, Cwa, and Dwa, which represent the largest share of global lettuce cultivation areas in climate classes (~45% of the total cultivated area), the total annual average energy requirement was 2,352 MJ/m^2 (cooling: 1,832 MJ/m^2, heating: 519 MJ/m^2), 2,448 MJ/m^2 (cooling: 2,096 MJ/m^2, heating: 352 MJ/m^2), and 2,465 MJ/m^2 (cooling: 1,335 MJ/m^2, heating: 1,130 MJ/m^2), respectively.

Our calculations for the United Arab Emirates (BWh climate) indicated a sensible cooling load of 2,872 MJ/m^2, which was ~6% lower than the values reported by Graamans et al. (Graamans et al., 2018). Additionally, they estimated the combined sensible and dehumidification load to be around 5,500 MJ/m^2. The reduced cooling load could be due to the higher cooling set-point in our study (28°C) compared to theirs (24°C). However, our analysis also identified a heating load of 117 MJ/m^2, which was reported as negligible in their study. This discrepancy could also arise from differences in greenhouse structures and covering materials (glass vs. polycarbonate).

In contrast, for Sweden, our calculations indicated a heating load of approximately 1,391 MJ/m² and a cooling load of 754 MJ/m^2. The heating load was higher by 20%, but the cooling value was more than double that of Graamans et al. (2018) due to not considering free ventilation. Reported heating leads for Scandinavia (1,908 MJ/m^2) and the United Kingdom (1,584 MJ/m^2) were ~20% higher than our estimates (Costantino et al., 2021). In Bucharest, Romania, reported heating and cooling loads for lettuce production (19/25 °C night/day setpoints) were 2,297 and 2,329 MJ/m^2, nearly double our estimates (Brækken et al., 2023). For closed greenhouse cucumber production in Tehran, Iran, reported heating and cooling demands were 425 and 1,124 MJ/m^2, compared to our estimates of 715 and 1,682 MJ/m^2. The discrepancy likely stems from higher day/night setpoints (25/18 °C in cold months, 29/25 °C in hot months) used in their study (Banakar et al., 2021). Lowering the setpoint from 29 °C to 27 °C increased cooling demand by ~46% in Almería, Spain (Prieto et al., 2023). In Pakdasht, Iran, heating and cooling demands for tomato production (24/20 °C Day/night) were reported as 4,536 and 10,879 MJ/m^2, respectively. Our analysis for lettuce at the exact location yielded substantially lower values: 674 and 1,743 MJ/m^2. A comparison between the estimates from this study and the greenhouse heating energy demand map for Iran (Nasrollahi et al., 2021) is presented in Figure 31.

In Switzerland, tomato production in multi-span Venlo greenhouses (20/16 °C day/night) required 645 MJ/m^2 for heating (Golzar et al., 2018), while our analysis for lettuce at 28/20 °C showed 1,084 MJ/m^2—suggesting that a 4 °C increase in setpoint may raise heating demand by ~68%. The heating demand in Saskatoon, Canada, was approximately 1,834 MJ/m^2, exceeding the cucumber requirement reported by Ahamed et al. (2019) by 10%. Despite utilizing the same heating set point (20°C), the extended photoperiod of 18 hours for cucumber cultivation, as opposed to 16 hours for lettuce production, necessitated operating more supplementary lighting, which led to mitigating

the heating load. In London, Ontario (Canada), reported heating and cooling demands for tomato production (18/26 °C night/day) were 1,367 and 1,283 MJ/m^2, respectively. Our estimates for lettuce under 20/28 °C setpoints were 1,482 MJ/m^2 for heating and 884 MJ/m^2 for cooling.

Reported annual energy demands for tomato cultivation in semi-closed greenhouses in Almería, Spain, range from 620/140 MJ/m^2 (heating/cooling) to 223/95 MJ/m^2 (Prieto et al., 2023; Van 'T Ooster et al., 2012). Under comparable conditions for lettuce, our simulations estimated 350 MJ/m^2 for heating and 1,479 MJ/m^2 for cooling. The markedly higher cooling demand observed in our study can be attributed to several key differences in assumptions and system configurations. First, prior assessments often excluded tomato cultivation during the peak summer months (July–August), when extreme internal temperatures render greenhouse production unviable under conventional practices. Second, our model assumed a fully closed greenhouse, thereby excluding the mitigating effect of free cooling through natural ventilation. Third, there were notable differences in thermal setpoints: while Prieto et al. (2023) assumed a constant cooling threshold of 29 °C for tomatoes, and Van 'T Ooster et al. (2012) employed seasonal setpoints ranging from 16 °C (winter) to 26 °C (summer), our lettuce model adopted intermediate setpoints of 20 °C for heating and 28 °C for cooling. Additional variations in greenhouse design and climate control strategies—such as ventilation protocols and cover materials (e.g., a single-layer plastic film combined with internal shading screens and NIR whitewashing)—further contributed to the divergence in energy demand profiles.

Interestingly, the most suitable greenhouse production seemed to be in the Cwc (temperate, dry winter, cold summer) with a total annual energy requirement of 1,854 MJ/m^2 (heating: 1,141 MJ/m^2, and cooling: 713 MJ/m^2) which was 21% lower compared to total annual energy requirement in Cfa (temperate, without dry season, hot summer), which was the climate class with

the highest lettuce cultivated areas. Figure 9 illustrates the energy demand for major lettuce-producing countries by cultivation area, revealing whether heating or cooling requirements dominate. For instance, in India, Mexico, Bangladesh, and Egypt, the cooling load was considerably dominant (9.5, 7.4, 20.7, and 6.5 times higher than the heating load, respectively), while for other countries, a balance was seen in both heating and cooling or slightly heating dominated energy load (i.e., the United Kingdom, Germany, and Chile). Similarly, the trend seemed more apparent when climate-based analysis was applied. The results showed that all tropical regions (i.e., Af, Aw, and Am), some arid regions (i.e., BWh, BSh), and temperate regions (i.e., Csa) were mainly dominated by the cooling load (Figure 9b). The highest density of lettuce production per pixel was in Salinas, California, United States, where heating and cooling loads corresponded to 505 and 1,677 MJ/m^2, respectively. It is also important to note the fluctuations of the energy fluxes in various months. For example, the Cfa cooling load depicted in Figure 27 showed the domination of the cooling load from March to December. The year-round fluctuation of other energy fluxes seemed more uniform for CO_2, motor, soil conduction, and perimeter (Figure 27).

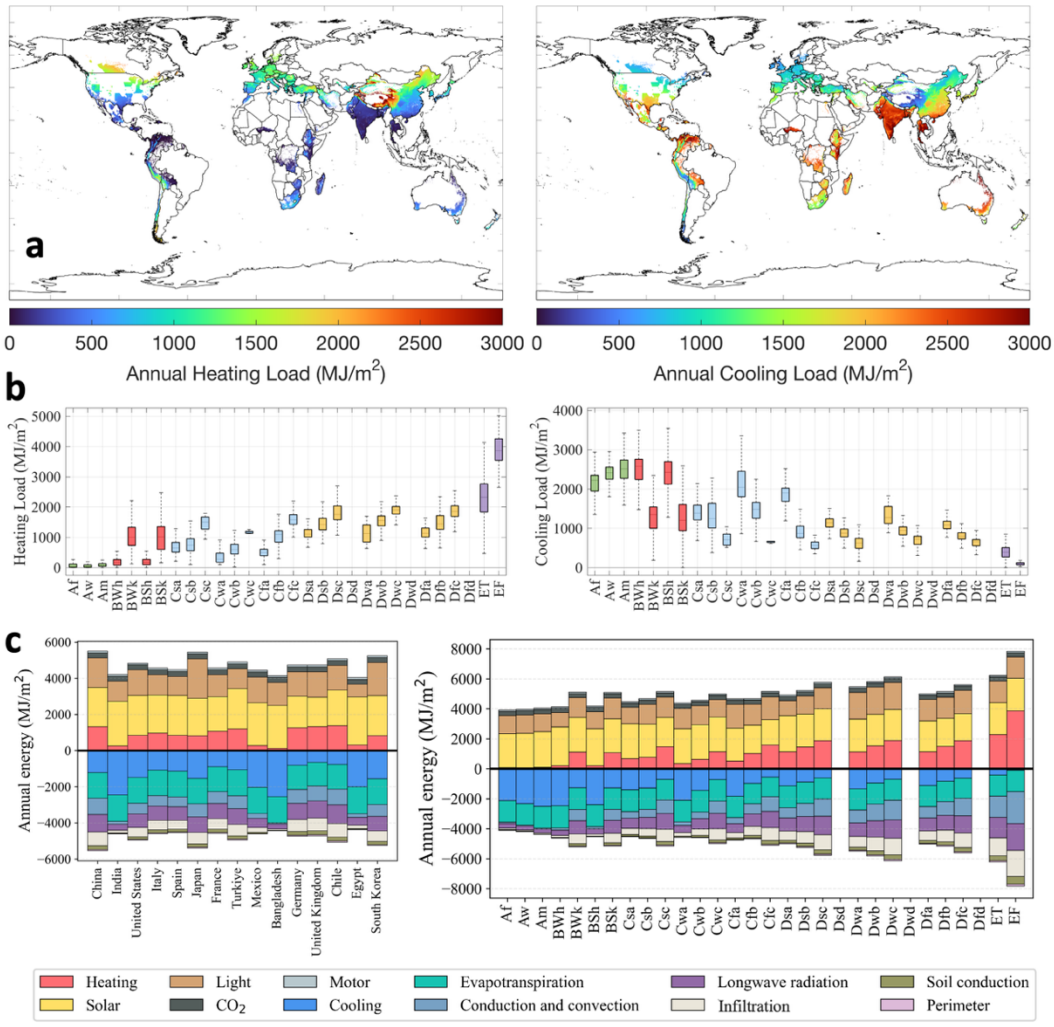


Figure 8. Global spatial distribution of heating and cooling energy requirements for lettuce greenhouse production (a), accompanied by boxplots detailing the load distribution for climate classes (b), average energy fluxes in countries, and climate classes (c).

2.4.4. Estimation of specific water and energy use

The spatial distribution of SWU and SEU for lettuce production highlighted the global contrasting efficiencies between greenhouse and open-field systems. In greenhouse systems (Figure 8a), SWU was significantly lower, ranging from 30 to 100 L/kg, with most regions demonstrating values below 90 L/kg, reflecting the high water-use efficiency of hydroponic systems. The SWU in this

study was higher than that of other reported studies (Graamans et al. 2018; Barbosa et al. 2015), which ranged between 15-20 L/kg of fresh lettuce in Arizona and the Netherlands. Compared with others, high water consumption could be due to two main reasons. First, the empirical Hargreaves model used in this study tends to overestimate the ET under humid conditions (Martinez & Thepadia, 2010; Tabari, 2010; Todorovic et al., 2013; Trajkovic, 2007). Secondly, this study considered closed greenhouses, so the water consumption in the fan-pad from increased cooling load could also be high. Generally, the water used for evapotranspiration could be half or more than the water used for cooling. For example, Sabeh et al. (2011) reported that 1 kg of fresh tomatoes was produced at the expense of 77 L of water in Arizona, of which more than 80% (62 L) was used for evaporative cooling.

SEU in greenhouses, however, was relatively high, spanning from 40 to 90 MJ/kg, with regions such as China, India, and Australia showing the highest energy demand. In this study, the estimated SEU for greenhouses in Finland, ranging from 50 to 55 MJ/kg, was almost half of the previously reported value of 97 MJ/kg (Frans & Juha-Matti, 2021). Similarly, the estimated SEU of 60 MJ/kg for greenhouse production was lower than the value (90 ± 11 MJ/kg) reported for Yuma, Arizona (Barbosa et al., 2015). In their analysis, the direct electricity costs associated with supplementary lighting were incorporated, a factor that was excluded in the definition of SEU used in this study (see Section 2.4.4). Additionally, the comparison was influenced by differing climate set points of the greenhouses. Meanwhile, SEU could be low for low-tech ventilated greenhouses, low indoor set-points for heating, and greenhouses without supplemental lighting.

Open-field systems exhibited more considerable variability and less efficiency in water use (Figure 9b). SWU under drip irrigation could be up to 800 L/kg, with arid and semi-arid regions, such as the Middle East and parts of Sub-Saharan Africa, experiencing the highest water demands. Furrow

irrigation further exacerbated water inefficiencies, with SWU reaching up to 1500 L/kg in some areas, such as India and South America, due to the compounding effect of high water requirement and low lettuce yield per unit cultivated area. These results underscored the inefficiencies associated with traditional irrigation methods compared to hydroponics in greenhouses.

Table 4 compares SWU and SEU for lettuce production in open fields and hydroponic greenhouses. China achieved an average SWU of 62 L/kg through hydroponics in greenhouses, a significant improvement over open-field practice. Similarly, India showed substantial inefficiencies under furrow irrigation, with an average SWU of 1,146 L/kg, which was reduced to 85 L/kg once hydroponics in greenhouses was adopted. Meanwhile, in terms of energy use, India demonstrated an SEU of 66 MJ/kg for growing lettuce in greenhouses. The trade-offs between water and energy use were evident: greenhouse systems significantly reduced water consumption and demanded higher energy inputs, particularly in regions requiring extensive climate control. Conversely, open-field systems conserve energy but faced substantial water, nutrient leaching, and salinity management challenges, particularly in water-scarce regions.

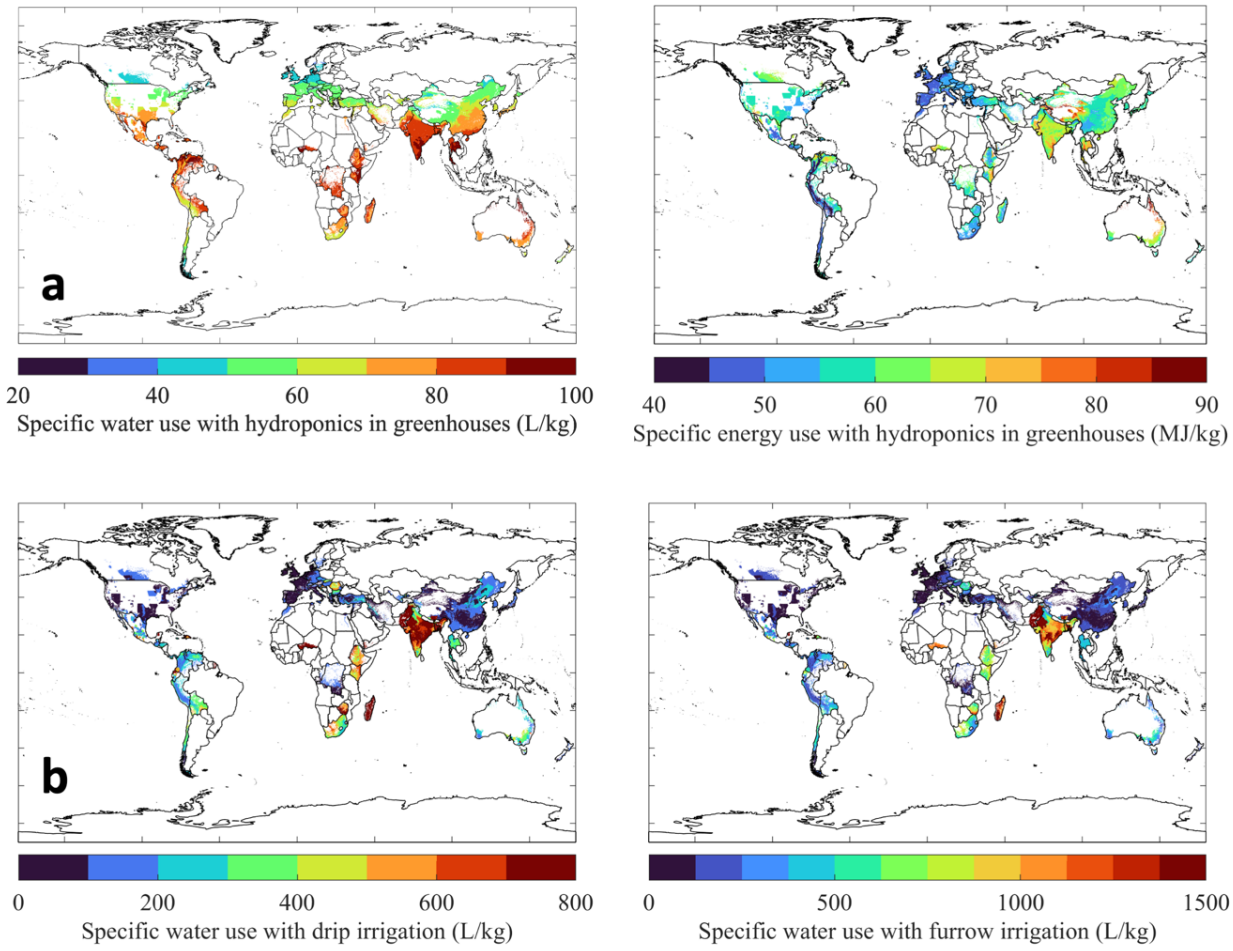


Figure 9. Global spatial variation in specific water (SWU) and energy (SEU) use per kilogram of lettuce production under contrasting agricultural systems, a) specific water use (L/kg) and specific energy use (MJ/kg) in greenhouses using hydroponic systems, b) specific water use (L/kg) in open-field agriculture under drip and furrow irrigation systems.

Table 4. Comparative analysis of country-level average specific water use (SWU) and specific energy use (SEU) for lettuce production under open-field cultivation systems (furrow and drip irrigation) vs. closed-loop hydroponics in greenhouse systems across the top 15 major lettuce-producing countries.

Country	**Open field**		**Greenhouse**	
	Furrow	**Drip**	**Closed-loop hydroponics**	
	Average SWU (L/kg)	Average SWU (L/kg)	Average SWU (L/kg)	Average SEU (MJ/kg)
China	152	130	62	62
India	1,146	821	85	66
United States	96	79	65	57
Italy	103	94	57	50
Spain	62	56	59	48
Japan	108	99	66	59
France	58	56	53	48
Turkiye	155	135	59	55
Mexico	352	257	79	56
Bangladesh	1,474	1,066	85	65
Germany	118	108	49	51
United Kingdom	45	48	45	48
Chile	282	211	55	53
Egypt	147	107	77	57
South Korea	157	138	67	59

2.4.5. Uncertainty analysis

2.4.5.1. Water-nutrient-salinity nexus in open-field

To evaluate the legacy effects of El Niño-associated climatic anomalies, a comparative assessment between the El Niño period (2015–2016) and a subsequent five-year climatological baseline (2019–2023) was conducted. Relative differences (%) were calculated across key indicators, including water source contributions (BW and GW), nutrient leaching, and salinity stress. Positive relative values indicated an increase in the post-El Niño period, while negative values denoted a decline compared to El Niño conditions.

Under open-field cultivation, BW remained virtually unchanged (−0.3%), regardless of irrigation method (furrow or drip), suggesting that irrigation inputs during 2015–2016 were comparable to the five-year average (Table 15; Figure 28). This stability likely reflected a buffering capacity in irrigation scheduling or minimal variation in crop evapotranspiration demand across the two periods. In contrast, GW showed modest increases (+1.7% under furrow and +1.8% under drip irrigation), indicating a greater reliance on precipitation-driven soil moisture in the post-El Niño climate (Figure 28).

These field-scale trends aligned with observed global precipitation patterns (Figure 30a–b), in which large regions—including northern India, the Sahel, and the eastern United States—experienced positive rainfall anomalies. A global average precipitation increases of +7.5% during 2019–2023 (Table 15) likely contributed to the reduced dependence on BW and enhanced GW uptake.

In greenhouse hydroponic systems, HBW declined by 4.4% (Table 15), despite the overall rise in precipitation. This counterintuitive response was attributed to changes in precipitation intensity

and frequency. Although total rainfall increased, a higher proportion of daily precipitation events fell below the 1 mm threshold for effective harvesting. Thus, the El Niño period was characterized by higher-intensity, lower-frequency rainfall events, in contrast to the five-year baseline which exhibited more frequent but less intense precipitation.

Nutrient leaching exhibited substantial changes. Losses decreased by 31–32% under furrow irrigation and by up to ~49% under drip irrigation across varying EC levels (Figure 29). These reductions were likely driven by (i) improved temporal rainfall distribution, which reduced the number of irrigation events and subsequent deep percolation, and (ii) lower nutrient concentration under more stable and favorable root-zone moisture conditions driven by precipitation.

Salinity stress coefficients also declined, particularly at higher EC levels (Figure 29). In furrow systems, stress levels decreased by up to 7.0%, while drip systems showed reductions reaching 8.8% at EC 2.0 (Table 15). These patterns suggested that even modest shifts in climatic water balance—primarily due to increased precipitation—facilitated root-zone salt dilution, particularly when combined with the spatial precision of drip irrigation. Although the observed reductions were moderate, they reflected enhanced soil flushing capacity and diminished salinity accumulation under wetter conditions.

2.4.5.2. Energy analysis sensitivity in greenhouses

The global sensitivity analysis revealed substantial shifts in greenhouse energy requirements when the temperature setpoints were narrowed from 20–28 °C to 22–26 °C (Figure 32). The relative changes varied both spatially and functionally, increasing and decreasing depending on regional climate conditions and energy transfer components.

Cooling energy requirements increased across most tropical and subtropical regions, particularly in equatorial Africa, South Asia, and northern South America (Figure 32a). Table 16 showed that tropical zones such as Af, Aw, and Am experienced increases in cooling demand by 8–10%. For example, in the Af zone, cooling energy rose from 2,123 MJ/m^2 to 2,328 MJ/m^2, corresponding to a 10% increase.

In contrast, heating energy requirements generally decreased, especially in temperate, continental, and polar regions (Figure 32b). The higher minimum setpoint reduced the frequency and intensity of nighttime heating, resulting in energy savings. In climates such as BWk and BSk, heating demand declined by 10%, with absolute reductions exceeding 100 MJ/m^2. The largest relative reduction occurred in the Cfa climate zone, where heating demand dropped by 16%.

The total energy requirement—the combined effect of cooling and heating—showed divergent patterns across climate zones (Table 16). In hot arid climates such as BWh, total energy demand increased by 7%, whereas in colder zones such as Dfc and ET, it declined by 6–8%. In equatorial climates (Af, Aw), the increased cooling energy more than offset the modest heating reductions, leading to a net energy increase of 8–9%.

At the component level, changes in energy fluxes were driven by shifts in environmental heat exchange. Longwave radiation losses increased in most low-latitude regions (Figure 32c), contributing to higher cooling requirements. Convective and conductive heat exchange also intensified in warmer climates (Figure 32d), while infiltration (Figure 32f) and greenhouse floor interactions (Figure 32e) showed modest changes, with slight reductions observed in some temperate areas. These findings highlighted the trade-offs involved in greenhouse climate control strategies and emphasized the importance of regional customization to optimize energy use efficiency.

2.4.6. Study limitations

While this study effectively modeled the spatial and temporal impact of various irrigation methods on lettuce production in open fields and greenhouses, it is essential to acknowledge its limitations. One notable limitation is the absence of consideration for the influence of salinity on the fertilizer uptake rate by plants in the current model. To address this, incorporating a more intricate biological sub-model becomes imperative, taking into account nutrient uptake and its effect on maintaining a balanced salt equation (Eq. 2.3). This enhanced sub-model is expected to contribute to a decrease in nutrient leaching and mitigate the occurrence of salt-affected areas.

Furthermore, it is crucial to note that the existing model only addressed the vertical water flow, neglecting the lateral flow. The consideration of salt inputs solely through irrigation water excluded other processes associated with the presence of shallow water tables. Salinity issues in irrigated agriculture, particularly those influenced by a shallow water table (<2 meters from the surface), are significant. Capillary rise brings water close to the ground surface, leading to evaporation and subsequent salt accumulation (Runyan & D'Odorico, 2010). Effectively managing the impact of a shallow water table through soil drainage is critical for successful salinity management and the long-term sustainability of irrigated agriculture. Additionally, it is pertinent to consider seawater intrusion into coastal aquifers as a significant factor in the enduring salinization trends observed in arid coastal regions (Jasechko et al., 2020). Also, in real-world applications, farmers may implement deficit irrigation strategies or utilize sprinkler systems to enhance water use and improve efficiency in the open-field.

Unlike traditional open-field soil-based cultivation, the effectiveness of hydroponic systems relies substantially on the correct application of a nutrient solution directly to the system (Kwon et al., 2021). In closed-loop NFT hydroponic systems, imbalances in macro- and micronutrient

concentrations could necessitate frequent disposal of the nutrient solution (Miller et al., 2020)—estimated at 30% in this study. However, this proportion could vary substantially depending on the nutrient formulation practices employed by growers and the quality of the irrigation water source. Nevertheless, treating nitrogen and phosphorous-rich discarded nutrient solutions in hydroponic systems proves to be more controllable than the uncontrollable nutrient leaching observed in conventional open-field settings.

Soilless greenhouse systems offer the advantage of operating independently of soil and climate constraints. However, their scalability remains dependent on key infrastructural factors such as water availability, transport networks, and market access. This study examined the transition from open-field agriculture to greenhouse systems, rather than the expansion into currently unsuitable areas.

Also, the approach for estimating water consumption for evaporative cooling was considered based on the closed concept of greenhouses, which does not always present the actual operation. However, the closed concept will increase water consumption as the load was estimated without considering the free cooling using outside cooled air (ventilation). Also, free cooling is commonly used for low-tech greenhouses and closed greenhouses in high northern latitudes. Consequently, actual water consumption for evaporative cooling in cold and dry regions would be substantially lower because of the advantages of free cooling. Similarly, the energy cost, especially for the cooling, was estimated for closed greenhouses, whereas most commercial greenhouses are ventilated in cooling mode to take advantage of free cooling. Thus, the SEU would be lower for low-tech and ventilated greenhouses. For farm-level assessments, high-resolution simulation methodologies—such as those implemented in EnergyPlus or TRNSYS (Çam et al., 2024; Hou et al., 2024; Opeyemi Ogunlowo et al., 2023)—are recommended to capture the detailed energy

dynamics and system interactions with sufficient precision. Besides, the ET model (Hargreaves & Samani Model) for the greenhouse is an empirical equation that tends to overestimate under humid conditions. Therefore, the actual water usage could be lower than the ported value in this study.

While this study primarily focused on the nexus of water, nutrient, salinity, and energy, overlooking the role of capital investment risks oversimplifying the complex decision-making landscape associated with transitioning to greenhouse production. Although climate resilience, yield enhancement, and resource-use efficiency are often foregrounded as key drivers, the substantial initial capital required to establish greenhouse systems represents a considerable barrier—particularly for smallholder and resource-constrained producers. Capital investment, therefore, is not merely a secondary consideration but a foundational determinant of feasibility and scalability. Current estimates of capital costs for greenhouse construction vary dramatically across contexts, ranging from ~1 USD/m^2 in developing countries to ~20 USD/m^2 in Qatar (Sajid & Bicer, 2021), ~300 USD/m^2 in Australia (Benke & Tomkins, 2017) and exceeding 900 USD/m^2 for high-tech dutch greenhouses (Zhuang et al., 2024). These disparities underscore the need for nuanced techno-economic assessments and inclusive policy frameworks that reflect the heterogeneity of production systems globally. However, a benefit-cost ratio (BCR) of 2.13 reported in the literature highlights the strong economic viability of the system (Souza et al., 2019). In line with this, Majid et al. (2021) demonstrated that deep water culture and NFT techniques outperformed conventional soil-based greenhouse lettuce cultivation, both achieving BCRs exceeding 2—justifying investment despite the initial capital burden.

2.4.7. Prospective avenues for future research

The adaptable methodology employed herein for the open-field water-nutrient-salinity-energy nexus holds promise for evaluating diverse crops and regions due to well-established crop

parameters from existing literature (Allen et al., 1998; Rhoades et al., 1992). The proposed crop-specific irrigation salt-score (ISS) is potentially applicable in national-level irrigated crop production. This innovative metric offers insights into the significance of applied EC for various crops beyond the scope of the current investigation. Its utility extends to providing valuable information on the relevance and impact of irrigation EC for diverse crop types, contributing to a more comprehensive understanding of irrigation practices. Besides, the proposed complete mix reactor-based agrohydrology model surpasses conventional EC monitoring by operating at the ion level. It can offer a detailed examination of nitrogen and phosphorus dynamics, providing enhanced agricultural and environmental management precision. Future research endeavors should explore the intricate interactions between various irrigation practices, especially those involving larger rhizosphere and soil characteristics, to enrich our understanding and guide sustainable agricultural practices.

The deep-water culture technique is also commonly used for hydroponic lettuce production. However, its high water consumption necessitates meticulous monitoring of EC and pH for optimal performance and can be evaluated. Concurrently, the emergence of vertical farms operating exclusively under artificial lighting raises questions about their global sustainability. Therefore, a comprehensive seasonal and spatial analysis of the energy footprint associated with vertical farms is imperative to assess their feasibility on a larger scale. The scope of our analysis was confined to the present lettuce cultivated areas, whereas soil-less greenhouses theoretically possess limitless potential. Consequently, the global applicability extends beyond the boundaries of existing lettuce fields. This suggests that our approach may have broader implications and could be extrapolated to diverse geographical regions where soil-less greenhouses could be implemented.

Recent advancements include semi-transparent photovoltaic systems installed on greenhouse roofs, demonstrating commendable environmental and economic feasibility (Agostini et al., 2021; Giri & Mohanty, 2022). These systems effectively fulfill greenhouse operations' heating, cooling, and lighting requirements (Fernández et al., 2022). Nevertheless, a thorough global analysis is warranted to delve deeper into their performance and potential implications, contributing valuable insights for broader adoption.

Rainwater harvested from greenhouse rooftops offers a high-quality source for hydroponic nutrient solutions; however, precise ion-specific nutrient management, particularly for key macronutrients such as nitrogen, phosphorus, potassium, calcium, magnesium, and sulfur, is essential to maintain optimal nutrient concentrations and mitigate imbalances. Thus, developing advanced ion-dosing algorithms is required to transition from open-loop to closed-loop hydroponic systems. Additionally, achieving enhanced water and nutrient efficiency in hydroponic greenhouse systems critically depends on advancing fault detection and diagnosis of sensors such as EC and pH sensors (Karimzadeh et al., 2025). Reliable sensor performance is essential for ensuring accurate farm-scale management. Therefore, future studies should focus on fault detection and diagnosis of sensors and the development of dosing algorithms to improve water and nutrient usage efficiency in hydroponic production.

2.5. Summary and conclusions

This study provided a holistic evaluation of the transition from conventional open-field, soil-based lettuce production systems (furrow and drip irrigation) to advanced closed-loop hydroponic systems in greenhouses through the lens of the water-salinity-nutrient-energy nexus. The findings revealed that soil salinity and nutrient leaching were persistent and inevitable challenges in open-field agriculture, necessitating mitigation strategies in arid and semi-arid regions. In contrast,

closed-loop hydroponic systems within greenhouses offered considerable advantages by curbing nutrient loss and salinity accumulation in soil; however, these benefits came at the cost of substantial energy requirements for greenhouse operations. Importantly, the results demonstrated that closed-loop hydroponics can achieve notable water savings and promote sustainable resource use in certain climatic regions and countries, underscoring its potential to address future challenges of food security and water scarcity. Key findings encompass:

Water management: Tailored irrigation strategies are essential. A well-designed furrow irrigation schedule can reduce water withdrawal by up to 22 mm compared to daily drip irrigation across seven climate zones (Af, Cwc, Cfa, Cfb, Cfc, Dfa, ET, and EF). This is achieved by optimizing the capture of green water within the plant's root zone, despite the lower irrigation efficiency of furrow irrigation. However, this requires accurate estimations of crop evapotranspiration and soil moisture.

Salinity control: By introducing the irrigation salt-score (ISS), the performance of irrigation systems and water salinity can be evaluated nationally. Countries such as the United Kingdom and France exhibit high resilience to high salinity irrigation for both drip and furrow irrigation systems. In contrast, countries such as India, Bangladesh, Chile, Mexico, and Egypt show notable sensitivity to the tested irrigation systems and applied water EC (0.5–2.0 dS/m) due to low precipitation and high evapotranspiration during the lettuce growing cycle. Drip irrigation demonstrates high efficacy in salt management, requiring lower water withdrawal to ensure non-saline areas. Meanwhile, furrow irrigation shows higher salt management effectiveness due to its greater leaching capacity.

Nutrient leaching: While effective in salt management, furrow irrigation causes substantial nutrient leaching. This effect is particularly noticeable in regions with low precipitation and high

evapotranspiration, as more irrigation is needed to manage water stress. Adopting drip irrigation is crucial for Egypt, where nutrient leaching with furrow irrigation at an EC of 0.5 dS/m is 483 g/m^2 (32%) higher than with drip irrigation using water at an EC of 2.0 dS/m. This contrasts with the United Kingdom, where the impact is less significant. The results help achieve a reasonable trade-off between minimizing water consumption, ensuring high salt management efficacy, and reducing nutrient leaching in open-field lettuce production.

Greenhouse water sustainability: Greenhouse-based lettuce cultivation emerges as a strategic solution to address water scarcity, soil salinity, and nutrient leaching. Harvested blue water (precipitation collected from greenhouse rooftops) can sustain the irrigation needed for closed-loop hydroponic lettuce production systems, meeting around 100% of the hydroponics water demand in Bangladesh during June, July, and August, in the United Kingdom during December, in Germany during January and November, and in Japan during November. In Cfa climates (temperate, without dry season, hot summer), where most global lettuce is produced, sustainable closed-loop hydroponics can occur up to 55% in June.

Greenhouse energy demand: The results provide a comprehensive global spatial distribution of heating and cooling energy needs, which are crucial for greenhouse adaptation and maintaining energy security. The minimum annual energy required for adapting high-tech closed greenhouses is 1,854 MJ/m^2, occurring in Cwc climates (temperate, dry winter, cold summer). Of this, 62% is dedicated to heating, and 38% to cooling.

Specific water and energy use per kg of lettuce: Closed-loop hydroponic systems drastically reduce specific water use (SWU), achieving savings from 6% in the United Kingdom (48 L/kg in a drip system to 45 L/kg in greenhouses) to over 90% in Bangladesh (from 1,474 L/kg in furrow systems to 85 L/kg) and India (from 1,146 L/kg in furrow systems to 85 L/kg). However, higher

specific energy use (SEU) offsets these gains, ranging from 48 MJ/kg in the United Kingdom to 66 MJ/kg in India.

The results indicate that hydroponic greenhouse production could be beneficial globally, especially in regions facing water and salinity challenges. This work highlights the trade-offs between sustainability and resource use in transitioning agricultural systems, providing actionable insights for policymakers, researchers, and practitioners.

2.6. References

Agostini, A., Colauzzi, M., & Amaducci, S. (2021). Innovative agrivoltaic systems to produce sustainable energy: An economic and environmental assessment. *Applied Energy*, *281*, 116102. https://doi.org/10.1016/j.apenergy.2020.116102

Ahamed, M. S., Guo, H., & Tanino, K. (2018). A quasi-steady state model for predicting the heating requirements of conventional greenhouses in cold regions. *Information Processing in Agriculture*, *5*(1), 33–46. https://doi.org/10.1016/j.inpa.2017.12.003

Ahamed, M. S., Guo, H., Taylor, L., & Tanino, K. (2019). Heating demand and economic feasibility analysis for year-round vegetable production in Canadian Prairies greenhouses. *Information Processing in Agriculture*, *6*(1), 81–90. https://doi.org/10.1016/j.inpa.2018.08.005

Ahmadi, A., Daccache, A., Snyder, R. L., & Suvočarev, K. (2022). Meteorological driving forces of reference evapotranspiration and their trends in California. *Science of The Total Environment*, *849*, 157823. https://doi.org/10.1016/j.scitotenv.2022.157823

Ahmadi, A., Kazemi, M. H., Daccache, A., & Snyder, R. L. (2024). SolarET: A generalizable machine learning approach to estimate reference evapotranspiration from solar radiation. *Agricultural Water Management*, *295*, 108779. https://doi.org/10.1016/j.agwat.2024.108779

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56*. *300*(9).

Allen, R. G., Smith, M., Pereira, L. S., Raes, D., & Wright, J. L. (2001). Revised FAO Procedures for Calculating Evapotranspiration: Irrigation and Drainage Paper No. 56 with Testing in Idaho. *Watershed Management and Operations Management 2000*, 1–10. https://doi.org/10.1061/40499(2000)125

Almeida, A. P., Liberalesso, T., Silva, C. M., & Sousa, V. (2021). Dynamic modelling of rainwater harvesting with green roofs in university buildings. *Journal of Cleaner Production*, *312*, 127655. https://doi.org/10.1016/j.jclepro.2021.127655

Alnaimat, F. (2022). Heat transfer analysis of air-mist evaporative cooling in heat sink. *International Journal of Thermofluids*, *14*, 100145. https://doi.org/10.1016/j.ijft.2022.100145

AlShrouf, Ali. (2017). *Hydroponics, aeroponic and aquaponic as compared with conventional farming*. *27*(1), 247–255.

Araya, N. A., Moremi, M. G., Mokgehle, S., Mofokeng, M. M., Malaka, M. J., Makgato, M., Araya, H. T., & Mampholo, B. (2024). Sustainable Soilless Recirculating Hydroponics for Productive Use of Marginal Lands: A South African Context. In A. D. Nciizah, A. Roopnarain, B. Ndaba, & M. E. Malobane (Eds.), *The Marginal Soils of Africa* (pp. 279–306). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-55185-7_15

Asseng, S., Guarin, J. R., Raman, M., Monje, O., Kiss, G., Despommier, D. D., Meggers, F. M., & Gauthier, P. P. G. (2020). Wheat yield potential in controlled-environment vertical farms. *Proceedings of the National Academy of Sciences*, *117*(32), 19131–19135. https://doi.org/10.1073/pnas.2002655117

Atekwana, E., Atekwana, E., Rowe, R., Werkemajr, D., & Legall, F. (2004). The relationship of total dissolved solids measurements to bulk electrical conductivity in an aquifer contaminated with hydrocarbon. *Journal of Applied Geophysics*, *56*(4), 281–294. https://doi.org/10.1016/S0926-9851(04)00057-6

Banakar, A., Montazeri, M., Ghobadian, B., Pasdarshahri, H., & Kamrani, F. (2021). Energy analysis and assessing heating and cooling demands of closed greenhouse in Iran. *Thermal Science and Engineering Progress*, *25*, 101042. https://doi.org/10.1016/j.tsep.2021.101042

Barbosa, G., Gadelha, F., Kublik, N., Proctor, A., Reichelm, L., Weissinger, E., Wohlleb, G., & Halden, R. (2015). Comparison of Land, Water, and Energy Requirements of Lettuce Grown Using Hydroponic vs. Conventional Agricultural Methods. *International Journal of Environmental Research and Public Health*, *12*(6), 6879–6891. https://doi.org/10.3390/ijerph120606879

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 180214. https://doi.org/10.1038/sdata.2018.214

Benke, K., & Tomkins, B. (2017). Future food-production systems: Vertical farming and controlled-environment agriculture. *Sustainability: Science, Practice and Policy*, *13*(1), 13–26. https://doi.org/10.1080/15487733.2017.1394054

Brækken, A., Sannan, S., Jerca, I. O., & Bădulescu, L. A. (2023). Assessment of heating and cooling demands of a glass greenhouse in Bucharest, Romania. *Thermal Science and Engineering Progress*, *41*, 101830. https://doi.org/10.1016/j.tsep.2023.101830

Brouwer, J., Hoevenaars, J. P. M., & van Bosch, B. E. (1992). *Scheme Irrigation Water Needs and Supply: A Manual* (Vols. 1–6). Food & Agriculture Org.

Cai, W., McPhaden, M. J., Grimm, A. M., Rodrigues, R. R., Taschetto, A. S., Garreaud, R. D., ... & Vera, C. (2020). *Climate impacts of the El Niño–southern oscillation on South America. Nature Reviews Earth & Environment, 1*(4), 215-231.

Çam, N. Y., Ezan, M. A., & Biçer, Y. (2024). An integrated transient model in TRNSYS for thermal management of the tomato growth process in a greenhouse with a PV-aided heat pump-assisted HVAC system. *International Journal of Thermofluids*, *22*, 100592. https://doi.org/10.1016/j.ijft.2024.100592

Carmassi, G., Incrocci, L., Maggini, R., Malorgio, F., Tognoni, F., & Pardossi, A. (2007). An aggregated model for water requirements of greenhouse tomato grown in closed rockwool culture with saline water. *Agricultural Water Management*, *88*(1–3), 73–82. https://doi.org/10.1016/j.agwat.2006.10.002

Carotti, L., Graamans, L., Puksic, F., Butturini, M., Meinen, E., Heuvelink, E., & Stanghellini, C. (2021). Plant Factories Are Heating Up: Hunting for the Best Combination of Light Intensity, Air Temperature and Root-Zone Temperature in Lettuce Production. *Frontiers in Plant Science*, *11*, 592171. https://doi.org/10.3389/fpls.2020.592171

Casey, L., Freeman, B., Francis, K., Brychkova, G., McKeown, P., Spillane, C., Bezrukov, A., Zaworotko, M., & Styles, D. (2022). Comparative environmental footprints of lettuce supplied by

hydroponic controlled-environment agriculture and field-based supply chains. *Journal of Cleaner Production*, *369*, 133214. https://doi.org/10.1016/j.jclepro.2022.133214

Castillo, Sánchez-Del, Esaú del C. Moreno-Pérez, Joel Pineda-Pineda, José M. Osuna, Juan E. Rodríguez-Pérez, & Tomás Osuna-Encino. (2014). *Hydroponic tomato (Solatium lycopersicum L.) production with and without recirculation of nutrient solution. 48*(2), 185–197.

Çaylı, A., Akyüz, A., Üstün, S., & Yeter, B. (2021). Efficiency of two different types of evaporative cooling systems in broiler houses in Eastern Mediterranean climate conditions. *Thermal Science and Engineering Progress*, *22*, 100844. https://doi.org/10.1016/j.tsep.2021.100844

Choab, N., Allouhi, A., Maakoul, A. E., Kousksou, T., Saadeddine, S., & Jamil, A. (2021). Effect of Greenhouse Design Parameters on the Heating and Cooling Requirement of Greenhouses in Moroccan Climatic Conditions. *IEEE Access*, *9*, 2986–3003. https://doi.org/10.1109/ACCESS.2020.3047851

Copernicus Climate Change Service. (2019a). *Agrometeorological indicators from 1979 up to 2019 derived from reanalysis* [Dataset]. ECMWF. https://doi.org/10.24381/CDS.6C68C9BB

Copernicus Climate Change Service. (2019b). *ERA5-Land hourly data from 1950 to present* [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/CDS.E2161BAC

Costantino, A., Comba, L., Sicardi, G., Bariani, M., & Fabrizio, E. (2021). Energy performance and climate control in mechanically ventilated greenhouses: A dynamic modelling-based assessment and investigation. *Applied Energy*, *288*, 116583. https://doi.org/10.1016/j.apenergy.2021.116583

Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F. N., & Leip, A. (2021). Food systems are responsible for a third of global anthropogenic GHG emissions. *Nature Food*, *2*(3), 198–209. https://doi.org/10.1038/s43016-021-00225-9

Cristiano, E., Deidda, R., & Viola, F. (2021). The role of green roofs in urban Water-Energy-Food-Ecosystem nexus: A review. *Science of The Total Environment*, *756*, 143876. https://doi.org/10.1016/j.scitotenv.2020.143876

Dimitropoulou, A.-M. N., Maroulis, V. Z., & Giannini, E. N. (2023). A Simple and Effective Model for Predicting the Thermal Energy Requirements of Greenhouses in Europe. *Energies*, *16*(19), 6788. https://doi.org/10.3390/en16196788

D'Odorico, P., Davis, K. F., Rosa, L., Carr, J. A., Chiarelli, D., Dell'Angelo, J., Gephart, J., MacDonald, G. K., Seekell, D. A., Suweis, S., & Rulli, M. C. (2018). The Global Food-Energy-Water Nexus. *Reviews of Geophysics*, *56*(3), 456–531. https://doi.org/10.1029/2017RG000591

Elliott, J., Deryng, D., Müller, C., Frieler, K., Konzmann, M., Gerten, D., Glotter, M., Flörke, M., Wada, Y., Best, N., Eisner, S., Fekete, B. M., Folberth, C., Foster, I., Gosling, S. N., Haddeland, I., Khabarov, N., Ludwig, F., Masaki, Y., … Wisser, D. (2014). Constraints and potentials of future irrigation water availability on agricultural production under climate change. *Proceedings of the National Academy of Sciences*, *111*(9), 3239–3244. https://doi.org/10.1073/pnas.1222474110

Fathidarehnijeh, E., Nadeem, M., Cheema, M., Thomas, R., Krishnapillai, M., & Galagedara, L. (2023). Current perspective on nutrient solution management strategies to improve the nutrient and water use efficiency in hydroponic systems. *Canadian Journal of Plant Science*, cjps-2023-0034. https://doi.org/10.1139/cjps-2023-0034

Fernández, E. F., Villar-Fernández, A., Montes-Romero, J., Ruiz-Torres, L., Rodrigo, P. M., Manzaneda, A. J., & Almonacid, F. (2022). Global energy assessment of the potential of photovoltaics for greenhouse farming. *Applied Energy*, *309*, 118474. https://doi.org/10.1016/j.apenergy.2021.118474

Fernández, M. D., Bonachela, S., Orgaz, F., Thompson, R., López, J. C., Granados, M. R., Gallardo, M., & Fereres, E. (2010). Measurement and estimation of plastic greenhouse reference evapotranspiration in a Mediterranean climate. *Irrigation Science*, *28*(6), 497–509. https://doi.org/10.1007/s00271-010-0210-z

Frans, S., & Juha-Matti, K. (2021). Reduction of the Climate Impact of Finnish Greenhouse Vegetables Achieved by Energy Acquisitions between 2004 and 2017. *Journal of Horticultural Science and Research*, *4*(1). https://doi.org/10.36959/745/408

Fu, Z., Ciais, P., Wigneron, J.-P., Gentine, P., Feldman, A. F., Makowski, D., Viovy, N., Kemanian, A. R., Goll, D. S., Stoy, P. C., Prentice, I. C., Yakir, D., Liu, L., Ma, H., Li, X., Huang, Y., Yu, K., Zhu, P., Li, X., … Smith, W. K. (2024). Global critical soil moisture thresholds of plant water stress. *Nature Communications*, *15*(1), 4826. https://doi.org/10.1038/s41467-024-49244-7

Fuchs, M., Dayan, E., & Presnov, E. (2006). Evaporative cooling of a ventilated greenhouse rose crop. *Agricultural and Forest Meteorology*, *138*(1–4), 203–215. https://doi.org/10.1016/j.agrformet.2006.05.002

Garreaud, R. D., Boisier, J. P., Rondanelli, R., Montecinos, A., Sepúlveda, H. H., & Veloso-Aguila, D. (2020). The Central Chile Mega Drought (2010–2018): A climate dynamics

perspective. *International Journal of Climatology*, *40*(1), 421–439. https://doi.org/10.1002/joc.6219

Giri, N. C., & Mohanty, R. C. (2022). Agrivoltaic system: Experimental analysis for enhancing land productivity and revenue of farmers. *Energy for Sustainable Development*, *70*, 54–61. https://doi.org/10.1016/j.esd.2022.07.003

Golzar, F., Heeren, N., Hellweg, S., & Roshandel, R. (2018). A novel integrated framework to evaluate greenhouse energy demand and crop yield production. *Renewable and Sustainable Energy Reviews*, *96*, 487–501. https://doi.org/10.1016/j.rser.2018.06.046

Goodman, W., & Minner, J. (2019). Will the urban agricultural revolution be vertical and soilless? A case study of controlled environment agriculture in New York City. *Land Use Policy*, *83*, 160–173. https://doi.org/10.1016/j.landusepol.2018.12.038

Graamans, L., Baeza, E., Van Den Dobbelsteen, A., Tsafaras, I., & Stanghellini, C. (2018). Plant factories versus greenhouses: Comparison of resource use efficiency. *Agricultural Systems*, *160*, 31–43. https://doi.org/10.1016/j.agsy.2017.11.003

Graamans, L., van den Dobbelsteen, A., Meinen, E., & Stanghellini, C. (2017). Plant factories; crop transpiration and energy balance. *Agricultural Systems*, *153*, 138–147. https://doi.org/10.1016/j.agsy.2017.01.003

Grogan, D., Frolking, S., Wisser, D., Prusevich, A., & Glidden, S. (2022). Global gridded crop harvested area, production, yield, and monthly physical area data circa 2015. *Scientific Data*, *9*(1), 15. https://doi.org/10.1038/s41597-021-01115-2

Hargreaves & Samani. (1985). Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*, *1*(2), 96–99. https://doi.org/10.13031/2013.26773

Hoekstra, A. Y. (2019). Green-blue water accounting in a soil water balance. *Advances in Water Resources*, *129*, 112–117. https://doi.org/10.1016/j.advwatres.2019.05.012

Hopmans, J. W., Qureshi, A. S., Kisekka, I., Munns, R., Grattan, S. R., Rengasamy, P., Ben-Gal, A., Assouline, S., Javaux, M., Minhas, P. S., Raats, P. A. C., Skaggs, T. H., Wang, G., De Jong Van Lier, Q., Jiao, H., Lavado, R. S., Lazarovitch, N., Li, B., & Taleisnik, E. (2021). Critical knowledge gaps and research priorities in global soil salinity. In *Advances in Agronomy* (Vol. 169, pp. 1–191). Elsevier. https://doi.org/10.1016/bs.agron.2021.03.001

Hou, Z., Wang, Z., Sun, Z., Gao, P., Hu, Y., & Fan, J. (2024). Annual energy performance simulation of ground source heat pump in a solar greenhouse. *Case Studies in Thermal Engineering*, *59*, 104556. https://doi.org/10.1016/j.csite.2024.104556

Incrocci, L., Thompson, R. B., Fernandez-Fernandez, M. D., De Pascale, S., Pardossi, A., Stanghellini, C., Rouphael, Y., & Gallardo, M. (2020). Irrigation management of European greenhouse vegetable crops. *Agricultural Water Management*, *242*, 106393. https://doi.org/10.1016/j.agwat.2020.106393

Jasechko, S., Perrone, D., Seybold, H., Fan, Y., & Kirchner, J. W. (2020). Groundwater level observations in 250,000 coastal US wells reveal scope of potential seawater intrusion. *Nature Communications*, *11*(1), 3229. https://doi.org/10.1038/s41467-020-17038-2

JISAO (2024). *Elevation data in netCDF* [Dataset]. http://research.jisao.washington.edu/data_sets/elevation/

Karimzadeh, S., Hartman, S., Chiarelli, D. D., Rulli, M. C., & D'Odorico, P. (2024). The tradeoff between water savings and salinization prevention in dryland irrigation. *Advances in Water Resources*, *183*, 104604. https://doi.org/10.1016/j.advwatres.2023.104604

Karimzadeh, S., Daccache, A., Rulli, M. C., & Ahamed, M. S. (2025). Global water-nutrient-salinity-energy nexus in lettuce production: from open-field irrigation to closed-loop hydroponics in greenhouses. *Journal of Agriculture and Food Research*, 101935.

Köchy, M., Hiederer, R., & Freibauer, A. (2015). Global distribution of soil organic carbon – Part 1: Masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world. *SOIL*, *1*(1), 351–365. https://doi.org/10.5194/soil-1-351-2015

Kummu, M., Fader, M., Gerten, D., Guillaume, J. H., Jalava, M., Jägermeyr, J., Pfister, S., Porkka, M., Siebert, S., & Varis, O. (2017). Bringing it all together: Linking measures to secure nations' food supply. *Current Opinion in Environmental Sustainability*, *29*, 98–117. https://doi.org/10.1016/j.cosust.2018.01.006

Kwon, M. J., Hwang, Y., Lee, J., Ham, B., Rahman, A., Azam, H., & Yang, J.-S. (2021). Waste nutrient solutions from full-scale open hydroponic cultivation: Dynamics of effluent quality and removal of nitrogen and phosphorus using a pilot-scale sequencing batch reactor. *Journal of Environmental Management*, *281*, 111893. https://doi.org/10.1016/j.jenvman.2020.111893

Lavigueur, Clémence. (2022). *Improving greenhouse modelling in building performance simulation tools* [Master's thesis]. Ecole Polytechnique, Montreal.

Libardi, L. G. P., De Faria, R. T., Dalri, A. B., De Souza Rolim, G., Palaretti, L. F., Coelho, A. P., & Martins, I. P. (2019). Evapotranspiration and crop coefficient (Kc) of pre-sprouted sugarcane

plantlets for greenhouse irrigation management. *Agricultural Water Management*, *212*, 306–316. https://doi.org/10.1016/j.agwat.2018.09.003

Liu, W., Liu, X., Yang, H., Ciais, P., & Wada, Y. (2022). Global Water Scarcity Assessment Incorporating Green Water in Crop Production. *Water Resources Research*, *58*(1), e2020WR028570. https://doi.org/10.1029/2020WR028570

Londra, P. A., Kotsatos, I.-E., Theotokatos, N., Theocharis, A. T., & Dercas, N. (2021). Reliability Analysis of Rainwater Harvesting Tanks for Irrigation Use in Greenhouse Agriculture. *Hydrology*, *8*(3), 132. https://doi.org/10.3390/hydrology8030132

Magán, J. J., Gallardo, M., Fernández, M. D., García, M. L., Granados, M. R., Padilla, F. M., & Thompson, R. B. (2019). Showcasing a fertigation management strategy for increasing water and nitrogen use efficiency in soil-grown vegetable crops in the FERTINNOWA project. *Acta Horticulturae*, *1253*, 17–24. https://doi.org/10.17660/ActaHortic.2019.1253.3

Mahendra, N., Nagaraju, C., Chowdary, J. S., Ashok, K., & Singh, M. (2023). A curious case of the Indian Summer Monsoon 2020: The influence of Barotropic Rossby Waves and the monsoon depressions. *Atmospheric Research*, *281*, 106476. https://doi.org/10.1016/j.atmosres.2022.106476

Majid, M., Khan, J. N., Ahmad Shah, Q. M., Masoodi, K. Z., Afroza, B., & Parvaze, S. (2021). Evaluation of hydroponic systems for the cultivation of Lettuce (Lactuca sativa L., var. Longifolia) and comparison with protected soil-based cultivation. *Agricultural Water Management*, *245*, 106572. https://doi.org/10.1016/j.agwat.2020.106572

Martinez, C. J., & Thepadia, M. (2010). Estimating Reference Evapotranspiration with Minimum Data in Florida. *Journal of Irrigation and Drainage Engineering*, *136*(7), 494–501. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000214

Mekonnen, M. M., & Hoekstra, A. Y. (2020). Sustainability of the blue water footprint of crops. *Advances in Water Resources*, *143*, 103679. https://doi.org/10.1016/j.advwatres.2020.103679

Miller, A., Adhikari, R., & Nemali, K. (2020). Recycling Nutrient Solution Can Reduce Growth Due to Nutrient Deficiencies in Hydroponic Production. *Frontiers in Plant Science*, *11*, 607643. https://doi.org/10.3389/fpls.2020.607643

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(1), 2007GB002947. https://doi.org/10.1029/2007GB002947

Montero, J. I., Antón, A., Muñoz, P., & Lorenzo, P. (2001). Transpiration from geranium grown under high temperatures and low humidities in greenhouses. *Agricultural and Forest Meteorology*, *107*(4), 323–332. https://doi.org/10.1016/S0168-1923(01)00215-5

Nasrollahi, H., Ahmadi, F., Ebadollahi, M., Nobar, S. N., & Amidpour, M. (2021). The greenhouse technology in different climate conditions: A comprehensive energy-saving analysis. *Sustainable Energy Technologies and Assessments, 47*, 101455.

Opeyemi Ogunlowo, Q., Denen Akpenpuun, T., Ho Na, W., Aderemi Adesanya, M., Rabiu, A., Dutta, P., Kim, H.-T., & Lee, H.-W. (2023). Simulation of greenhouse energy and strawberry (Seolhyang sp.) yield using TRNSYS DVBES: A base case. *Solar Energy*, *266*, 112196. https://doi.org/10.1016/j.solener.2023.112196

Paucek, I., Durante, E., Pennisi, G., Quaini, S., Gianquinto, G., & Orsini, F. (2023). A methodological tool for sustainability and feasibility assessment of indoor vertical farming with artificial lighting in Africa. *Scientific Reports*, *13*(1), 2109. https://doi.org/10.1038/s41598-023-29027-8

Perri, S., Molini, A., Hedin, L. O., & Porporato, A. (2022). Contrasting effects of aridity and seasonality on global salinization. *Nature Geoscience*, *15*(5), 375–381. https://doi.org/10.1038/s41561-022-00931-4

Postel, S. L., Daily, G. C., & Ehrlich, P. R. (1996). Human Appropriation of Renewable Fresh Water. *Science*, *271*(5250), 785–788. https://doi.org/10.1126/science.271.5250.785

Prieto, J., Ajnannadhif, R. M., Fernández-del Olmo, P., & Coronas, A. (2023). Integration of a heating and cooling system driven by solar thermal energy and biomass for a greenhouse in Mediterranean climates. *Applied Thermal Engineering*, *221*, 119928. https://doi.org/10.1016/j.applthermaleng.2022.119928

Rhoades, J. D. (2018). Salinity: Electrical Conductivity and Total Dissolved Solids. In D. L. Sparks, A. L. Page, P. A. Helmke, R. H. Loeppert, P. N. Soltanpour, M. A. Tabatabai, C. T. Johnston, & M. E. Sumner (Eds.), *SSSA Book Series* (pp. 417–435). Soil Science Society of America, American Society of Agronomy. https://doi.org/10.2136/sssabookser5.3.c14

Rhoades, J. D., Kandiah, A., & Mashali, A. M. (1992). *The use of saline waters for crop production-FAO irrigation and drainage paper 48*.

Rosa, L., & Gabrielli, P. (2023). Energy and food security implications of transitioning synthetic nitrogen fertilizers to net-zero emissions. *Environmental Research Letters*, *18*(1), 014008. https://doi.org/10.1088/1748-9326/aca815

Rosa, L., Rulli, M. C., Ali, S., Chiarelli, D. D., Dell'Angelo, J., Mueller, N. D., Scheidel, A., Siciliano, G., & D'Odorico, P. (2021). Energy implications of the 21st century agrarian transition. *Nature Communications*, *12*(1), 2319. https://doi.org/10.1038/s41467-021-22581-7

Runyan, C. W., & D'Odorico, P. (2010). Ecohydrological feedbacks between salt accumulation and vegetation dynamics: Role of vegetation-groundwater interactions. *Water Resources Research*, *46*(11), 2010WR009464. https://doi.org/10.1029/2010WR009464

Sabeh, N. C., Giacomelli, G. A., & Kubota, C. (2011). Water Use in a Greenhouse in a Semi-Arid Climate. *Transactions of the ASABE*, *54*(3), 1069–1077. https://doi.org/10.13031/2013.37098

Sajid, M. U., & Bicer, Y. (2021). Comparative life cycle cost analysis of various solar energy-based integrated systems for self-sufficient greenhouses. *Sustainable Production and Consumption*, *27*, 141–156. https://doi.org/10.1016/j.spc.2020.10.025

Schmitt, R. J. P., Rosa, L., & Daily, G. C. (2022). Global expansion of sustainable irrigation limited by water storage. *Proceedings of the National Academy of Sciences*, *119*(47), e2214291119. https://doi.org/10.1073/pnas.2214291119

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*(1), 249–263. https://doi.org/10.1002/2013MS000293

Silva, C. M., Sousa, V., & Meireles, I. (2017). Water Savings in Rainwater Harvesting Systems in Portugal: Influence of Weather and Type of Building. In *Frontiers in Civil Engineering* (pp. 196–250). BENTHAM SCIENCE PUBLISHERS. https://doi.org/10.2174/9781681084831117020008

Soufi, H. R., Roosta, H. R., & Hamidpour, M. (2023). The plant growth, water and electricity consumption, and nutrients uptake are influenced by different light spectra and nutrition of lettuce. *Scientific Reports*, *13*(1), 20766. https://doi.org/10.1038/s41598-023-48284-1

Soussi, M., Chaibi, M. T., Buchholz, M., & Saghrouni, Z. (2022). Comprehensive Review on Climate Control and Cooling Systems in Greenhouses under Hot and Arid Conditions. *Agronomy*, *12*(3), 626. https://doi.org/10.3390/agronomy12030626

Souza, S. V., Gimenes, R. M. T., & Binotto, E. (2019). Economic viability for deploying hydroponic system in emerging countries: A differentiated risk adjustment proposal. *Land Use Policy*, *83*, 357–369. https://doi.org/10.1016/j.landusepol.2019.02.020

Stout, D. T., Walsh, T. C., & Burian, S. J. (2017). Ecosystem services from rainwater harvesting in India. *Urban Water Journal*, *14*(6), 561–573. https://doi.org/10.1080/1573062X.2015.1049280

Tabari, H. (2010). Evaluation of Reference Crop Evapotranspiration Equations in Various Climates. *Water Resources Management*, *24*(10), 2311–2337. https://doi.org/10.1007/s11269-009-9553-8

Tawalbeh, M., Aljaghoub, H., Alami, A. H., & Olabi, A. G. (2023). Selection criteria of cooling technologies for sustainable greenhouses: A comprehensive review. *Thermal Science and Engineering Progress*, *38*, 101666. https://doi.org/10.1016/j.tsep.2023.101666

Todorovic, M., Karic, B., & Pereira, L. S. (2013). Reference evapotranspiration estimate with limited weather data across a range of Mediterranean climates. *Journal of Hydrology*, *481*, 166–176. https://doi.org/10.1016/j.jhydrol.2012.12.034

Trabucco, A., & Zomer, R. (2022). *Global Aridity Index and Potential Evapotranspiration (ET0) Database: Version 3* (p. 6086879943 Bytes) [Dataset]. figshare. https://doi.org/10.6084/M9.FIGSHARE.7504448.V5

Trajkovic, S. (2007). Hargreaves versus Penman-Monteith under Humid Conditions. *Journal of Irrigation and Drainage Engineering*, *133*(1), 38–42. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:1(38)

Tsafaras, I., Campen, J. B., Stanghellini, C., De Zwart, H. F., Voogt, W., Scheffers, K., Harbi, A. A., & Assaf, K. A. (2021). Intelligent greenhouse design decreases water use for evaporative cooling in arid regions. *Agricultural Water Management*, *250*, 106807. https://doi.org/10.1016/j.agwat.2021.106807

University of California, Salinity measurement. (2023). Salinity Measurement and Unit Conversion. https://ucanr.edu/sites/Salinity/Salinity_Management/Salinity_Basics/Salinity_measurement_and_unit_conversions/

Van Ginkel, S. W., Igou, T., & Chen, Y. (2017). Energy, water and nutrient impacts of California-grown vegetables compared to controlled environmental agriculture systems in Atlanta, GA. *Resources, Conservation and Recycling*, *122*, 319–325. https://doi.org/10.1016/j.resconrec.2017.03.003

Van 'T Ooster, A., Van Ieperen, W., & Kalaitzoglou, P. (2012). Model study on applicability of a semi closed greenhouse concept in Almeria: effects on greenhouse climate, crop yield and resource use efficiency. *Acta Horticulturae*, *927*, 51–58. https://doi.org/10.17660/ActaHortic.2012.927.4

Villarreal-Guerrero, F., Kacira, M., Fitz-Rodríguez, E., Kubota, C., Giacomelli, G. A., Linker, R., & Arbel, A. (2012). Comparison of three evapotranspiration models for a greenhouse cooling strategy with natural ventilation and variable high pressure fogging. *Scientia Horticulturae*, *134*, 210–221. https://doi.org/10.1016/j.scienta.2011.10.016

Weidner, T., Yang, A., Forster, F., & Hamm, M. W. (2022). Regional conditions shape the food–energy–land nexus of low-carbon indoor farming. *Nature Food*, *3*(3), 206–216. https://doi.org/10.1038/s43016-022-00461-7

Weidner, T., Yang, A., & Hamm, M. W. (2021). Energy optimisation of plant factories and greenhouses for different climatic conditions. *Energy Conversion and Management*, *243*, 114336. https://doi.org/10.1016/j.enconman.2021.114336

Xu, J., Li, Y., Wang, R. Z., Liu, W., & Zhou, P. (2015). Experimental performance of evaporative cooling pad systems in greenhouses in humid subtropical climates. *Applied Energy*, *138*, 291–301. https://doi.org/10.1016/j.apenergy.2014.10.061

Yang, P., Wu, L., Cheng, M., Fan, J., Li, S., Wang, H., & Qian, L. (2023). Review on Drip Irrigation: Impact on Crop Yield, Quality, and Water Productivity in China. *Water*, *15*(9), 1733. https://doi.org/10.3390/w15091733

Zhangzhong, L., Gao, H., Zheng, W., Wu, J., Li, J., & Wang, D. (2023). Development of an evapotranspiration estimation method for lettuce via mobile phones using machine vision: Proof of concept. *Agricultural Water Management*, *275*, 108003. https://doi.org/10.1016/j.agwat.2022.108003

Zhuang, Y., Lu, N., Kikuchi, M., Takagaki, M., & Tamashiro, T. (2024). Productivity potential and economic feasibility of small-sized tomato production in plant factories with artificial lighting: A comparative study with high-tech greenhouse production. *Journal of Cleaner Production*, *470*, 143171. https://doi.org/10.1016/j.jclepro.2024.143171

Chapter 3: Machine learning-based fault detection and diagnosis of electrical conductivity and pH sensors in hydroponic systems

Reference: Karimzadeh, S., Li, Z., & Ahamed, M. S. (2025). Machine learning-based fault detection and diagnosis of electrical conductivity and pH sensors in hydroponic systems. *Computers and Electronics in Agriculture*, *237*, 110544.

3.1. Abstract

Fault detection and diagnosis (FDD) is critical for the precise operation of many complex systems. This study evaluates the effectiveness of data-driven FDD for electrical conductivity (EC) and the potential of hydrogen (pH) sensors in a closed-loop nutrient film technique (NFT) hydroponic system. Five machine learning models—random forest (RF), artificial neural networks (ANN), long short-term memory (LSTM), k-nearest neighbors (KNN), and support vector machine (SVM)—were assessed for their ability to (i) detect fault readings and (ii) diagnose five fault types: bias, drift, precision degradation, spike, and stuck in EC and pH sensors. RF demonstrated superior accuracy in fault detection, achieving 93.7% for EC and 96.5% for pH sensors. In fault diagnosis, RF also consistently outperformed other models for the EC sensor, achieving accuracies reaching up to 100% for spike and 97.1% for bias. In contrast, the pH sensor fault diagnosis showed varied performance: SVM excelled in bias with 100% accuracy, while RF led in normal, drift, and precision degradation, with accuracies up to 92.1%. Additionally, we introduce a novel, sensor-independent fault detection framework for hydroponic systems, leveraging machine learning models to predict EC using only environmental and operational variables. This approach departs from conventional sensor-based diagnostics by enabling fault

detection without relying on the target sensor's readings. Among the models tested, ANN achieved a fault detection accuracy of 93.2%, markedly outperforming support vector regression (SVR), which reached only 49.4%. Overall, the advancement in FDD for hydroponics underscores the potential of machine learning to enhance the robustness and sustainability of the precision operation of hydroponic production.

Keywords: Fault Detection and Diagnosis; Machine Learning; Hydroponic Systems; Nutrient Management; Precision Agriculture

3.2. Introduction

Sensors are fundamental to precision indoor farming and the intelligent control of hydroponic systems. These devices are often placed in challenging conditions (extreme heat or underwater), making them susceptible to malfunctions and damage. Such issues can compromise the accuracy, stability, and reliability of operations (Yin & Hou, 2016). In the event of a sensor failure, operational efficiency could be negatively impacted, potentially leading to severe repercussions. Given the system's dependence on accurate data for optimal performance, any anomalies or corruption in the sensor data could induce detrimental effects, particularly in precisely controlling crop production and resource use efficiency. Specifically, the most critical aspect is the swift detection of emerging faults—identifying issues at the earliest stage of slow development and promptly pinpointing the problem. Implementing a real-time identification system is essential to facilitate such rapid detection instead of relying on experience-based monitoring.

Fault detection identifies deviations from normal system behavior, while fault diagnostics determines their underlying causes (Chen et al., 2023). Because this process can be complex, multiple techniques may be required to handle it effectively (Yin & Hou, 2016). Methods for

sensor fault detection and diagnosis (FDD) can be categorized into three types: model-based (i.e., using a dynamic decision function; Willsky & Jones, 1976), knowledge-based (i.e., applying fuzzy logic analysis and historical data alongside expert knowledge; Li et al., 2020), and data-driven (or machine learning-based) approaches (Wu & Zhao, 2018). Data-driven FDD is an evolving technique that does not require a quantitative mathematical model but uses process data directly (Li et al., 2020). This allows the extraction of concealed information by carefully processing and examining large data sets (Jiang & Yin, 2018). The quality of the training dataset predominantly influences the performance of data-driven models. Elements such as temporal resolution, completeness, extensiveness (ensuring data reflects diverse operating conditions), and accuracy (in terms of measurement precision or errors) are central to the success of data-driven FDD strategies (Chen et al., 2023).

Data-driven FDD tool reduces the need for prior experience, adapting to environmental shifts. Artificial neural networks address complex non-linear challenges (Baharvand et al., 2021) and have seen notable growth and application in FDD studies over the past decade (Yin & Hou, 2016). Machine learning-based fault detection is widely applied in various fields including agricultural soil moisture sensor (Ludeña-Choez et al., 2019), water quality monitoring and control in aquaculture (Yang et al., 2017), wastewater treatment plant's sensors (Yang et al., 2025), and leakage detection in pipelines (Banjara et al., 2020). Meanwhile, most FDD research focuses on the heating, ventilation, and air conditioning (HVAC) components of buildings, including air handling unit (AHU), chillers, boilers, AC units, heat pumps, and other subsystems, to identify a range of faults. Numerous supervised techniques—including support vector machine (SVM), support vector regression (SVR), k-nearest neighbor (KNN), decision trees (DT), random forest (RF), gradient boosting (GB), neural networks (NN), deep neural networks (DNN), deep recurrent

neural networks (DRNN), fuzzy deep neural network (FDNN), and long short-term memory (LSTM)—have shown notable effectiveness in various sectors including building and process control and operations. Likewise, unsupervised methods such as clustering algorithms and principal component analysis (PCA) have also been proven for diagnosis (Capozzoli et al., 2015; Dehestani et al., 2013; Ebrahimifakhar et al., 2020; Jiang & Yin, 2018; Lee et al., 2019; Pita et al., 2021; Ragab et al., 2021; Shohet et al., 2020; Taheri et al., 2021; Yan et al., 2020; Yin & Hou, 2016; Zhou et al., 2020).

Leveraging computer vision and machine learning has emerged as a widely researched area in automated anomaly detection. Manual object detection and identification of irregularities in video data are both time-intensive and costly, underscoring the need for autonomous systems capable of efficiently analyzing processes and identifying anomalies across extensive video datasets (Anushka et al., 2021; Roka et al., 2023). Within controlled environment agriculture (CEA), Pita et al. (2021) applied computer vision techniques to address the complexities of fault diagnosis in greenhouse systems, including humidifiers, windows, and temperature and humidity sensors. Four novel architectures—micro output manifold fault detection (MOMFD), output manifold fault detection (OMFD), attention fault long short-term memory (AFLSTM), and fault long short-term memory (FLSTM)—were compared against the established LSTM, with MOMFD achieving the highest validation accuracy (95.56%).

Despite the recently emphasized global potential of closed-loop hydroponic systems in CEA (Karimzadeh et al., 2025), the field remains limited by a scarcity of FDD studies, most of which were conducted nearly two decades ago. For example, Ferentinos et al. (2003a, 2003b) utilized NN for FDD of mechanical faults (e.g., malfunction of the nutrient solution circulation pump), actuators faults (e.g., malfunction of the pH control pump), sensors faults (e.g., failure of pH and

EC sensors), and biological faults (e.g., concerns about plant health) in deep water culture lettuce production. Recently, Shekarian et al. (2024) implemented a framework for real-time fault detection and automated control within greenhouse environments. Their findings revealed that convolutional neural networks (CNN) and decision tree models outperformed linear regression in accurately identifying faults across key environmental parameters, including temperature, humidity, luminosity, and carbon monoxide levels.

Precision water and nutrient management in hydroponics is essential for maximizing yield, improving resource efficiency, and reducing nutrient-driven environmental pollution. Closed-loop hydroponic systems rely on continuous monitoring of EC and pH for proper nutrient management (Bamsey et al., 2012), as these systems lack the buffering capacity of soil-based agriculture. However, sensor faults can result in under- or over-dosing, leading to plant stress, reduced quality, nutrient solution wastage, increased costs, and environmental impact (Bouchaaba et al., 2015). Although data-driven fault detection is well-established across diverse fields, its application to nutrient management in CEA remains underexplored. This study bridges this gap by evaluating advanced data-driven models specifically tailored for hydroponic systems, aiming to enhance precision. By enabling timely FDD in sensors, these models enhance operational reliability, reduce resource waste, and support the sustainability of CEA operations.

This study used experimentally measured data with artificially induced faults to evaluate multiple supervised machine learning approaches for FDD across EC and pH sensors in the context of vertical farming. Specifically, we implemented SVM, KNN, ANN, RF, and LSTM models to identify failures in EC and pH sensors based on their readings (sensor-based FDD). Additionally, we applied the selected models to diagnose five faults (drift, bias, precision, spike, and stuck) in both EC and pH sensors. Furthermore, to establish a novel and sensor-independent fault detection

strategy for hydroponic systems, we employed SVR and ANN to predict EC solely from climate and nutrient dosing variables, eliminating dependence on EC sensor readings and introducing a new paradigm in fault detection for CEA.

3.3. Materials and methods

3.4.1. Cultivation conditions and data collection

3.4.1.1. Experimental setup

The experiment was conducted in an indoor farming system at the University of California, Davis, USA (38.54°N, 121.75°W). The closed-loop nutrient film technique (NFT) hydroponic system (Figure 10f) was used. Seedlings (14 days old) of butterhead lettuce (*Lactuca sativa var. capitata*) were transplanted into the system on August 30th, 2023. These plants were cultivated until their harvest on September 23rd, 2023. The plants were grown on two racks (70 lettuce heads in total), each measuring 1.0 m × 1.4 m, for 24 days (Figure 10). We measured the climatic and rootzone parameters during the growing period. Figure 10 presents the time series data for temperature, relative humidity, nutrient solution volume, and lighting spectrum and intensity details. The temperatures inside the growing tiers varied: 24.9 ± 3.4 °C for the upper trays and 24.4 ± 3.0 °C for the lower trays. The relative humidity was 60 ± 8.5% for the upper tiers and 53 ± 8.3% for the lower tiers. The photoperiod was set to 16 hours, from 4:30 to 20:30. Lighting intensity in the center of the racks was recorded at 250 μmol/m^2.s (LICOR-180, USA), resulting in a daily light integral (DLI) of 14.4 mol/m^2. The plants were illuminated with cool-white, fluorescent lights (6500K T5, VIVOSUN, USA).

The nutrient solution was prepared using industrial water with an EC of 0.23 dS/m and a pH of 8.2. The composition included ions in mg/L as follows: N: 182.0, P: 136.4, K: 234.0, Ca: 120.0,

Mg: 20.4, and S: 25.3, following the recipe by Carotti et al. (2021). Micronutrients were added only once on the transplanting day at the following concentrations in μg/L: B: 0.137, Cu: 0.410, Fe: 0.820, Mn: 0.684, Zn: 1.230. During the experiment, periodic additions of water and nutrients were required due to the limited tank capacity of ~95 liters. Details of these supplementation events are shown in Figure 10d.

3.4.1.2. Data collection

EC and pH sensors were calibrated using a two-point calibration: pH at 4.0 and 10.0, and EC at 1 and 50 dS/m. To emulate practical hydroponic conditions, manual nutrient dosing was performed seven times during the experiment, guided by real-time measurements of EC to maintain optimal nutrient availability for plant growth under standard hydroponic practices. The system's initial pH was adjusted to 5.9. Data on variables such as EC and pH (Vernier, USA), as well as climate variables such as temperature (T) and relative humidity (RH) (DHT22, AOSONG, China), and water level (HC-SR04, MaxBotix, USA) were continuously recorded at 5-second intervals. Observational data were subsequently averaged over 5-minute intervals to reduce noise. These measurements were facilitated by Arduino Mega and Arduino Uno systems (Arduino, Italy). To ensure the reliability of EC and pH measurements, brand-new sensors were employed, and manual inspections of sensor probes were conducted regularly to prevent biofilm buildup. A UV disinfection system (Aquarium Filter, AQUANEAT, USA) was integrated into the nutrient solution to mitigate microbial activity that could compromise sensor performance. To account for fluctuations in water level within the nutrient solution tanks, the EC and pH sensors were mounted on a floating foam platform, ensuring a constant submergence depth and stable measurement conditions throughout the experiment. Furthermore, external EC and pH sensors (Digital pH and TDS Meter Kit, VIVOSUN, USA) were used for routine cross-validation of sensor readings,

ensuring data accuracy throughout the experiment. A detailed schematic of the experimental setup is shown in Figure 10a.

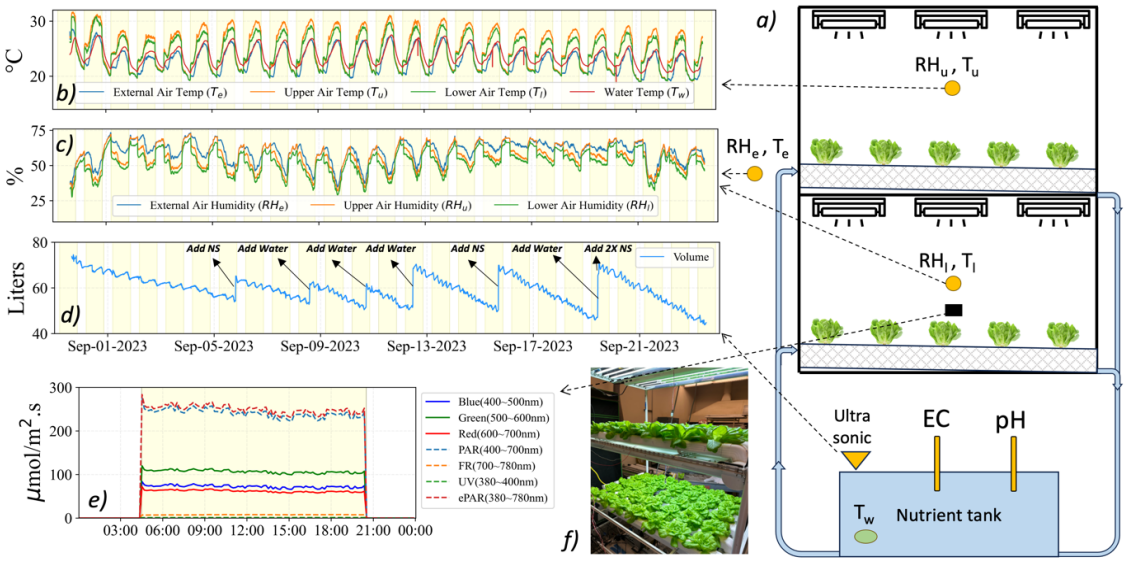


Figure 10. Overview of the experimental setup and sensor installations (a and f), time series data depicting climate variables, temperature (b) and relative humidity (c), nutrient solution volume and dosing (d), and spectral composition analysis of the fluorescent growing lights (e). The highlighted yellow section indicates the 16-hour photoperiod during which plants were exposed to light.

3.4.2. Normal and faulty operation in CEA systems

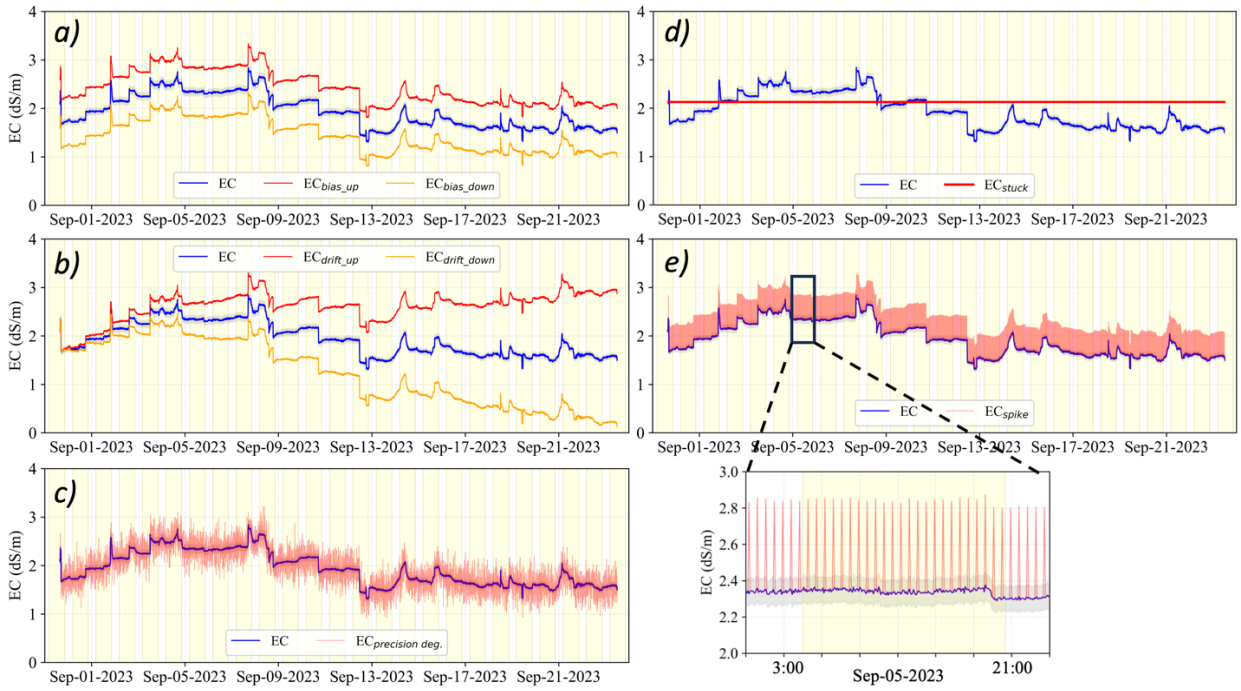
Any divergence from the anticipated performance, commonly referred to as 'faults,' can compromise the system's efficiency and overall productivity. These faults are categorized based on their deviation patterns from normal operation (Jan et al., 2021). Defining a 'normal' state in CEA is particularly challenging in the hydroponic context. This complexity arises from the ongoing interactions between plants, their environment, and the hydroponic solution. Additionally, supplementing the nutrient solution, replenishing water, and introducing nutrients at higher

concentrations further complicate the dynamics, influencing parameters such as EC, pH, and specific ion concentrations. Let $s(t) = v(t) + \eta$ represent the anticipated output of a sensor in normal operation (Table 5). In this equation, $v(t)$ denotes the sensor's output at time t, while η with a distribution of $\eta \sim N(0, \delta_\eta^2)$, signifies noise (dominated by the instrument's precision). The data collected from the described hydroponic system were considered normal data, which was ensured by the calibration of sensors twice before starting the experiment. This study considered five distinct synthetic faults in the sensor data: drift, bias, precision degradation, spike, and stuck (Table 5, Figure 11). The synthetic faults were generated using the methods described in Jan et al. (2021). Table 5 summarizes the mathematical formulations and the typical causes of these sensor faults, along with the parameter values used to synthetically generate faults for EC and pH sensors. These fault values were chosen to ensure non-negative EC and a pH range of 0 to 14.

Table 5. Summary of mathematical formulations, fault parameters, and typical causes of EC and pH sensor faults in the hydroponic nutrient solution (Jan et al., 2021).

Fault type	**Mathematical representation**	**Fault parameters**	**Common causes**
Normal (no fault)	$s(t) = v(t) + \eta$	-	-
Drift	$s(t) = v(t) + \eta + d(t)$ $d(t) = d(t-1) + \vartheta$ $\vartheta = constant$	$EC: \vartheta = \pm 0.0002 \frac{dS}{m}$ pH: $\vartheta = \pm 0.0005$	Material corrosion, environmental damage
Bias	$s(t) = v(t) + \eta + \vartheta$ $\vartheta = constant$	$EC: \vartheta = \pm 0.5 \frac{dS}{m}$ $pH: \vartheta = \pm 1.0$	Short circuits, calibration errors
Precision degradation	$s(t) = v(t) + \eta + \gamma$ $\gamma \sim N(0, \delta_\gamma^2),\ \delta_\gamma^2 \gg \delta_\eta^2$	$EC:\ \delta_\gamma^2 = \pm 0.04 \frac{dS}{m}$ $pH:\ \delta_\gamma^2 = \pm 0.09$	Chemical interference, algae growth on the sensor surface, circuitry issues
Spike	$s(t) = v(t) + \eta + \vartheta(t)$	$EC: \vartheta(t) = +0.5 \frac{dS}{m}\ every\ 40\ minutes$	Vibration or jolt, air bubbles

		$pH: \vartheta(t) = +0.5\ every\ 40\ minutes$	
Stuck	$s(t) = \vartheta$ $\vartheta = constant$	$EC\ and\ pH: \vartheta = First\ observation$	Depleted electrodes, saturation



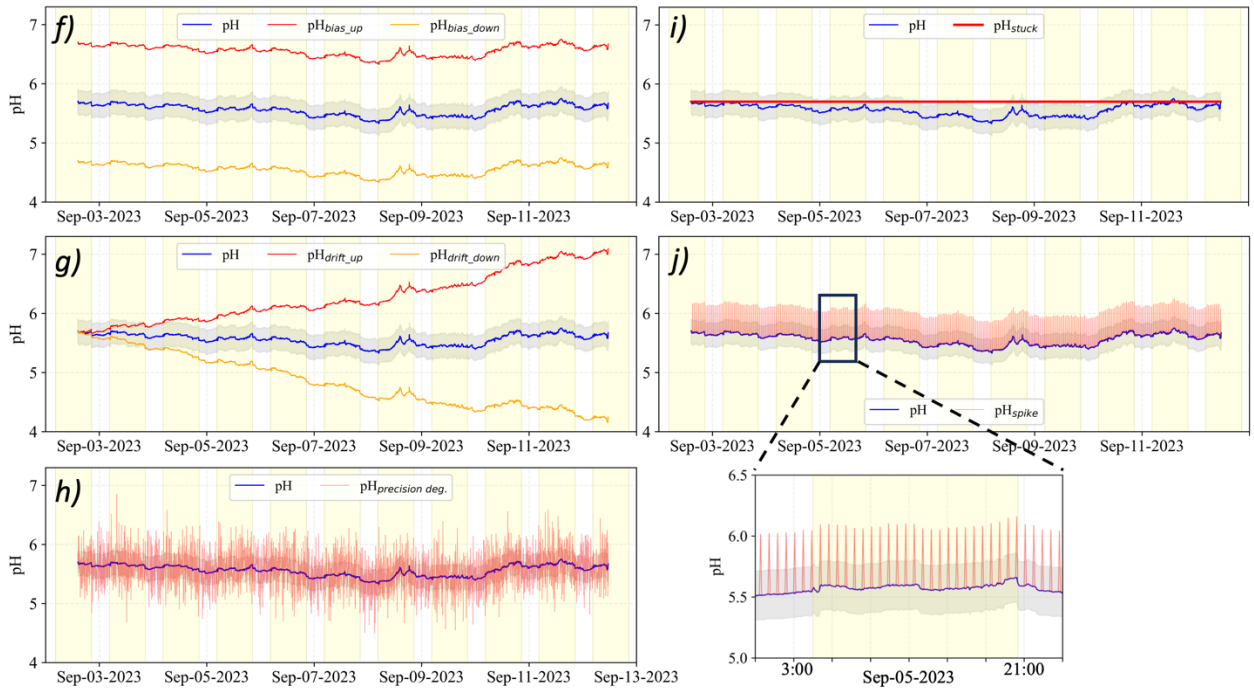


Figure 11. EC and pH sensing in the lettuce hydroponic system, illustrating normal data acquisition and synthetic faults. Subfigures (a) through (e) depict simulated faults for EC: bias, drift, precision degradation, stuck, and spike. Subfigures (f) through (j) illustrate the corresponding faults for pH.

3.4.3. Feature analysis

We employed the Pearson correlation coefficient (Pearson, 1896) to assess the significance and impact of meteorological and dosing signal variables on EC. The Pearson correlation coefficient (Pearson's r) quantifies the linear relationship between two random variables. This method is instrumental in determining how a machine-learning model prioritizes various input variables. Pearson's r ranges from −1 to 1, where 1 indicates a perfect positive linear correlation, −1 indicates a perfect negative linear correlation, and 0 indicates no linear correlation. Mathematically, the Pearson correlation coefficient between two random variables X and Y is expressed in Eq. 3.1:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \qquad (3.1)$$

where $cov(X,Y)$ represents the covariance of X and Y, while σ_X and σ_Y are the standard deviations of X and Y, respectively.

3.4.4. Develop machine learning algorithms

Data-driven FDD techniques require an extensive and inclusive training dataset that encompasses normal and faulty states. In this study, we evaluated five machine learning models for sensor FDD, balancing complexity and performance. SVM and KNN offered robustness with smaller datasets and computational efficiency, whereas RF provided interpretability and handled non-linear relationships. ANN and LSTM excelled at capturing temporal dependencies in sequential data but required higher computational resources. By assessing these trade-offs, we identified models suitable for resource-constrained scenarios and those optimized for capturing complex fault patterns, which enables practical and accurate FDD deployment. The process for sensor-based FDD using machine learning models is depicted in Figure 12.

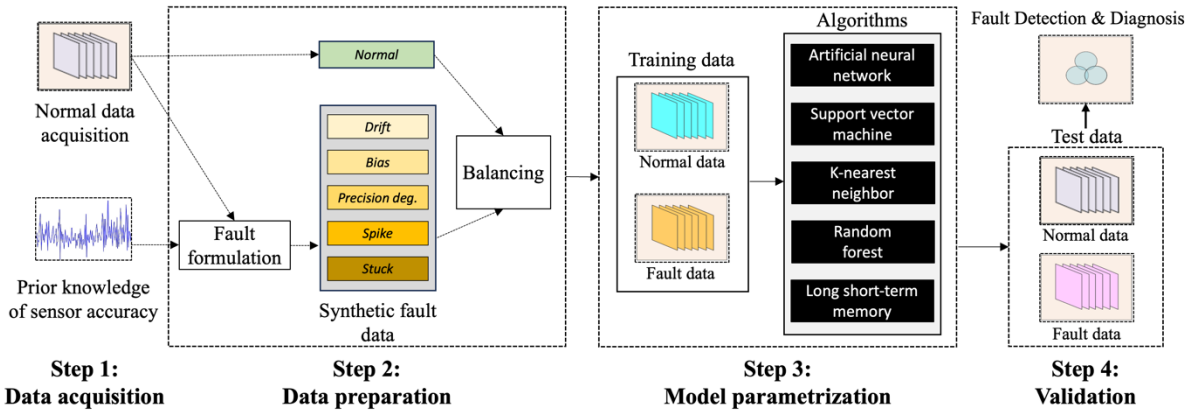


Figure 12. Comprehensive methodological workflow for sensor-based fault detection and diagnosis in hydroponic sensors, consisting of four steps: (1) data acquisition from sensors, (2)

data preparation with synthetic fault generation and class balancing, (3) model training using five machine learning algorithms, and (4) validation for fault detection and diagnosis.

For sensor-independent fault detection, we employed SVR and ANN to predict EC in hydroponic systems. SVR is particularly effective with smaller datasets and offers robust noise tolerance, ensuring reliable predictions and strong generalization. Conversely, ANN demonstrates a superior ability to capture intricate multivariate relationships and scale effectively with larger datasets, making it well-suited for dynamic and high-complexity scenarios.

3.4.4.1. Support vector machine (SVM)

The fundamental concept behind SVM is to use a linear hyperplane to separate the training samples into two distinct classes. This method establishes a linear boundary between the two classes, with the maximum margin is defined as the distance between two hypothetical lines. Yin and Hou (2016) presented a thorough and detailed mathematical explanation for this method. The mathematical representation of the optimized hyperplane is as follows (Haykin, 1999):

$$w^T x + b = 0 \qquad (3.2)$$

where w denotes the weight vector, x stands for an input vector, and b represents the bias. The formulas for the support vectors of each class in the fault detection process are given by:

$$w^T x + b = +1 \quad for\ d_i = +1$$

$$w^T x + b = -1 \quad for\ d_i = -1 \qquad (3.3)$$

where d_i corresponds to the respective class, with d_i being $+1$ for the normal and -1 for the fault class. In the diagnosis step (multi-class classification), the SVM can be applied using a one-versus-rest approach. This method creates m separate classifiers for m-class classification (i.e., m = 6),

representing normal, drift, bias, precision degradation, spike, and stuck. In each m-th classifier, the data associated with the m-th class are treated as the true values, while the remaining m − 1 classes are treated as false values (Jan et al., 2017).

3.4.4.2. k-nearest neighbor (KNN)

This approach relies on memorizing the training data and calculating the Euclidean distance between a query point and its nearest data points. Based on this distance, the k-nearest neighbors (KNN) algorithm identifies the k closest data points in the training dataset, which may belong to different classes. Subsequently, a voting strategy is employed to determine the data category using the labels of the neighboring data points (Song et al., 2020). The value of 'k' is typically chosen as an odd number (e.g., 1, 3, 5, etc.) to ensure a clear majority when determining the votes from the nearest neighbors. The fundamental concept behind KNN is that if a significant proportion of the nearest neighbors share the same class, the data point is classified into that class. Although the algorithm is straightforward in its design, it exhibits strong performance and serves as a significant benchmark method. For distance evaluation, the Euclidean distance metric is a straightforward and readily applicable approach for measuring distances in multidimensional input spaces. The Euclidean distance between two points, 'p' and 'q', represents the length of the line connecting them. In the Cartesian coordinate system, the distance from point 'p' to point 'q' is calculated as follows (Moosavian et al., 2013):

$$d_E = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \qquad (3.4)$$

where p_i and q_i represent two points in Euclidean n-dimensional space.

3.4.4.3. Artificial neural network (ANN)

ANN models are widely employed for tasks such as pattern recognition, classification, regression, and decision-making, mimicking how biological neural networks process information (Yan et al., 2020). This study built the neural network architecture using the sequential model (Tutz, 1991), a widely used framework for constructing feedforward neural networks. It utilizes stacked hidden layers to classify sensor readings within set limits, incorporating accuracy metrics and fault-type labels as inputs. The activation function used in this layer was a rectified linear unit (ReLU), known for its effectiveness in handling non-linear data. The hidden layer played a crucial role in extracting meaningful patterns and relationships from the input data, enhancing the network's ability to learn and generalize from the training data effectively (Panchal et al., 2011). The basic artificial neural network model is as follows:

$$F(x) = \sigma(W^{[4]}\sigma\left(W^{[3]}\sigma\left(W^{[2]}\sigma + b^{[2]}\right) + b^{[3]}\right) + b^{[4]}) \qquad (3.5)$$

where $\{(W^{[2]},b^{[2]}), (W^{[3]},b^{[3]}), (W^{[3]},b^{[3]})\}$, σ is the softmax activation function:

$$\sigma(Z)_i = \frac{e^{Z_i}}{\sum_{j=1}^{k} e^{Z_i}} \qquad (3.6)$$

where $i = 1, \ldots, k,$ and $Z \in R^k$

To address fault detection, the network performs binary classification to determine whether the sensor readings indicate normal operation or the presence of a fault, utilizing an output layer containing two neurons. The output layer comprises six neurons for fault diagnosis, designed to classify different fault labels, categorizing normal, drift, bias, precision degradation, spike, and stuck conditions. By utilizing the softmax function, the network furnishes a probability distribution

across various fault classes, offering insight into the likelihood of each fault type for a given input data point.

3.4.4.4. Random forest (RF)

RF is a versatile and powerful machine learning algorithm that belongs to the ensemble learning method family. It operates by constructing a multitude of decision trees during training and outputting the class based on the majority vote or mean prediction of the individual trees. The fundamental concept behind RF is to combine the predictions from multiple decision trees, which increases the overall result's accuracy and reduces the likelihood of overfitting. This algorithm is particularly effective for a wide range of datasets because it can handle classification tasks with large data sets and missing values (Wang et al., 2018). One of the key strengths of RF is its ability to handle large data sets with high dimensionality. It can manage thousands of input variables without the need for variable deletion, making it a highly scalable option.

3.4.4.5. Long short-term memory (LSTM)

LSTM is a type of recurrent neural network (RNN) architecture used in deep learning (Hochreiter, 1998). Unlike standard feedforward neural networks, LSTMs have feedback connections, making them effective for processing sequences of data (Hochreiter, 1998). They are particularly well-suited for classifying, processing, and making predictions based on time series data. This is because LSTMs can learn long-term dependencies in data, overcoming the vanishing gradient problem often encountered in traditional RNNs (Gers et al., 2002). An LSTM unit typically consists of a cell, an input gate, an output gate, and a forget gate. The cell stores values over arbitrary time intervals, and the three gates regulate the flow of information into and out of the cell. In terms of fault detection and diagnosis, LSTM networks perform similarly to ANNs by leveraging its ability to model temporal dependencies.

3.4.5. Data preprocessing and model parametrization

Data were collected at 5-second intervals and averaged to generate 5-minute readings, resulting in 6,870 observations for EC and 2,845 observations for pH. Data collection for pH was interrupted, leading to fewer observations than EC measurements. To address data imbalance, we employed the synthetic minority over-sampling technique (SMOTE) (Elreedy & Atiya, 2019). Subsequently, the variables were scaled. The dataset was then partitioned, with 70% allocated for training and 30% reserved for testing. For training the sensor-based FDD models, the last three observations (t–1, t–2, and t–3) were used. Conversely, for the sensor-independent FDD, all variables — including climate, volume of water, and dosing events— were utilized to predict EC and pH, as the feature analysis (Section 3.3.4) did not reveal substantial information regarding the importance of individual variables for prediction. Subsequently, hyperparameter tuning was performed for the machine learning models using grid search.

3.4.6. Model performance evaluation metrics

For sensor-based FDD, we evaluated model performance using accuracy, precision, recall, and F1-score to ensure reliable fault detection and minimize false alarms. Classification outcomes were analyzed using confusion matrices, providing a comprehensive understanding of model efficacy. This multi-metric approach allows stakeholders, such as farmers, to prioritize metrics based on specific use cases and farm management objectives. For instance, low precision, which increases false alarms, may hinder adoption due to operational disruptions, whereas low recall in detecting sensor faults could lead to inaccuracies in restoring nutrient solutions to optimal ranges, potentially impacting crop productivity. These metrics are described as follows:

Confusion Matrix: A confusion matrix provides a detailed breakdown of the model's predictions, allowing for a granular understanding of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Table 6 illustrates a two-by-two confusion matrix used in fault detection.

Table 6. Confusion matrix in the context of a binary classification task.

		Actual class	
		Normal	**Fault**
Predicted class	Normal	True positives (TP)	False positive (FP)
	Fault	False negative (FN)	True negative (TN)

Accuracy: This metric measures the overall correctness of the model's predictions, representing the proportion of correctly classified sensor data instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3.7)$$

Precision: Precision signifies the model's capability to minimize false positives, which is crucial when avoiding unnecessary maintenance or intervention in response to false alarms.

$$Precision = \frac{TP}{TP + FP} \qquad (3.8)$$

Recall: Recall quantifies the model's ability to correctly identify true sensor faults among all actual fault instances, thus ensuring minimal false negatives and reducing the risk of undetected issues.

$$Recall = \frac{TP}{TP + FN} \qquad (3.9)$$

F1-score: The F1-score, a harmonic mean of precision and recall, offers a balanced assessment of the model's performance, striking a crucial equilibrium between precision and recall, achieving optimal fault detection and diagnosis results.

$$F1 - score = 2 \times \frac{Precision * Recall}{Precision + Recall} \qquad (3.10)$$

Sensor-independent fault detection: The effectiveness is assessed by evaluating whether the models parameterized with climate and dosing variables, can predict EC within the manufacturer's nominal sensor accuracy (±3%). Therefore, the model performance is evaluated based on the following criteria:

$$Sensor\ independent\ fault\ detection\ accuracy = \frac{\#\ of\ prediction\ within\ sensor's\ accuracy}{Total\ \#\ of\ readings} \qquad (3.11)$$

Hyperparameter optimization was conducted using a grid search approach across the parameter space defined in Table 19. A pseudocode summary of the implemented approaches is presented in Table 20.

3.4.7. Cross-validation

To assess model robustness, we employed cross-validation (CV). Common CV techniques include k-fold CV, the holdout method, leave-one-out CV, and sliding window approaches (Ahamed et al., 2020). Given relatively large datasets, we employed 4-fold CV, following Jan et al. (2017), to enhance the robustness of our findings. The dataset is partitioned into four subsets, each containing patterns from the respective classes. In each iteration, one subset is designated for testing while the remaining three are used for training (Figure 34). It ensures an unbiased evaluation by

averaging performance metrics across all folds, thereby improving model reliability and generalizability.

3.4. Results and discussion

3.4.1. Sensors' behavior

3.4.1.1. Observation period and sensor variability

In this experiment, the EC data collection period was longer than that for pH, resulting in more observations available for model training. The time series data indicate that the fluctuation behavior of EC differed from that of pH (Figure 11). In other words, EC data exhibited high variance and complex relationships (Figure 11).

3.4.1.2. Statistical summary of normal and fault conditions

The mean values and standard deviations for normal EC and pH were 1.93 ± 0.34 dS/m and 5.56 ± 0.09, respectively. Figure 13 illustrates the statistical characteristics of EC and pH measurements under normal and different fault scenarios. While precision faults closely mimicked normal conditions, making them harder to detect, stuck faults exhibited a complete loss of variability, providing a clear and distinct fault signature.

For EC (Figure 13a), fault conditions such as bias up and drift up led to substantial increases in mean values, reaching 2.43 dS/m and 2.67 dS/m, respectively, compared to the normal mean of 1.93 dS/m. Conversely, bias down and drift down showed reductions in EC, with the mean dropping to 1.43 dS/m and 1.24 dS/m, respectively. Spike faults exhibited a slight increase in mean EC to 1.99 dS/m.

Similarly, for pH (Figure 13b), a similar pattern emerged. Bias up and drift up caused elevations in pH values, with means of 5.66 and 6.21, respectively. In contrast, bias down and drift down resulted in notable decreases, with the mean pH dropping to 4.56 and 4.88, respectively. Spike faults produced a sharp peak at a mean pH of 6.22, while stuck faults maintained constant values with a mean and median of 5.70.

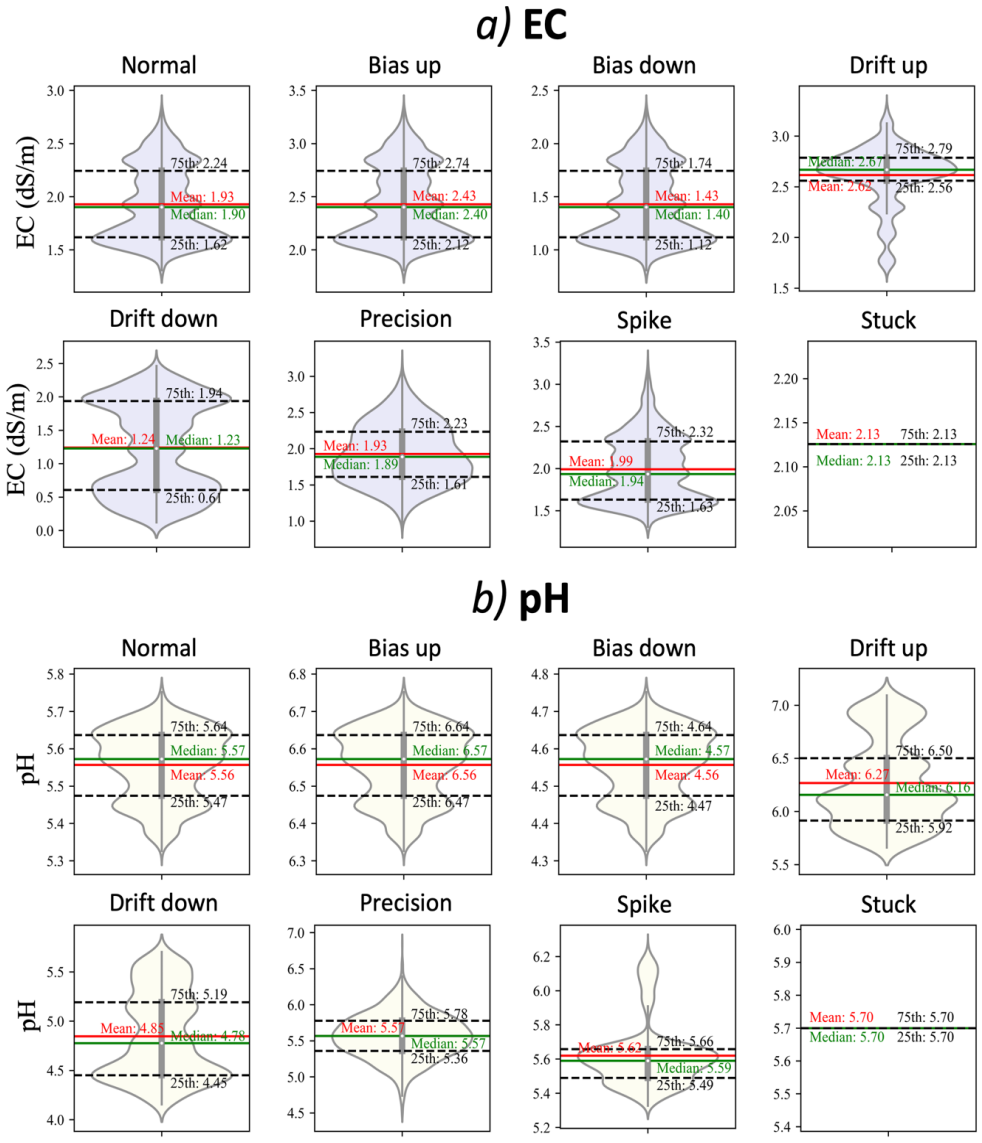


Figure 13. Violin plots comparing normal and faulty observations (including bias, drift, precision degradation, stuck, and spike) for EC (a) and pH (b) sensors in the hydroponic system. Each

subplot represents a specific fault scenario, including bias (up and down), drift (up and down), precision degradation, spike, and stuck sensor behavior. The red and green horizontal lines indicate the mean and median values, respectively, while the dashed black lines mark the 25th and 75th percentiles.

3.4.2. Hyperparameter optimization

3.4.2.1. Accuracy overview

Overall, all models showed higher fault detection accuracy for the pH sensor compared to the EC sensor (Table 18). Specifically, for fault diagnosis, SVM and KNN models performed better for pH, while ANN, RF, and LSTM models performed better for EC. Detection accuracy ranged from 0.7863 for SVM with the EC sensor to 0.9648 for RF with the pH sensor. Similarly, diagnosis accuracy varied, with SVM yielding the lowest accuracy for EC (0.7618) and RF the highest for pH (0.9295). Table 18 provides a summary of the best-performing models along with their respective hyperparameters.

3.4.2.2. Optimized parameters

For the SVM model, the combination of a high regularization parameter (C = 1000), the radial basis function (RBF) kernel, and specific gamma values (scale or auto) consistently provided the best accuracy across tasks, achieving up to 93.92% for pH detection. Since EC and pH sensor readings have different variability and clustering patterns in the feature space, KNN models demonstrated optimal parameters varying from n_neighbors = 3 and distance-weighted Manhattan metric for EC detection to n_neighbors = 11 with the same metric for pH detection, achieving a peak accuracy of 94.23%.

The ANN models required a more complex architecture, with hidden layers and hyperparameters customized to each task. Optimal configurations for detection tasks included smaller networks (e.g., [128, 64, 32] layers), while larger networks ([256, 128, 64] layers) were favored for pH diagnosis, yielding a maximum accuracy of 94.45%. RF models achieved high performance with n_estimators = 100, max_features set to either 'log2' or 'sqrt,' and small splits (min_samples_split = 3–5), reaching an accuracy of 96.48% for pH detection. LSTM models further demonstrated the importance of architecture depth, with hidden layer configurations like [32, 32, 32] and [64, 64, 64], combined with low dropout rates and learning rates of 0.0001, resulting in robust accuracies up to 93.56%. Performance metrics such as precision, recall, and F1-score followed a similar pattern to accuracy, as shown in Tables 18 and 19.

3.4.2.3. Handling data imbalance with SMOTE

Training the models on an imbalanced dataset had a more pronounced effect on other model performance metrics such as precision, recall, and F1-score. Imbalanced learning is a significant challenge in FDD because most data represent normal operations, leading to a scarcity of fault data (Zhou et al., 2020). This challenge is particularly relevant in diagnosis, where different faults occur at varying frequencies, exacerbating data imbalance among fault types (Gui et al., 2022). In this study, the minority class in both EC and pH datasets was the spike fault, constituting only 1.6% of the total observations. The models tended to misclassify these as normal, consistent with Mao et al. (2017) findings, which suggest that data imbalance can shift the classification hyperplane of machine learning algorithms toward the minority class, thereby increasing the likelihood of misclassification. Using SMOTE helped resolve imbalances, and similar results were observed for packaged rooftop units (FDD) in buildings (Ebrahimifakhar et al., 2020).

3.4.2.4. Sensor-based vs. sensor-independent fault detection

Due to the greater fluctuation observed in EC compared to pH, we employed a sensor-independent fault detection method for EC as an alternative to sensor-based fault detection. For fault diagnosis, however, we confined our analysis to sensor-based methods for both EC and pH.

3.4.3. Sensor-based fault detection for EC and pH

3.4.3.1. Model accuracy

The RF algorithm achieved the highest accuracy for both EC and pH sensors, with accuracies of 0.9369 and 0.9648, respectively. This high performance can be attributed to the ensemble nature of RF, which combines multiple decision trees to mitigate overfitting and improve generalization. Subsequently, both ANN and LSTM demonstrated superior fault detection performance; however, they also resulted in a higher incidence of false alarms compared to RF (Figure 14). This increases the need for manual inspection and may undermine farmer confidence in system performance. ANN achieved 0.9123 for EC and 0.9445 for pH, while LSTM achieved 0.9122 for EC and 0.9356 for pH. The slightly higher accuracy of ANN and LSTM for the pH sensor could be due to the nature of pH data, which comprises more stable and less noisy data, leading to a measurement to predict. While LSTM is powerful, its deployment in resource-constrained environments requires careful optimization and hardware considerations. For real-time fault detection in hydroponic systems, the simpler and better-performing model (e.g., RF) may strike a balance between performance and feasibility.

3.4.3.2. Comparative performance of models

The KNN algorithm showed relatively high accuracy for pH sensors (0.9423) but slightly lower performance for EC sensors (0.8896). This difference might be due to the KNN's sensitivity to the

data distribution. For EC sensors, the data seemed less clustered, leading to a reduction in accuracy. SVM showed a larger difference in performance between the two types of sensors, with an accuracy of 0.7863 for EC and 0.9392 for pH. SVM typically performs well with linearly separable data or can be effectively transformed into a linearly separable space through kernel functions. The lower accuracy for EC suggests that the data were not easily separable or that the chosen kernel was less effective in capturing the underlying patterns of EC sensor faults. While SVM is often reported to possess strong generalization capabilities, particularly with small sample sizes (Yin & Hou, 2016), our findings did not align with this expectation. This discrepancy may be attributed to the superior performance of other models in our study, resulting from extensive data collection.

Figure 14 illustrates the confusion matrices for all models. The results indicated that LSTM and ANN perform better for fault detection of EC and pH sensors, followed by RF, KNN, and SVM, which showed similar conclusions based on accuracy indices. The choice of model may vary depending on the farmers' preferences for monitoring sensor data. For instance, if the priority is to avoid missing any faulty data, fault recall is an important metric to consider. LSTM (recall: 0.992) and ANN (recall: 0.988) performed better than other models for EC sensors, while ANN (recall: 0.990) and RF (recall: 0.986) performed better in detecting faulty data for pH sensors (Table 18).

3.4.3.3. False alarm rates and practical implications

On average, all models misclassified normal observations as faulty approximately 13.9% of the time for EC sensors and 5.9% for pH sensors (Figure 14). KNN and SVM models perform better when detecting the normal operation of pH sensors, followed by RF, ANN, and LSTM. In contrast, KNN and RF perform relatively better for EC sensors, with an average misclassification rate of less than 10%. Based on the outcomes from the confusion matrix, the RF model performed relatively better for both sensors, but ANN and LSTM performed better for fault detection. Thus,

it can be concluded that ANN and LSTM models may require farmers to investigate more false alarms—normal signals misidentified as faults—while keeping the risk of missing actual faults below 2% (Figure 14).

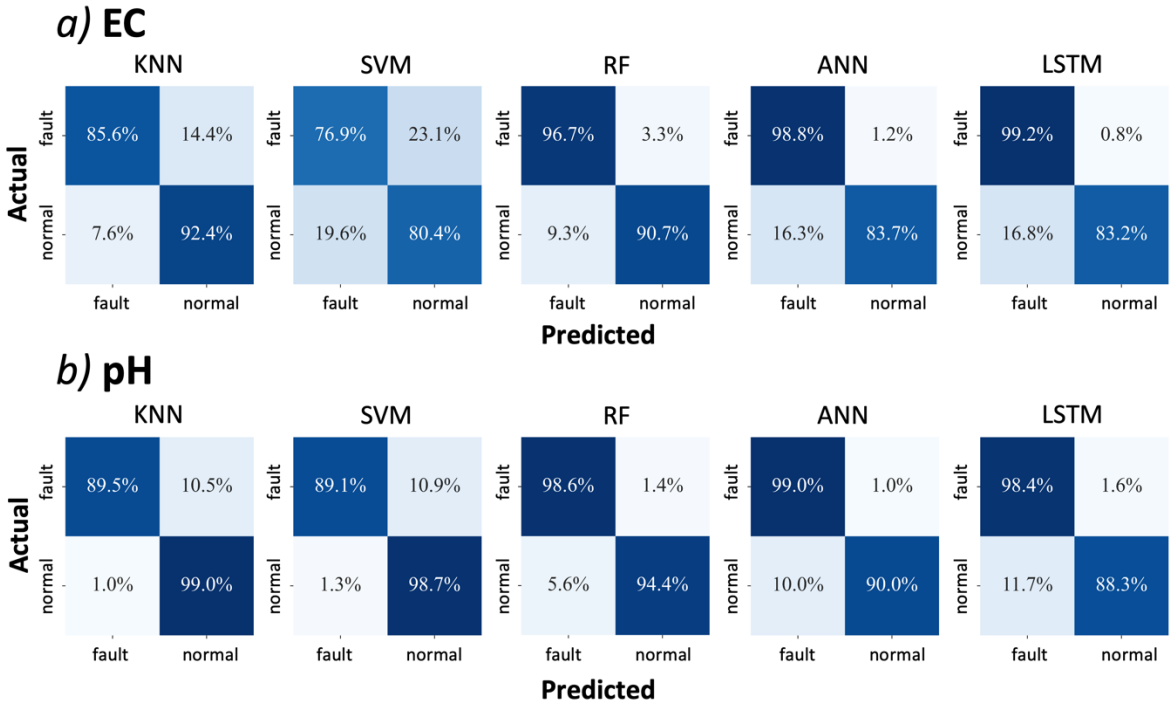


Figure 14. Confusion matrices of sensor-based fault detection in hydroponic solution monitoring systems using five selected machine learning models. The figure illustrates sensor faults compared to normal readings in two commonly used sensors: EC (a) and pH (b). Each confusion matrix provides a detailed comparison between the predicted observations (x-axis) and the actual observations (y-axis). RF provided the most robust sensor-based fault detection performance in hydroponic systems.

3.4.4. Sensor-independent fault detection for EC

3.4.4.1. Model selection and performance metrics

Sensor-independent fault detection for EC prediction was assessed using SVR and ANN. These models were selected based on their contrasting performance, with SVR identified as a low-performing and ANN as a high-performing model. The primary criterion for evaluating these models was their ability to predict EC within the sensor's nominal accuracy. Predictions that remained within this accuracy threshold were considered effective for distinguishing between normal and faulty sensor observations.

3.4.4.2. Analysis of feature correlations

The Pearson correlation coefficient analysis (Figure 33) revealed that EC did not exhibit a strong or linear relationship with the collected environmental (e.g., water temperature, air temperature, and relative humidity) and operational variables (e.g., lighting schedule, adding nutrients and water, nutrient volume). This lack of correlation was unexpected, especially considering that, based on the principle of mass conservation, variables such as the volume of nutrient solution and the concentration of added nutrients should logically influence EC. Given these insights, the Pearson correlation coefficient was not used to eliminate any features. Instead, all collected variables were included as inputs for the ANN and SVR models.

3.4.4.3. Prediction accuracy and fault identification

Both ANN and SVR models demonstrated strong performance in predicting EC (Figure 15), with respective root mean square error (RMSE) values of 0.0742 dS/m and 0.0354 dS/m. The coefficient of determination (R^2) further confirmed this, with values of 0.989 for ANN and 0.953

for SVR (Figure 15). These results suggest that both models effectively captured the underlying dynamics of EC within the nutrient solution.

However, a closer examination of the predictions reveals key differences in model performance. While generally accurate, the SVR model exhibited greater fluctuations around the observed EC values (Figure 15a). This increased variability suggests that SVR may be more sensitive to noise or less effective at generalizing across different conditions, leading to less stable predictions. On the other hand, the ANN model produced predictions that were more consistently aligned with the observed EC values (Figure 15b), particularly under stable conditions. Despite this overall better alignment, the ANN model sometimes struggled to predict abrupt changes in EC, often caused by adding large volumes of water or nutrients to a relatively small volume of nutrient solution in the tank.

When assessing sensor-independent fault detection accuracy, SVR achieved an accuracy rate of 49.4%, which is akin to the probability of flipping a coin. This low accuracy underscores the model's limitations in distinguishing between normal and faulty sensor readings. In contrast, the ANN model achieved an accuracy of 93.2%, substantially outperforming SVR and proving to be as reliable as other sensor-based fault detection methods, including those using synthetic fault data with SVM and KNN.

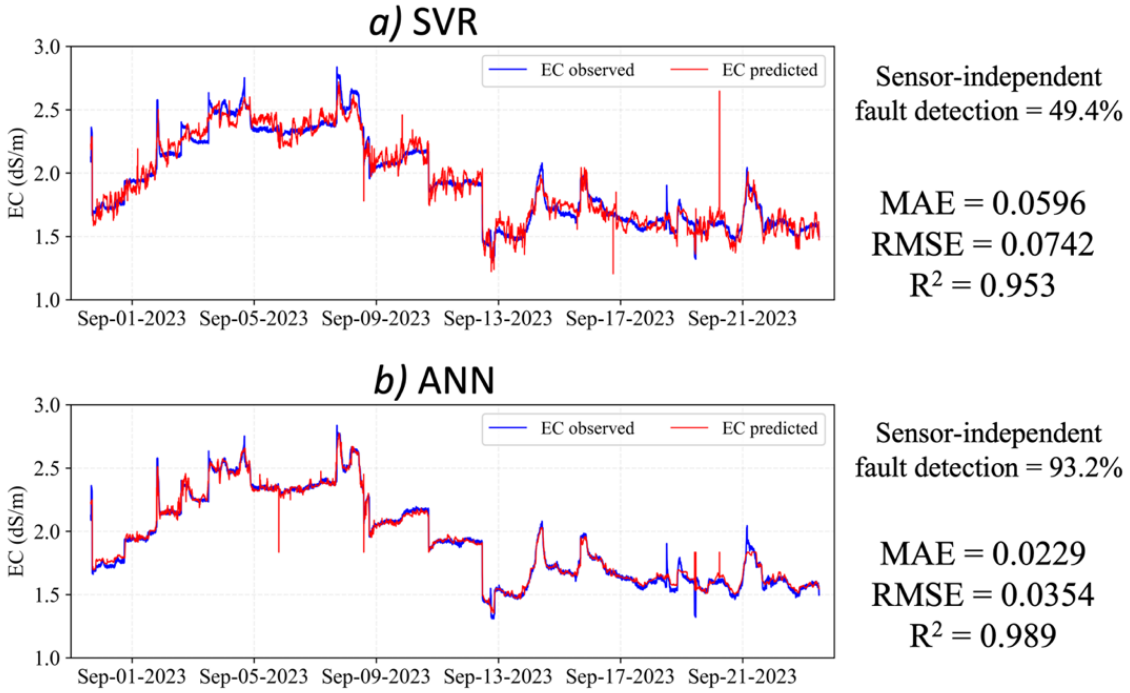


Figure 15. Time-series comparison of observed and predicted EC in a hydroponic lettuce system using (a) SVR and (b) ANN models, based on climate variables and dosing signals. Predicted values (red) closely follow observed values (blue), with ANN showing higher accuracy and better sensor-independent fault detection. Performance metrics include MAE, RMSE, and R^2.

3.4.4.4. Application scenarios and model limitations

Sensor-independent fault detection can serve as a complementary tool for estimating the expected range of EC in hydroponic solutions. While ANN models require greater computational resources, their integration into industrial hydroponic systems—alongside sensor-dependent approaches—may represent an optimal strategy for fault diagnosis. Importantly, even as a standalone method, the sensor-independent approach offers a robust and reliable means to approximate EC values, enhancing system resilience and supporting early fault detection in sensor networks.

3.4.5. Fault diagnosis for EC and pH

3.4.5.1. Classifier accuracy per fault type

The integration of machine learning models presented a promising advancement in fault diagnosis beyond traditional methods, which often rely heavily on manual inspection and farmer experience. We evaluated the efficacy of five machine learning models—SVM, KNN, RF, ANN, and LSTM—in diagnosing faults in the EC and pH sensors of hydroponic solutions used for lettuce growth. Table 7 summarizes the highest and lowest accuracy achieved by five tested machine learning classifiers for diagnosing six distinct fault categories. The highest-performing classifier(s) and their accuracy percentages are provided for each category, followed by the lowest-performing classifier(s).

For EC sensors, RF emerged as the most consistently high-performing model across several fault categories, aligning with previous findings (Anwar et al., 2025), with achieving accuracies of 97.1% for bias, 95.1% for drift, 82.7% for normal operation, 93.5% for precision degradation, 91% for stuck, and perfect 100% for spike diagnosis. The RF model's ability to achieve 100% accuracy in detecting the spike class is due to its capability to effectively handle sharp and distinct variations in the time-series input data (EC_t, EC_{t-1},EC_{t-2},EC_{t-3}). Spikes in EC represented abrupt deviations that are well captured by the decision trees within the RF ensemble. Each tree independently learned patterns from the data, and when combined, they created a robust model capable of identifying such anomalies with precision.

On the other hand, ANN and LSTM struggled with slightly lower accuracy in diagnosing normal observations and often confuse them with stuck faults. The underlying reason lied in the overlapping temporal patterns between normal and stuck conditions. In the case of stuck faults,

the EC values remained constant over time, resulting in a temporal profile (EC_t=EC_{t-1}=EC_{t-2}=EC_{t-3}) that resembled stable periods in normal operation. ANN, relying on learned weights from the provided data, may fail to differentiate these subtle variations due to insufficiently engineered features. LSTM, while capable of modeling temporal dependencies, may not effectively distinguish static patterns when there were no significant changes across the time window, leading to misclassifications between stuck and normal faults. Using derived features such as rate of change (ΔEC) can improve model performance and fault differentiation.

Conversely, the performance of models for pH sensors showed a distinct pattern. SVM excelled in detecting bias, spike, and stuck faults, with perfect accuracy of 100% due to its ability to establish clear decision boundaries in feature space, while RF achieved the highest accuracy for drift (85.3%), normal operation (92.1%), and precision degradation (87.8%), reflecting its robustness in modeling complex, noisy, or gradual temporal variations. Drift faults involved incremental changes over time, which RF effectively identified through its ensemble of decision trees, each capturing distinct feature relationships. Precision degradation, marked by increased variability around baseline pH values, was similarly well-addressed by RF's capacity to integrate multiple perspectives on feature interactions. Similar to SVM, KNN achieved perfect scores for spike and stuck fault detection by leveraging its distance-based classification mechanism, which exceled in identifying localized deviations and uniform patterns. The sharp transitions in spike faults and the static nature of stuck faults resulted in well-defined clusters, facilitating KNN’s precise classification.

ANN exhibited poor performance in normal conditions (50.1% for EC) and bias detection (45.9% for pH). Besides, SVM performed poorly in EC bias (48.8%) and drift detection (65.3% for EC, 63.5% for pH). The weak performance of SVM in drift conditions suggests a limitation in

capturing gradual sensor deviations, likely due to its inherent reliance on static decision boundaries.

Table 7. Comparison of five machine learning classifiers for fault diagnosis of EC and pH sensors across six fault categories. RF emerged as the most reliable classifier for diagnosing faults in EC and pH sensors, as well as distinguishing normal operational states. Most models demonstrated high accuracy in diagnosing sensor stuck faults compared to other fault types.

Classes	**Highest accuracy classifier**		**Lowest accuracy classifier**	
	EC	**pH**	**EC**	**pH**
Normal	RF (82.7%)	RF (92.1%)	ANN (50.1%)	KNN and SVM (66.3%)
Bias	RF (97.1%)	SVM (100%)	SVM (48.8%)	ANN (45.9%)
Drift	ANN (95.1%)	RF (85.3%)	KNN (64.5%)	SVM (63.5%)
Precision degradation	RF (93.5%)	RF (87.8%)	KNN (67.9%)	LSTM (78.8%)
Spike	RF (100%)	KNN and SVM (100%)	ANN and LSTM (98.3%)	LSTM (93.8%)
Stuck	KNN, SVM, ANN, LSTM (100%)	KNN and SVM (100%)	RF (91.0%)	RF (76.2%)

3.4.5.2. Insights from confusion matrices

Figure 16 illustrates the confusion matrices for EC and pH sensors, revealing critical insights into the models' error patterns. Notably, RF, ANN, and LSTM exhibited significant confusion between bias and stuck faults in pH sensors, while ANN and LSTM struggled to distinguish between normal and stuck faults in EC sensors. In general, most models perform better for spike and stuck faults. For spike fault, model selection exhibited minimal influence on classification performance, as all models achieved consistently high accuracy. Even the least effective model maintained an accuracy exceeding 93.8%, regardless of the selected fault parameters. The induced spike fault

parameter of 0.5 resulted in a various deviation, leading to a 26% increase in EC mean (1.93 dS/m) and a 9% variation in pH mean (5.56). Despite these variations, all models accurately diagnosed the spike fault patterns by analyzing the three most recent observations. Additionally, it could be concluded that the RF model outperforms for diagnosing all fault classes with the EC sensor, except in the case of bias fault in the pH sensor. These observations underscore the sensitivity of model performance to the defined fault parameters and suggest that certain fault types are more challenging to differentiate, depending on the sensor and the model employed.

The results align with existing literature, which suggests that the nature of the fault parameters influences the performance of machine learning models in temperature sensor fault diagnosis (Jan et al., 2021). This study confirms that the defined fault parameters (specified in Table 5) are crucial for clustering and distinguishing fault types (expect spike fault), affecting model accuracy and reliability. Here, we selected fault parameters for which the synthetically generated faults are not excessively divergent from normal observations (e.g., avoiding negative faults) in order to rigorously challenge the models' performance.

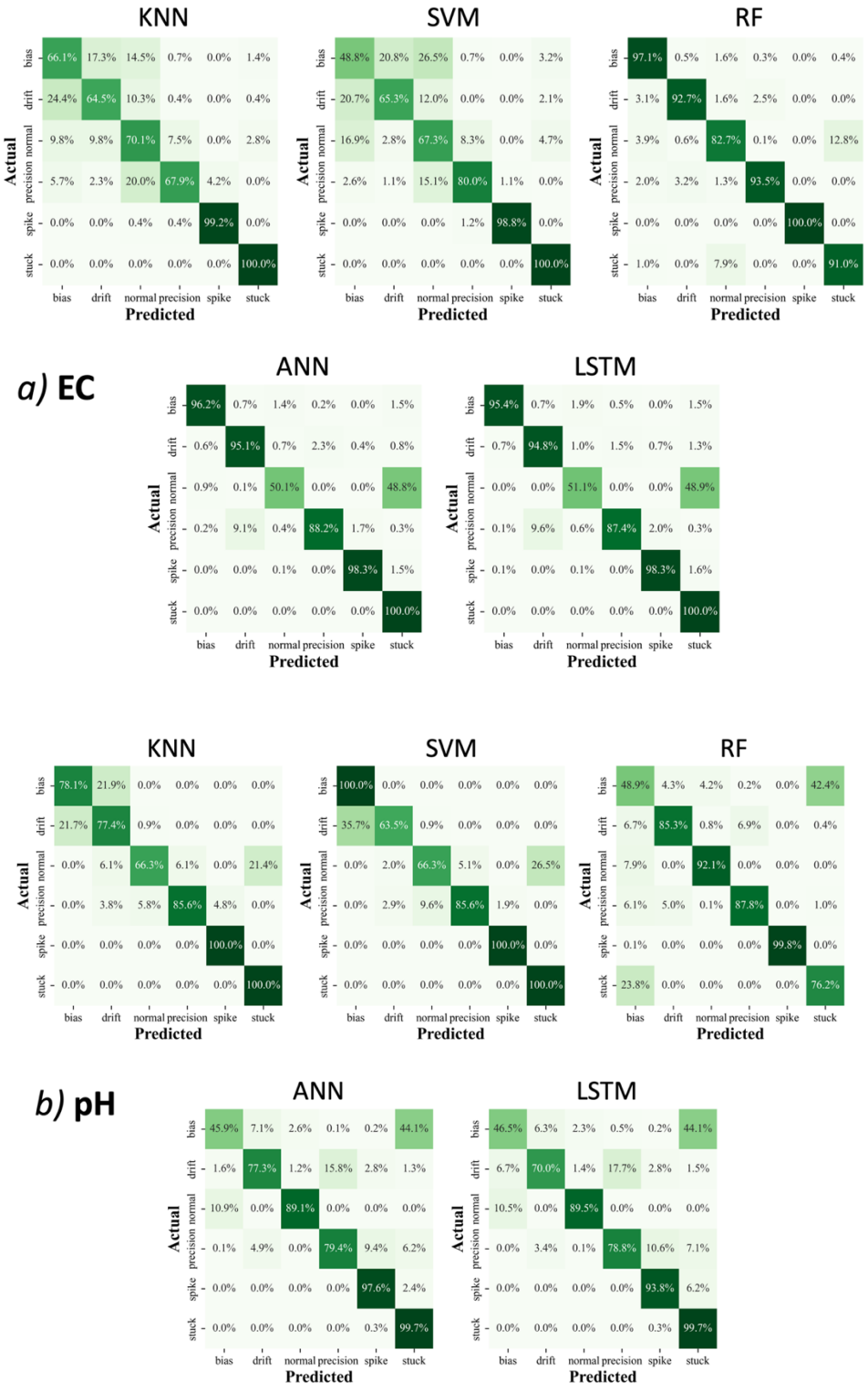


Figure 16. Confusion matrices of sensor-based fault diagnosis for EC (a) and pH (b) sensors in hydroponic solution monitoring systems using five selected machine learning models. Each

confusion matrix compares predicted fault categories (x-axis) against actual fault categories (y-axis), including normal, bias, drift, precision degradation, spike, and stuck conditions. RF is the most reliable choice for multi-class fault diagnosis in hydroponic systems.

3.4.6. Comparative performance evaluation with cross-validation

LSTM and ANN achieve the highest accuracy (0.925 and 0.999, respectively) for EC and pH, highlighting their superior capability in capturing non-linear patterns and temporal dependencies. Notably, RF also demonstrated strong performance (0.873 for EC and 0.999 for pH), making it a reliable alternative to deep learning. In contrast, SVM exhibited sub-optimal performance, particularly in EC detection, where its accuracy was 0.788.

For fault diagnosis (Figure 17b), RF emerged as the most reliable traditional algorithm, achieving 0.943 accuracy for EC and 0.859 for pH. However, ANN and LSTM performed nearly identically (≥0.879), indicating that deep learning architectures are well-suited for diagnosing sensor faults. KNN, in contrast, struggled significantly, particularly in EC diagnosis, where it achieved only 0.646 accuracy, suggesting its limitations in distinguishing fault patterns.

To further assess the sensor-independent EC prediction (Figure 17c), we compared ANN and SVR using three key metrics: R^2, MAE, and RMSE. The results indicate that SVR outperformed ANN across all metrics, achieving a higher (0.957 vs. 0.864), lower MAE (0.059 dS/m vs. 0.081 dS/m), and lower RMSE (0.072 dS/m vs. 0.125 dS/m). This suggests that, despite the strong learning capabilities of ANN, SVR provides a more stable and accurate prediction model for EC in a sensor-independent setting. Given the average EC of 1.93 dS/m and the nominal accuracy of the EC sensor (±3%), maintaining an MAE below 0.058 dS/m would ensure robust sensor-independent EC detection. Although both models demonstrated strong predictive performance, the observed MAE

exceeded 0.058 dS/m during cross-validation, potentially limiting their efficacy in EC sensor fault detection.

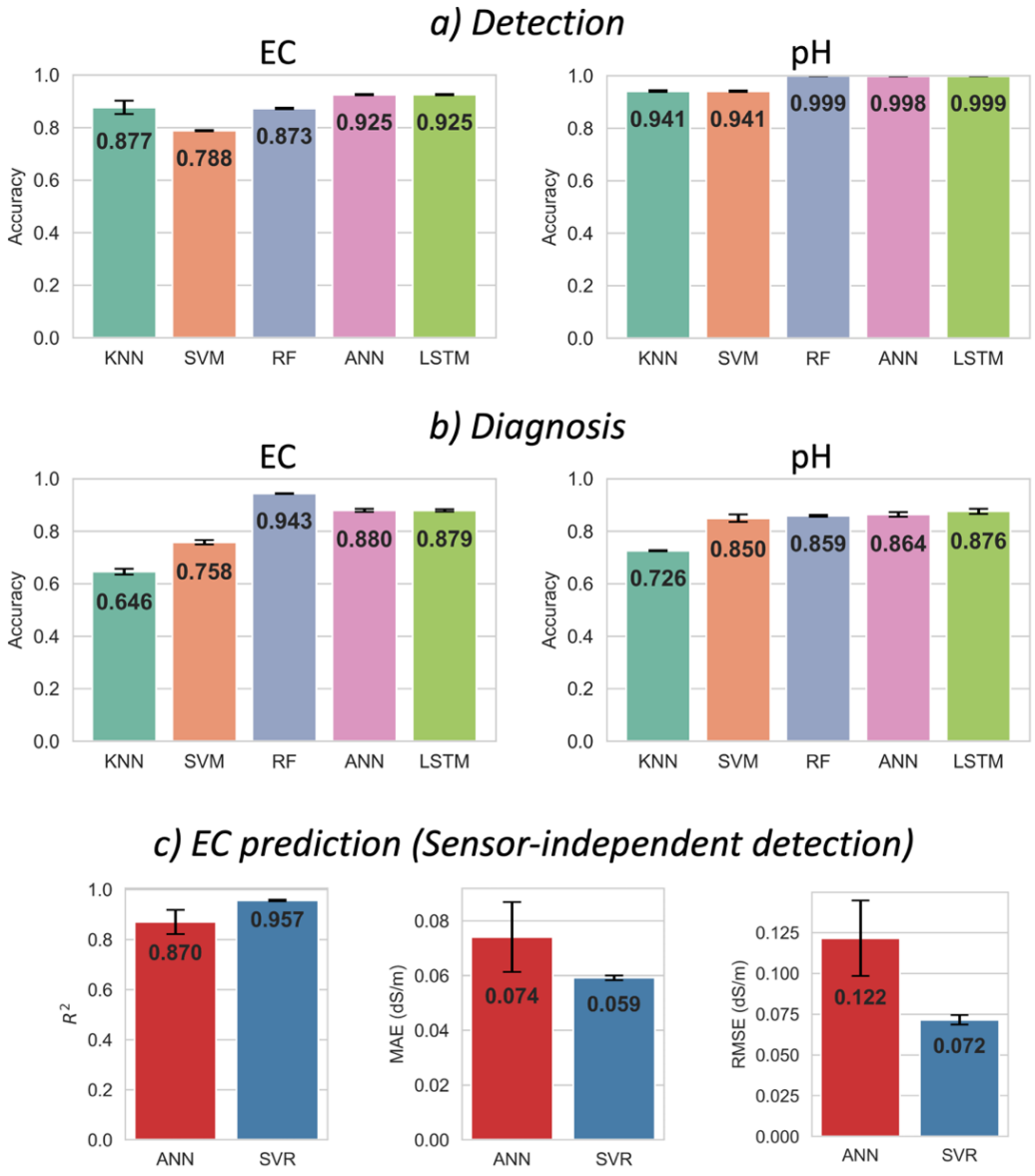


Figure 17. Cross-validation performance of machine learning models for EC and pH sensor's fault detection (a), diagnosis (b), and EC prediction (c). RF demonstrates the highest robustness and reliability for both fault detection and multi-class fault diagnosis in EC and pH sensors. For EC prediction, SVR shows lower uncertainty compared to ANN.

3.4.7. Model selection trade-offs in hydroponic systems

3.4.7.1. Suitability across farm scales

Integrating machine learning FDD in hydroponic systems necessitates balancing model accuracy with deployment feasibility, particularly across varying farm scales. Small-scale farms, constrained by limited computational resources and sparse data, benefit from lightweight models such as SVM and RF, which offer moderate-to-high accuracy with low inference costs. KNNs are unsuitable due to high memory and latency requirements, while deep models are generally infeasible without external infrastructure. Medium-scale farms can accommodate shallow ANNs and, where temporal resolution permits, LSTM networks. These models improve predictive performance but require greater computational support and model maintenance.

In contrast, large-scale farms equipped with high-frequency sensing and cloud infrastructure can leverage deep ANNs and LSTMs for time-series forecasting and control optimization. Despite their superior accuracy, these models demand extensive data, tuning, and deployment pipelines. RF remains a robust and interpretable alternative for ensemble or baseline integration. Sensor-independent fault detection offers a valuable complement to sensor-based methods. A hybrid approach combining sensor-based models (e.g., RF, ANN, LSTM) with sensor-independent ANN models may represent an optimal strategy for large-scale hydroponic systems.

3.4.7.2. Model efficiency and deployment feasibility

Generally, models such as LSTM and ANN tend to benefit significantly from larger datasets due to their higher capacity for learning complex patterns, whereas models such as KNN and SVM may plateau or even degrade in performance if the dataset becomes too large without appropriate optimization. In contrast, RF can handle both small and large datasets effectively but may become

computationally intensive at scale. Our current dataset provided a robust basis for evaluation; however, future work could explore model performance under varied data scales and evaluation metrics to ensure generalizability and scalability across different deployment scenarios.

Real-time deployment of data-driven fault detection models for EC and pH sensors in IoT systems is constrained by latency and limited computational resources. While SVM and KNN offer robust accuracy, their inference times are often prohibitive for edge deployment. RF provides faster predictions but can be memory intensive. ANN and LSTM models capture nonlinear and temporal patterns effectively, yet their high computational and power demands challenge real-time use. These trade-offs highlight the need for lightweight, optimized models or hybrid edge-cloud architectures to ensure both responsiveness and reliability in sensor monitoring. RF can enable a balanced and robust FDD model, offering continuous sensor performance assessment beyond traditional threshold-based approaches. While RF performed well, we acknowledge the limited adaptability to novel fault types compared to deep learning methods.

3.4.8. Study limitations and future research

Machine learning models for FDD of EC and pH sensors in hydroponic systems demonstrate substantial potential for broader application across different CEA systems, including aquaponics, greenhouses, and vertical farms. Also, this research supports advancements in sensor technology and AI-driven fault detection in agriculture and can be applied to ion-selective electrode sensors enable precise ion-level nutrient management (e.g., NO_3^-, NH_4^+, K^+, Ca^{2+}, PO_4^{3-}, Cl^-). The proposed machine learning framework for FDD of EC and pH sensors is well-suited for integration into commercial hydroponic systems. With low computational overhead, models such as RF can operate in real time on edge devices within existing nutrient management or IoT-based control platforms. The novel sensor-independent EC fault detection method enables hardware-agnostic

deployment, enhancing interoperability across diverse system architectures. However, deployment may be limited by hardware incompatibilities, sensor calibration variability, and inconsistent data protocols across hydroponic systems, requiring careful consideration. Addressing current limitations and exploring future research directions is essential to improve their robustness and applicability. The following points address the key limitations and possible future studies to increase the model's robustness in practical applications:

1. Specific management practices may influence the results. For instance, the impulsive behavior of the control signal in this study might change if low-flow rate peristaltic pumps are used instead of manually adding large doses of water, nutrients, and pH-down solutions. Additionally, EC fluctuations in the hydroponic solution can be affected by factors such as evapotranspiration rate (varying with vapor pressure deficit (VPD), radiation, and wind speed), nutrient interactions, nutrient volume, and dosing frequency. Future studies should investigate these variables in detail and use different management practices.
2. This study assumes that the nutrient solution behaves as a completely mixed reactor, where EC and pH measurements are consistent across different locations. However, inadequate mixing can lead to zones with varying characteristics. Future research should consider more sensor placements within the nutrient solution to evaluate these uncertainties. In our experiment, EC and pH sensors floated on the water with probes submerged about 6-8 cm from the water level.
3. The fluctuations in EC and pH suggest that incorporating additional variables could aid the generalization of machine learning models. For example, including buffer capacity variables (such as phosphate and bicarbonate buffers) could improve pH prediction accuracy. Additionally, ensemble machine learning techniques, which combine predictions from multiple models, could be explored to improve overall performance and reliability. The

application of hybrid models could further improve fault detection accuracy, as demonstrated in the literature on HVAC systems (Ebrahimifakhar et al., 2020) and temperature sensors (Jan et al., 2021).

4. Assessing the performance of the models under unforeseen experimental conditions, particularly with varying hydroponic solution management practices, is crucial for evaluating their generalizability; however, this remains unexamined in the current study. Furthermore, incorporating a broader range of cultivars, growing conditions, environmental factors, and sensor age in the training and testing phases is essential to enhance the robustness and applicability of the models across diverse agricultural systems. Thus, the dataset's potential biases, such as its limited representativeness for different hydroponic types and crops, environmental conditions, and real-world fault scenarios, may impact its generalizability. To enhance robustness, we recommend a hybrid approach that integrates sensor-independent and sensor-based fault detection methods for initial screening. Alternatively, ensemble methods (which combine the predictions of multiple individual models) are strongly recommended to enhance predictive performance and robustness in complex analytical frameworks.

3.5. Conclusions

Hydroponic farms depend on EC and pH sensors to maintain optimal nutrient solutions. Malfunctions in these sensors can result in inaccurate nutrient and pH levels, leading to the unnecessary disposal of nutrient solutions and fertilizers, thereby reducing resource efficiency and increasing environmental impact. Here, we evaluated five machine learning models—KNN, SVM, RF, ANN, and LSTM—for EC and pH sensor fault detection and diagnosis, alongside sensor-independent fault detection using SVR and ANN models to predict EC based on climate and dosing variables.

RF emerged as the most robust and reliable model, achieving the highest accuracy across both EC and pH sensors, making it the optimal choice for sensor fault detection. Although ANN and LSTM performed well, particularly with pH data, RF's consistency, lower computational complexity, and memory requirements solidify its superiority. In contrast, the SVM model's lower performance for EC sensors underscores the importance of model selection tailored to specific sensor characteristics.

Sensor-independent fault detection offered a robust means to estimate expected EC ranges in hydroponic systems. ANN significantly outperformed SVR for sensor-independent fault detection, achieving a 93.2% accuracy in identifying normal versus faulty sensor readings, compared to the SVR model's 49.4%, though with limitations in cross-validation. Thus, as detection is the first step in addressing sensor malfunctions, the proposed sensor-independent fault detection methodology offers a supportive approach applicable across diverse nutrient hydroponic management practices that farmers use.

Regarding sensor-based fault diagnosis, RF achieved outstanding results, with 97.1% diagnosis accuracy for bias, 95.1% for drift, and 100% for diagnosing both spike and stuck faults in EC sensors. For pH sensors, SVM excelled in diagnosing bias with perfect 100% accuracy, while RF was most effective in identifying drift (85.3%) and normal operation (92.1%).

These findings underscore the potential of machine learning models, particularly RF and ANN, to improve the accuracy and efficiency of hydroponic farming by enhancing sensor FDD and advancing precision agriculture in closed-loop, soil-less greenhouses and vertical farms. With careful consideration of hardware compatibility or data acquisition challenges, the models can be embedded within existing nutrient management software platforms or IoT-based controllers to improve hydroponic efficiency and reduce manual monitoring. Future research should focus on

validating the proposed sensor-independent fault detection framework across a broader range of hydroponic systems and crop types to enhance model robustness. Moreover, exploring the scalability and integration of this approach into commercial hydroponic operations will be critical for assessing its practical applicability and facilitating adoption in real-world agricultural settings.

3.7. References

Ahamed, M. S., Zmeureanu, R., Cortrufo, N., & Candanedo, J. (2020). Gray-box virtual sensor of the supply air temperature of air handling units. *Science and Technology for the Built Environment*, *26*(8), 1151–1162. https://doi.org/10.1080/23744731.2020.1785812

Anushka, Arya, C., Tripathi, A., Singh, P., Diwakar, M., Sharma, K., & Pandey, H. (2021). Object Detection using Deep Learning: A Review. *Journal of Physics: Conference Series*, *1854*(1), 012012. https://doi.org/10.1088/1742-6596/1854/1/012012

Anwar, T., Mu, C., Yousaf, M. Z., Khan, W., Khalid, S., Hourani, A. O., & Zaitsev, I. (2025). Robust fault detection and classification in power transmission lines via ensemble machine learning models. *Scientific Reports*, *15*(1), 2549. https://doi.org/10.1038/s41598-025-86554-2

Baharvand, S., Jozaghi, A., Fatahi-Alkouhi, R., Karimzadeh, S., Nasiri, R., & Lashkar-Ara, B. (2021). Comparative Study on the Machine Learning and Regression-Based Approaches to Predict the Hydraulic Jump Sequent Depth Ratio. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, *45*(4), 2719–2732. https://doi.org/10.1007/s40996-020-00526-2

Bamsey, M., Graham, T., Thompson, C., Berinstain, A., Scott, A., & Dixon, M. (2012). Ion-Specific Nutrient Management in Closed Systems: The Necessity for Ion-Selective Sensors in

Terrestrial and Space-Based Agriculture and Water Management Systems. *Sensors*, *12*(10), 13349–13392. https://doi.org/10.3390/s121013349

Banjara, N. K., Sasmal, S., & Voggu, S. (2020). Machine learning supported acoustic emission technique for leakage detection in pipelines. *International Journal of Pressure Vessels and Piping*, *188*, 104243. https://doi.org/10.1016/j.ijpvp.2020.104243

Bouchaaba, Z., Santamaria, P., Choukr-Allah, R., Lamaddalena, N., & Montesano, F. F. (2015). Open-cycle drip vs closed-cycle subirrigation: Effects on growth and yield of greenhouse soilless green bean. *Scientia Horticulturae*, *182*, 77–85. https://doi.org/10.1016/j.scienta.2014.11.007

Capozzoli, A., Lauro, F., & Khan, I. (2015). Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Systems with Applications*, *42*(9), 4324–4338. https://doi.org/10.1016/j.eswa.2015.01.010

Carotti, L., Graamans, L., Puksic, F., Butturini, M., Meinen, E., Heuvelink, E., & Stanghellini, C. (2021). Plant Factories Are Heating Up: Hunting for the Best Combination of Light Intensity, Air Temperature and Root-Zone Temperature in Lettuce Production. *Frontiers in Plant Science*, *11*, 592171. https://doi.org/10.3389/fpls.2020.592171

Chen, Z., O'Neill, Z., Wen, J., Pradhan, O., Yang, T., Lu, X., Lin, G., Miyata, S., Lee, S., Shen, C., Chiosa, R., Piscitelli, M. S., Capozzoli, A., Hengel, F., Kührer, A., Pritoni, M., Liu, W., Clauß, J., Chen, Y., & Herr, T. (2023). A review of data-driven fault detection and diagnostics for building HVAC systems. *Applied Energy*, *339*, 121030. https://doi.org/10.1016/j.apenergy.2023.121030

Dehestani, D., Su, S., Nguyen, H., & Guo, Y. (2013). Robust fault tolerant application for HVAC system based on combination of online SVM and ANN black box model. *2013 European Control Conference (ECC)*, 2976–2981. https://doi.org/10.23919/ECC.2013.6669140

Ebrahimifakhar, A., Kabirikopaei, A., & Yuill, D. (2020). Data-driven fault detection and diagnosis for packaged rooftop units using statistical machine learning classification methods. *Energy and Buildings*, *225*, 110318. https://doi.org/10.1016/j.enbuild.2020.110318

Elreedy, D., & Atiya, A. F. (2019). A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. *Information Sciences*, *505*, 32–64. https://doi.org/10.1016/j.ins.2019.07.070

Ferentinos, K. P., & Albright, L. D. (2003). Fault Detection and Diagnosis in Deep-trough Hydroponics using Intelligent Computational Tools. *Biosystems Engineering*, *84*(1), 13–30. https://doi.org/10.1016/S1537-5110(02)00232-5

Gers, F., A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*, *3*(Aug), 115–143.

Gui, X., Zhang, J., Tang, J., Xu, H., Zou, J., & Fan, S. (2022). A Quadruplet Deep Metric Learning model for imbalanced time-series fault diagnosis. *Knowledge-Based Systems*, *238*, 107932. https://doi.org/10.1016/j.knosys.2021.107932

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed). Pearson Education.

Hochreiter, S. (1998). The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *06*(02), 107–116. https://doi.org/10.1142/S0218488598000094

Jan, S. U., Lee, Y. D., & Koo, I. S. (2021). A distributed sensor-fault detection and diagnosis framework using machine learning. *Information Sciences*, *547*, 777–796. https://doi.org/10.1016/j.ins.2020.08.068

Jan, S. U., Lee, Y.-D., Shin, J., & Koo, I. (2017). Sensor Fault Classification Based on Support Vector Machine and Statistical Time-Domain Features. *IEEE Access*, *5*, 8682–8690. https://doi.org/10.1109/ACCESS.2017.2705644

Jiang, Y., & Yin, S. (2018). Recursive Total Principle Component Regression Based Fault Detection and Its Application to Vehicular Cyber-Physical Systems. *IEEE Transactions on Industrial Informatics*, *14*(4), 1415–1423. https://doi.org/10.1109/TII.2017.2752709

Lee, K.-P., Wu, B.-H., & Peng, S.-L. (2019). Deep-learning-based fault detection and diagnosis of air-handling units. *Building and Environment*, *157*, 24–33. https://doi.org/10.1016/j.buildenv.2019.04.029

Li, D., Wang, Y., Wang, J., Wang, C., & Duan, Y. (2020). Recent advances in sensor fault diagnosis: A review. *Sensors and Actuators A: Physical*, *309*, 111990. https://doi.org/10.1016/j.sna.2020.111990

Ludeña-Choez, J., Choquehuanca-Zevallos, J. J., & Mayhua-López, E. (2019). Sensor nodes fault detection for agricultural wireless sensor networks based on NMF. *Computers and Electronics in Agriculture*, *161*, 214–224. https://doi.org/10.1016/j.compag.2018.06.033

Mao, W., He, L., Yan, Y., & Wang, J. (2017). Online sequential prediction of bearings imbalanced fault diagnosis by extreme learning machine. *Mechanical Systems and Signal Processing*, *83*, 450–473. https://doi.org/10.1016/j.ymssp.2016.06.024

Mohammadhossein Shekarian, S., Aminian, M., Mohammad Fallah, A., & Akbary Moghaddam, V. (2024). AI-powered sensor fault detection for cost-effective smart greenhouses. *Computers and Electronics in Agriculture*, *224*, 109198. https://doi.org/10.1016/j.compag.2024.109198

Moosavian, A., Ahmadi, H., Tabatabaeefar, A., & Khazaee, M. (2013). Comparison of Two Classifiers; K-Nearest Neighbor and Artificial Neural Network, for Fault Diagnosis on a Main Engine Journal-Bearing. *Shock and Vibration*, *20*(2), 263–272. https://doi.org/10.1155/2013/360236

Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2011). Behaviour Analysis of Multilayer Perceptronswith Multiple Hidden Neurons and Hidden Layers. *International Journal of Computer Theory and Engineering*, 332–337. https://doi.org/10.7763/IJCTE.2011.V3.328

Pearson, K. (1896). VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *187*, 253–318. https://doi.org/10.1098/rsta.1896.0007

Pita, A. J., Leon, J. R., & Mendoza, O. B. (2021). Smart monitoring system to diagnose faults in a greenhouse. *2021 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, 1–6. https://doi.org/10.1109/ROPEC53248.2021.9668013

Ragab, A., El Koujok, M., Ghezzaz, H., & Amazouz, M. (2021). Fault diagnosis in industrial processes based on predictive and descriptive machine learning methods. In *Applications of Artificial Intelligence in Process Systems Engineering* (pp. 207–254). Elsevier. https://doi.org/10.1016/B978-0-12-821092-5.00002-4

Roka, S., Diwakar, M., Singh, P., & Singh, P. (2023). Anomaly behavior detection analysis in video surveillance: A critical review. *Journal of Electronic Imaging*, *32*(04). https://doi.org/10.1117/1.JEI.32.4.042106

Shohet, R., Kandil, M. S., Wang, Y., & McArthur, J. J. (2020). Fault detection for non-condensing boilers using simulated building automation system sensor data. *Advanced Engineering Informatics*, *46*, 101176. https://doi.org/10.1016/j.aei.2020.101176

Song, B., Tan, S., Shi, H., & Zhao, B. (2020). Fault detection and diagnosis via standardized k nearest neighbor for multimode process. *Journal of the Taiwan Institute of Chemical Engineers*, *106*, 1–8. https://doi.org/10.1016/j.jtice.2019.09.017

Taheri, S., Ahmadi, A., Mohammadi-Ivatloo, B., & Asadi, S. (2021). Fault detection diagnostic for HVAC systems via deep learning algorithms. *Energy and Buildings*, *250*, 111275. https://doi.org/10.1016/j.enbuild.2021.111275

Tutz, G. (1991). Sequential models in categorical regression. *Computational Statistics & Data Analysis*, *11*(3), 275–295. https://doi.org/10.1016/0167-9473(91)90086-H

Wang, Q., Nguyen, T.-T., Huang, J. Z., & Nguyen, T. T. (2018). An efficient random forests algorithm for high dimensional data classification. *Advances in Data Analysis and Classification*, *12*(4), 953–972. https://doi.org/10.1007/s11634-018-0318-1

Willsky, A., & Jones, H. (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, *21*(1), 108–112. https://doi.org/10.1109/TAC.1976.1101146

Wu, H., & Zhao, J. (2018). Deep convolutional neural network model based chemical process fault diagnosis. *Computers & Chemical Engineering*, *115*, 185–197. https://doi.org/10.1016/j.compchemeng.2018.04.009

Yan, K., Huang, J., Shen, W., & Ji, Z. (2020). Unsupervised learning for fault detection and diagnosis of air handling units. *Energy and Buildings*, *210*, 109689. https://doi.org/10.1016/j.enbuild.2019.109689

Yang, H., Hassan, S. G., Wang, L., & Li, D. (2017). Fault diagnosis method for water quality monitoring and control equipment in aquaculture based on multiple SVM combined with D-S evidence theory. *Computers and Electronics in Agriculture*, *141*, 96–108. https://doi.org/10.1016/j.compag.2017.05.016

Yang, J., Tian, K., Zhao, H., Feng, Z., Bourouis, S., Dhahbi, S., Khan, A. A., Berrima, M., & Por, L. Y. (2025). Wastewater treatment monitoring: Fault detection in sensors using transductive learning and improved reinforcement learning. *Expert Systems with Applications*, *264*, 125805. https://doi.org/10.1016/j.eswa.2024.125805

Yin, Z., & Hou, J. (2016). Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing*, *174*, 643–650. https://doi.org/10.1016/j.neucom.2015.09.081

Zhou, Z., Li, G., Wang, J., Chen, H., Zhong, H., & Cao, Z. (2020). A comparison study of basic data-driven fault diagnosis methods for variable refrigerant flow system. *Energy and Buildings*, *224*, 110232. https://doi.org/10.1016/j.enbuild.2020.110232

Chapter 4: OptiDose: A model predictive control framework for precise water and fertilizer dosing in closed-loop hydroponics

4.1. Abstract

Achieving closed-loop hydroponic systems necessitates precise regulation of individual macro- and micronutrients. However, nutrient management in hydroponics remains largely constrained to electrical conductivity (EC) and pH-based approaches due to the complexity of managing individual ions and the coupling inherent in multi-element fertilizer formulations. In this study, we develop and implement a model predictive control (MPC) framework, OptiDose, to optimize daily fertigation strategies for hydroponically grown lettuce. The system integrates six fertilizer sources—calcium nitrate, magnesium sulfate, monopotassium phosphate, potassium nitrate, magnesium nitrate, and potassium sulfate—to maintain the concentrations of nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and sulfur (S) within crop-specific sufficiency ranges. Five management scenarios are tested to evaluate system performance under varying operational constraints. Results demonstrate the capability of OptiDose to achieve zero nutrient deficiency or toxicity throughout the growth cycle while substantially improving resource-use efficiency. Compared to baseline conditions, water use efficiency increased nearly sixfold, and fertilizer use efficiency doubled, reaching up to 34 g/L 12.5 g/g, respectively. These enhancements translated into substantial reductions in both water and fertilizer input costs and highlighted the potential of MPC-based fertigation to advance precision nutrient management in hydroponic cultivation.

Keywords: Model Predictive Control (MPC); Nutrient Solution Optimization; Precision Nutrient Dosing; Precision Hydroponic; Closed-Loop Fertigation.

4.2. Introduction

Hydroponics, a form of soilless agriculture, involves cultivating plants in aqueous solutions enriched with essential mineral nutrients, thereby eliminating the need for soil as a growth medium. Hydroponic systems are broadly categorized into open and closed-loop systems. In open systems, the applied nutrient solution is discarded after a single use, whereas closed systems collect and recycle excess nutrient solution, enabling reuse throughout the cultivation cycle (Raviv et al., 2019). The environmental ramifications of open systems are significant, as the discharge of nutrient-laden effluent can contaminate surrounding ecosystems and groundwater sources (Pandey et al., 2023). In contrast, closed-loop hydroponics has demonstrated a 20–40% reduction in nutrient input requirements (Pandey et al., 2023). In some cases, water and fertilizer inputs can be reduced by approximately 20–35% and 40–50%, respectively, in comparison to conventional open soilless cultivation systems, while substantially minimizing—or in some instances entirely eliminating—nutrient losses to the environment and offering both economic and ecological advantages (Giannothanasis et al., 2024; Grewal et al., 2011; Massa et al., 2011). For optimal growth and productivity, plants require 17 essential elements (Trejo-Téllez & Gómez-Merino, 2012). Of these, carbon (C), hydrogen (H), and oxygen (O) are derived from atmospheric CO_2 and water. The remaining 14 elements must be supplied through the nutrient solution and are classified as macronutrients—nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and sulfur (S)—and micronutrients—iron (Fe), zinc (Zn), manganese (Mn), copper (Cu), boron (B), molybdenum (Mo), chlorine (Cl), and nickel (Ni) (Kim et al., 2023). Nutrient management in hydroponic systems is conventionally regulated via electrical conductivity (EC) measurements,

which serve as a proxy for total ion concentration. Nutrient replenishment is typically automated, with concentrated fertilizer solutions injected proportionally to reductions in EC caused by plant uptake (Bailey et al., 1988). However, EC provides no information on the individual ion concentrations and cannot reflect nutrient-specific imbalances (Kim et al., 2023). Such limitations hinder the precision management of individual nutrient formulations tailored to the dynamic demands of crops.

Advances in sensor networks and fertigation automation offer promising solutions for closed-loop nutrient optimization. Although several studies have demonstrated the viability of automated fertigation systems for soilless cultivation (Domingues et al., 2012; Pandey et al., 2023; Rahman et al., 2018; Steidle Neto et al., 2014), sustaining individual nutrient balance remains a persistent challenge. Conventional monitoring strategies, primarily based on EC and pH measurements (Cho et al., 2017), cannot detect ionic imbalances or the accumulation of non-absorbed solutes, which can arise due to differential uptake rates at the root interface. These deficiencies often necessitate manual interventions such as foliar sprays to mitigate nutrient disorders, potentially compromising yield and plant health.

To address these limitations, more refined analytical approaches are required. Recent advancements in nutrient concentration measurement techniques—including colorimetry, spectrophotometry, ion-selective electrodes, and emission spectroscopy (Langenfeld et al., 2022)—have enabled more precise characterization of individual ionic species within nutrient solutions. These innovations offer unprecedented opportunities to refine nutrient management strategies by aligning nutrient dosing more closely with plant-specific uptake dynamics. One such approach, implemented in the decision support system NUTRISENSE (Giannothanasis et al., 2024; Savvas et al., 2023), regulates the nutrient solution supplied to plants by analyzing the

chemical composition of the drainage solution. This feedback-based strategy aims to harmonize nutrient input with crop uptake, thereby maintaining target root zone concentrations for ions. The algorithm was evaluated in a pilot study within a commercial greenhouse cultivating tomatoes in a closed-loop soilless system. However, the system relies on the sequence calculation for individual nutrient mass balances to adjust dosing, which might be tedious and susceptible to error. A more systematic and scalable approach, such as the matrix-based method proposed by Jung et al. (2015) could streamline these calculations by simultaneously considering multiple nutrient interactions and fertilizer compositions. Notably, integrating such matrix-based nutrient balancing with optimization frameworks that adjust both water volumes and fertilizer dosing, accounting for the specific elemental composition of each fertilizer, has yet to be developed and applied in practice. To date, no widely adopted control algorithm enables real-time optimization of fertilizer inputs based on individual nutrient concentrations rather than aggregate EC values. Traditional control methods are also unaware of the economic/environmental costs of their actions. Developing such precision fertigation algorithms would significantly advance sustainable and high-efficiency hydroponic crop production.

Agricultural systems are inherently complex and subject to uncertainty (Ding et al., 2018; Kamilaris & Prenafeta-Boldú, 2018). Traditional control strategies, including on/off and proportional-integral-derivative (PID) controllers, are widely used due to their simplicity (Afram & Janabi-Sharifi, 2014; Christofides et al., 2013; Ding et al., 2018). Alternatively, fuzzy logic enables adaptive control of nutrient solutions under dynamic cultivation conditions (Catota-Ocapana et al., 2025; Mohamed et al., 2022). Fuzzy logic, in particular, has shown clear advantages over traditional PID control in hydroponics nutrient management. Zhang et al. (2019) reported that hybrid fuzzy PID reduced overshoot, improved response time, and enhanced system stability.

Other studies (Dela Vega et al., 2021; Nurmahaludin et al., 2020) successfully applied fuzzy logic to regulate pH and EC in nutrient solutions, with the latter incorporating the Sugeno method and servo motor adjustments. Mashumah et al. (2018) combined image processing with fuzzy logic to adjust EC based on plant age, achieving high accuracy. Puno et al. (2020) further expanded on sensor-driven fuzzy logic control systems to manage EC, pH, and water levels; however, their design included draining the mixing tank—a potentially problematic feature due to environmental concerns over nutrient-rich wastewater and the added operational costs for water and fertilizer replenishment. A recent study further revealed that while PID responded faster, fuzzy control minimized overshoot, and consensus algorithms offered the best overall balance (Nizam et al., 2024). Despite their contributions, none of these studies addressed nutrient solution preparation beyond basic EC and pH control. Sometimes overshooting led to the implementation of drainage systems (e.g., Puno et al., 2020). Besides, these methods struggle with dynamic processes involving time delays and require extensive tuning, making them impractical for real-time precision agriculture (Ya-Gang Wang et al., 2001), particularly for multiple-input multiple-output (MIMO) systems with operational constraints.

Model predictive control (MPC) is an advanced class of algorithms that leverages an explicit process model to anticipate a system's future behavior. MPC utilizes a dynamic process model to predict the system and solve constrained optimization problems in real time (Qin & Badgwell, 2003). The optimization-based framework enables MPC to minimize economic/environmental costs while achieving the specified control objectives. Its ability to manage multivariate processes and enforce state and input constraints has made it indispensable in modern industrial applications (Christofides et al., 2013).

MPC has been extensively applied in agriculture and open-field irrigation (Abioye et al., 2023), comprehensively reviewing its applications (Bwambale et al., 2023; Ding et al., 2018). Within the context of CEA, including greenhouses and vertical farms, MPC has been employed to enhance energy efficiency (Bersani et al., 2021; Hu & You, 2022b; Lin et al., 2020; Mahmood et al., 2023a; Ullah et al., 2022), optimize harvest timing and market strategies for crop production (García-Mañas et al., 2024), maximize yield within the spatial constraints of vertical farming systems (Bagnerini et al., 2023), minimize operational costs (Chen et al., 2022; Lin et al., 2021; Ren et al., 2022), and regulate climatic conditions (Blasco et al., 2007; Hamidane et al., 2024; Hu & You, 2022a; Jung et al., 2020; Mahmood et al., 2023b, 2025; Nakayama et al., 2024; Padmanabha et al., 2020; Qi et al., 2024). However, despite its extensive use in various aspects of CEA, the efficiency of MPC in optimizing water and fertilizer use remains limited. To our knowledge, its integration into water and fertilizer management in soilless production systems has not been comprehensively explored. Given its capability for dynamic optimization, MPC holds substantial potential for advancing precision agriculture, particularly in fine-tuning element-level nutrient management in hydroponic systems. MPC presents a robust and adaptive framework for optimizing fertilizer dosing in hydroponic systems, particularly given their inherently MIMO dynamics and stringent control constraints (Mayne, 2014).

This study applied an MPC framework to optimize water and nutrient dosing for lettuce cultivation in a CEA hydroponic system. The model accounted for daily plant uptake dynamics of water and macronutrients, enabling precise regulation of input delivery. Six commonly used fertilizers—calcium nitrate, magnesium sulfate, monopotassium phosphate, potassium nitrate, magnesium nitrate, and potassium sulfate—were incorporated to meet essential nutrient requirements. Five distinct farm management scenarios were evaluated, and a suite of standardized performance

metrics was used to assess the effectiveness of the MPC approach. The results demonstrate the potential of MPC to enhance water and nutrient use efficiency and enable individual nutrient-level management in hydroponic crop production.

4.3. Materials and methods

4.3.1. Experimental setup and conditions

The experiment was conducted in an indoor farming facility at the University of California, Davis, USA (38.54°N, 121.75°W). A closed-loop nutrient film technique (NFT) hydroponic system was employed (Figure 18f) to cultivate butterhead lettuce (*Lactuca sativa* var. *capitata*). Fourteen-day-old seedlings, initially grown in cocopeat, were transplanted into the system on February 22, 2024, and cultivated until harvest on March 22, 2024. The hydroponic system consisted of two vertical racks, each measuring 1.0 × 1.4 m, accommodating 70 lettuce plants.

4.3.1.1. Environmental conditions and lighting

Environmental conditions, including temperature, relative humidity, EC, and nutrient solution volume, were recorded throughout the experiment (Figure 18b–e). The temperature exhibited tier-dependent variations, averaging 22.9 ± 2.9°C in the upper tier and 21.6 ± 2.8°C in the lower tier. Relative humidity remained stable at 50 ± 7% in the upper tier and 49 ± 7% in the lower tier. A 16-hour photoperiod was maintained from 08:00 to 24:00, with an irradiance level of 250 μmol/m^2·s at the rack center, measured using a LI-COR-180 (LI-COR, USA). The corresponding daily light integral (DLI) was 14.4 mol/m^2·d. Illumination was provided by cool-white, fluorescent lamps (6500K T5, VIVOSUN, USA).

4.3.1.2. Nutrient solution and growth monitoring

The nutrient solution composition followed the formulation proposed by Carotti et al. (2021), with macro-nutrient concentrations (mg/L) of N: 182.0, P: 136.4, K: 234.0, Ca: 120.0, Mg: 20.4, S: 25.3. Micronutrients were supplemented on the transplanting day at the following concentrations (mg/L): B: 0.137, Cu: 0.410, Fe: 0.820, Mn: 0.684, Zn: 1.230. Due to the limited tank capacity **(**~95 L), the nutrient solution tank was replenished regularly to maintain volume and nutrient balance. Water additions (10 L each) were made on February 29, March 4, 15, 16, and 17, while nutrient solution was added on March 6, 8, 10, 12 (with calcium nitrate only), 14, and 20. On March 8 and 20, water and nutrient solution (10 L each) were added (Table 21). Daily fresh weight measurements were performed on 18 lettuce heads cultivated on the lower tier of the growing system to monitor biomass accumulation over time. Each measurement included the total weight of the plant (shoot and root) along with the cocopeat substrate. To isolate the net plant biomass, the saturated weight of the substrate, determined independently under full water retention conditions, was subtracted from the total measured weight.

4.3.1.3. Sensor calibration and data acquisition

Nutrient solution EC was monitored using a Vernier EC sensor (Vernier, USA), which was calibrated using a two-point calibration method (EC: 1 and 50 dS/m). The pH of the solution was maintained at ~5.5 through manual adjustments. EC and water levels were monitored daily, and nutrient solution adjustments were performed twelve times throughout the growing cycle to maintain optimal conditions.

Climate parameters, including air temperature and relative humidity, were measured using a DHT22 sensor (AOSONG, China), which provides a temperature accuracy of ±0.5 °C and a relative humidity accuracy of ±2% RH (maximum ±5% RH). The water level was monitored via

an HC-SR04 ultrasonic sensor (MaxBotix, USA), with a measurement accuracy of up to 3 mm. All sensors were integrated with an Arduino Mega system (Arduino, Italy), enabling continuous data logging at 15-second intervals. The data were then averaged to provide 30-minute interval measurements. A UV disinfection system (AQUANEAT, USA) was incorporated into the nutrient solution reservoir to mitigate microbial contamination that could interfere with sensor performance. Figure 18 illustrates the experimental setup alongside the range and variability of key growing parameters.

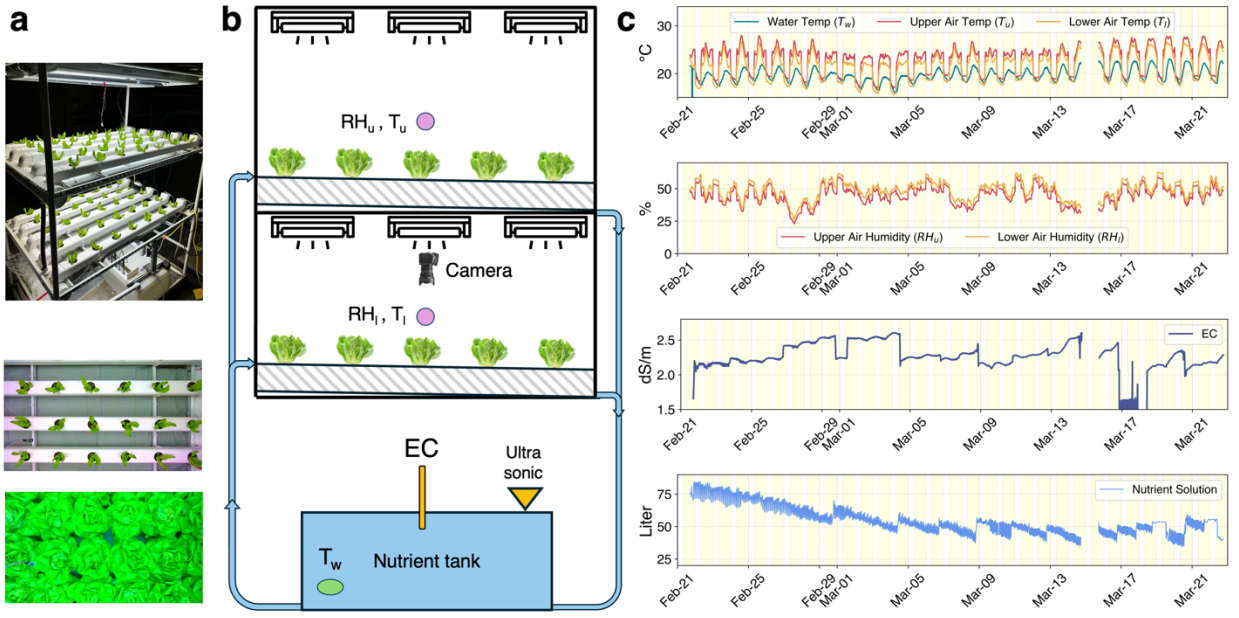


Figure 18. (a) Lettuce growth progression from transplanting (Feb 22, 2024) to near-harvest (Mar 18, 2024), (b) Schematic of the closed-loop NFT hydroponic system with two vertical tiers, (c) Time series of measured parameters: air temperature, humidity, EC, and nutrient solution volume, illustrating system dynamics and nutrient replenishment events.

4.3.2. Formulation of the dynamic model

The hydroponic system analyzed in this study was designed to integrate key physiological and biophysical processes, including lettuce biomass accumulation, evapotranspiration dynamics, and macronutrient uptake kinetics for six essential elements, i.e., N, P, K, Ca, Mg, and S. This comprehensive approach allowed for a dynamic assessment of crop performance and nutrient solution interactions under controlled environmental conditions.

4.3.2.1. Biomass and evapotranspiration models

The biomass part was the important factor that affected evapotranspiration and nutrient uptake rates. Figure 19 shows the biomass measured daily for the 18 lettuce heads in the bottom tier. A sigmoid function was used to model the average biomass per lettuce head:

$$y = \frac{y_{max}}{1 + e^{-r(t-t_0)}} \qquad (1)$$

Where y_{max} is the 163 g/head, r is the fitted growth rate (0.195), and t_0 is 17.5 and t is day after transplanting (DAT). The sigmoid model resulted in a coefficient of determination (R^2) of 0.9991 and a root mean square error (RMSE) of 1.4738 g.

The total cumulative evapotranspiration was ~173 liters. A sigmoid function was employed to model the daily crop evapotranspiration ($ET_{c,daily}$) as a function of biomass growth and maximum evapotranspiration. The model is expressed as:

$$ET_{c,daily} = \left(\frac{Biomass_i}{Total\ Biomass}\right).ET_{c,max} + E \quad (2)$$

where E is 0.3 pure evaporation measured from the surface of the tank, independent of the crop biomass.

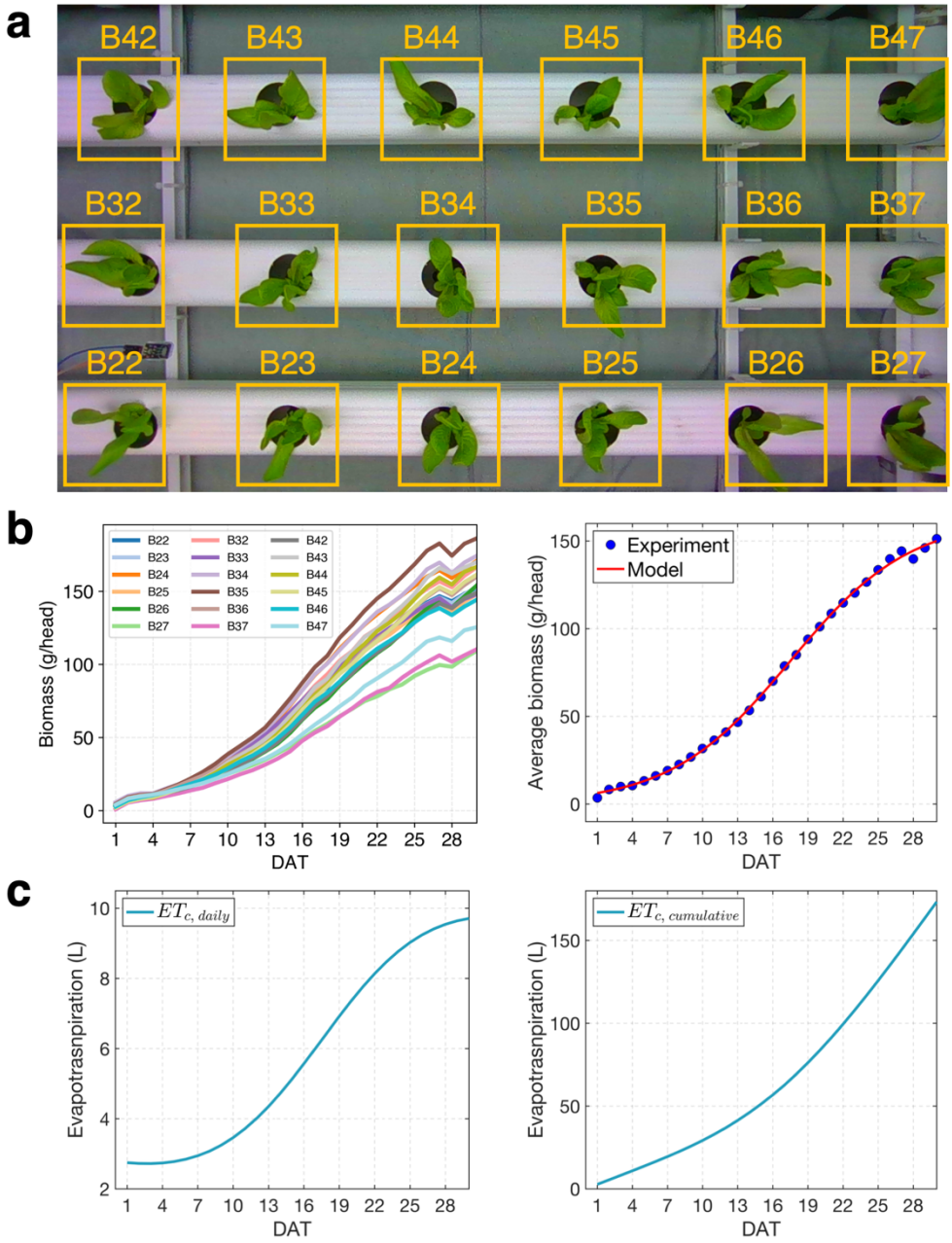


Figure 19. Monitoring and modeling of plant growth and evapotranspiration. (a) Images of the measured individual plants across three rows of NFT in the lower tier levels. (b) Biomass accumulation trends during the growing season. Left: Individual plant biomass growth curves. Right: Comparison between experimental average biomass (blue dots) and model-predicted values

(red line). (c) Temporal dynamics of evapotranspiration. Left: Daily evapotranspiration rate ($ET_{c,daily}$), and right: cumulate evapotranspiration per ~10 kg of lettuce ($ET_{c,\ cumulative}$).

4.3.3. Overview of the MPC framework

To enable autonomous fertigation that ensures crop nutrient sufficiency while minimizing water and input use, we developed OptiDose, a dynamic nutrient-water optimization framework grounded in MPC principles. MPC is particularly well-suited for complex fertigation dosing and scheduling due to its capacity to forecast system behavior, optimize control actions over a finite horizon, and incorporate updated system information in a feedback loop. In this study, we implemented a static, non-receding horizon formulation of MPC, solving the optimization problem once across a 30-day cultivation cycle. In this framework, the primary inputs include biomass accumulation, evapotranspiration dynamics, macronutrient uptake rates, and water and fertilizers unit prices. The simplified approach was considered to reduce the computation time and simplify implementation. OptiDose generated the optimal sequence of daily fertigation doses by minimizing a multi-objective cost function that jointly penalized nutrient imbalances, water overuse, and economic inefficiencies, subject to physiological and operational constraints, depicted in Figure 20. This formulation provides a benchmark for evaluating long-term fertigation strategies under ideal information conditions.

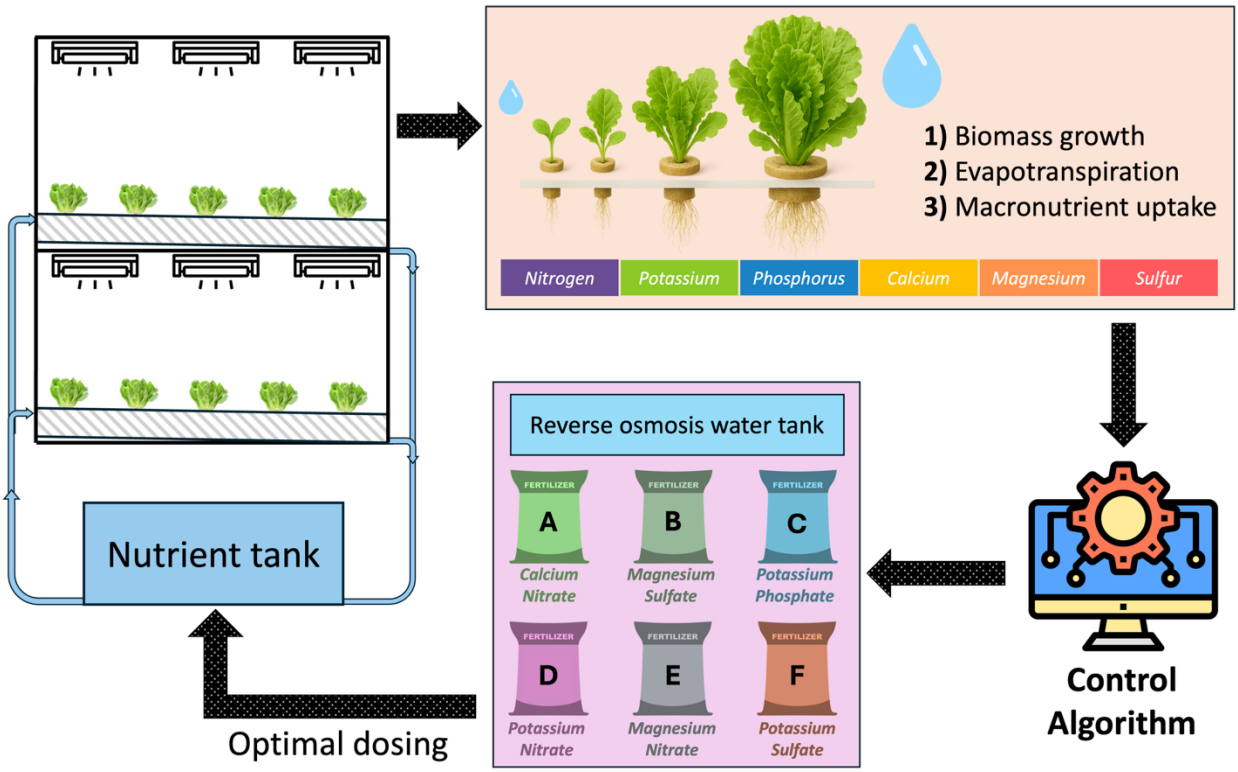


Figure 20. Overview of the proposed control framework for optimal macronutrient dosing for hydroponic production.

4.3.3.1. Problem formulation

The optimal control problem was formulated over a 30-day prediction horizon, with a daily control resolution. The system comprises seven state variables (inputs) representing cumulative mass or volume in the root zone (six macronutrients + nutrient solution water), and seven control outputs, which correspond to the application rates of six distinct fertilizers and reverse osmosis water. The model involved three primary decision variable matrices:

- $u(k) \in \mathbb{R}^7 \quad for \ k \in \{1, 2, \cdots, T\}$: dosing decisions for six fertilizers and one reserve osmosis water (7 components) across T=30 days.

- $x(k) \in \mathbb{R}^7 \ for \ k \in \{1, 2, \cdots, T\}$: state vector representing concentrations of nutrients and water volume on day k.
- $s(k) \in \mathbb{R}^{12} \ for \ k \in \{1, 2, \cdots, T\}$: the slack variables were used to manage constraint softening.

The state dynamics of the nutrient solution were governed by the discrete-time system:

$$x(k+1) = x(k) + B.u(k) - U_{uptake}(k) \quad (3)$$

where $u(k)$ denotes the control signal output, $U_{uptake}(k)$ is the crop nutrient and water uptake profile on day k, based on the reported (Albornoz & Lieth, 2016), which macronutrient uptake rates, expressed per gram of plant fresh weight, decrease as the plant matures (Table 9), $B \in \mathbb{R}^{7\times7}$ is the transformation matrix encoding fertilizer-to-nutrient mapping, it is based on the composition listed on the labels of the six selected fertilizers (Table 8), as well as the nutrient concentration of the dosing water (zero for reverse osmosis water). Rows represented the seven system states being tracked: 1) N, 2) P, 3) K, 4) Ca, 5) Mg, 6) S, and 7) water. Columns corresponded to the control inputs, which are the volumes (in liters) of six fertilizer stock solutions and water added to the system each day: 1) Tank A: Calcium Nitrate [$Ca(NO_3)_2$], 2) Tank B: Magnesium Sulfate [$MgSO_4$], 3) Tank C: Monopotassium Phosphate [KH_2PO_4], 4) Tank D: Potassium Nitrate [KNO_3], 5) Tank E: Magnesium Nitrate [$Mg(NO_3)_2$], 6) Tank F: Potassium Sulfate [K_2SO_4], 7) Tank G: Water (no nutrients). Each entry $B_{i,j}$ of the matrix quantifies the amount of nutrient i delivered per liter of solution j. These values were computed based on the fertilizer concentration (50 grams dissolved per liter of water) and the nutrient content of each fertilizer (from product specifications).

Table 8. List of selected fertilizer compounds used in the nutrient solution formulation, including their chemical names, compositions, and market price (USD/kg).

Fertilizer name	Chemical compound	Chemical composition by weight	Price (USD/kg)	Reference
Calcium Nitrate	$Ca(NO_3)_2$	$N\text{-}NO_3$: 15%, Ca: 18%	11	Green Leaf Aquariums
Magnesium Sulfate	$MgSO_4$	Mg: 10%, S: 13%		
Monopotassium Phosphate	KH_2PO_4	P_2O_5: 52%, K_2O: 34%		
Potassium Nitrate	KNO_3	$N\text{-}NO_3$: 13.7%, K_2O: 46%		
Magnesium Nitrate	$Mg(NO_3)_2$	$N\text{-}NO_3$: 11%, MgO: 16%, Mg: 9.6%		
Potassium Sulfate	K_2SO_4	K_2O: 52%, S: 18%		

The nutrient content expressed in oxide form (commonly used in fertilizer labeling) was converted to its elemental equivalent using molecular conversion factors to accurately quantify the amount of elemental nutrients delivered through each fertilizer source. The following conversion factors were applied to the fertilizers:

- Phosphorus (P): Fertilizers often report phosphorus content as phosphorus pentoxide (P_2O_5). To convert this to elemental phosphorus (P), a factor of 0.436 was used. This means that 1 gram of P_2O_5 provides 0.436 grams of elemental phosphorus.
- Potassium (K): Potassium content is typically listed as potassium oxide **(**K_2O**)**. A conversion factor of 0.830 was applied to derive the elemental potassium content, meaning that 1 gram of K_2O yields 0.830 grams of elemental potassium.

- Magnesium (Mg): Magnesium content is usually expressed as magnesium oxide (MgO). A factor of 0.603 was used, indicating that 1 gram of MgO provides 0.603 grams of elemental magnesium.

These conversion factors ensured consistency between the chemical composition of the fertilizer solutions and the required elemental concentrations defined by the nutrient bounds in the model. Applying these conversion factors allowed precise calculation of nutrient contributions from each stock solution, which were then encoded in the nutrient composition matrix B.

4.3.3.2. Objective Function

The objective function, as defined as the total cost function J, combines two main objectives, including the target for reducing water and fertilizers and constraints minimization within the allowable ranges. These two objective functions are expressed as follows:

Minimizing input costs: A linear cost is imposed on all control inputs, including fertilizers and water.

$$J_{dose} = \sum_{k=1}^{T-1} c_u \,.\, u(k) \quad (4)$$

where c_u is the cost function that includes the price of fertilizers and water (Table 8).

Minimizing constraint violations: Quadratic penalties are applied to slack variables to enforce concentration and operational constraints softly.

$$J_{slack} = \sum_{k=1}^{T} \|q_s(k) \circ s(k)\|_2^2 \quad (5)$$

where $s(k)$ is the slack variable one per nutrient, $q_s(k)$ is the weight of the normalized allowable range of each nutrient (i.e., the difference between its upper and lower concentration limits, Table 9). Using this approach, nutrients with tight allowable ranges (stricter control requirements) received higher penalties for constraint violations, while nutrients with wider acceptable ranges were penalized less. This ensured the model treats all nutrients fairly and consistently without favoring those that naturally have higher concentration values.

Thus, the total cost was:

$$J_{total} = J_{dose} + J_{slack} \quad (6)$$

The optimization framework was implemented using CasADi (Andersson et al., 2019) and solved using the interior point optimizer (IPOPT) solver in MATLAB, which is well-suited for large-scale problems involving quadratic programs.

Table 9. Suitable lower and upper concentration bounds (in parts per million, ppm) for essential macronutrients in nutrient solution formulations and uptake rate in lettuce production (Albornoz & Lieth, 2016; Mattson & Peters, 2014).

Element	**Lower Bound (ppm)**	**Upper Bound (ppm)**	**Range (ppm)**	**Plant age and uptake rate (mg/g FW.day)**			
				1-7 DAT	**8-14 DAT**	**15-21 DAT**	**21-30 DAT**
N	150	210	60	0.722	0.603	0.358	0.209
P	16	50	34	0.148	0.214	0.148	0.324
K	132	341	209	1.772	1.794	0.888	0.303
Ca	38	200	162	0.477	0.376	0.4	0.471
Mg	14	60	46	0.059	0.092	0.05	0.133
S	32	113	81	0.258	0.323	0.253	0.297

4.3.4. Scenario modeling

The OptiDose platform was used to evaluate the effect of nutrient dosing strategies on hydroponic lettuce production. Five distinct scenarios were developed and simulated to assess nutrient and water input strategies under varying operational constraints and grower behaviors.

4.3.4.1. Baseline scenario design and implementation

A baseline scenario was defined to evaluate system performance under minimal intervention by imposing hard constraints on the dosing schedule. Specifically, all nutrient and water inputs were applied exclusively on the first day of the growing cycle, with no further applications permitted for the remaining duration. Mathematically, this was formulated as:

$$u(k) = 0, \qquad for\ k = 2, \dots, 30 \qquad (7)$$

These constraints simulated a single-shot fertigation strategy, whereby all the required resources were delivered at the initiation of the growth period. The objective of this scenario is to assess the system's capacity to maintain optimal nutrient and nutrient solution levels throughout the 30-day cycle solely from an initial dose, without any subsequent control intervention. This approach serves as a reference point for evaluating the robustness of the system under extreme input limitations and highlights the temporal dynamics of nutrient availability and uptake under non-recurring fertigation regimes.

4.3.4.2. Scenario 1: Daily fertigation strategy without constraints on water and fertilizer storage capacity

In this scenario, a dynamic fertigation strategy was implemented, wherein nutrient and water inputs were optimized and applied daily over a 30-day growing cycle. The aim was to evaluate the potential of adaptive, time-resolved dosing to maintain optimal system performance while

minimizing resource input and constraint violations. Accordingly, no explicit constraints were imposed on the capacities of water and fertilizer tanks, allowing the optimization framework to explore the full range of feasible dosing strategies without infrastructural limitations.

4.3.4.3. Scenario 2: Alternate-day fertigation strategy without constraints on water and fertilizer storage capacity

In this scenario, we explored a limited intervention fertigation strategy, where nutrient and water dosing were only permitted on alternate days throughout the 30-day crop cycle. This constraint reflects practical limitations in some production systems, such as labor availability, automation cycles, or the desire to reduce the frequency of fertigation while still meeting crop needs. To implement the alternate-day dosing strategy, additional constraints were added to the model to forbid any nutrient or water application on even-numbered days:

$$u(k) = 0, \qquad for\ k = 2, 4, 6, \dots, 30 \qquad (8)$$

This ensures that dosing was only allowed on odd-numbered days (i.e., days 1, 3, 5, …, 29), effectively halving the number of intervention opportunities. This constraint was implemented directly within the optimization problem, ensuring that the solver accounted for non-dosing days when optimizing the dosing schedule. This alternate-day strategy evaluates the system's resilience to less frequent fertigation and explores whether adequate nutrient levels can still be maintained. It also supports evaluating trade-offs between dosing frequency, input use efficiency, and crop nutritional risk (deficiencies or toxicities) under limited intervention conditions.

4.3.4.4. Scenario 3: Daily fertigation under volume-constrained nutrient solution supply

In this scenario, additional hard constraints was imposed to simulate practical limitations on the volume of nutrient solution in a closed-loop reservoir system. This approach reflected real-world

CEA conditions, where fertigation systems have limited tank capacity and dosing equipment volume restrictions. This scenario introduced two critical physical constraints related to volume handling:

Reservoir volume constraint: The total solution volume in the reservoir (represented by the state variable x_7 was bound to remain within operational limits:

$$50 \leq x_7(k) \leq 100 \quad \forall k \quad (9)$$

These bounds reflect minimum and maximum allowable volumes in the system's nutrient reservoir, where 50 liters ensures sufficient solution for NFT operation, and 100 liters represents the reservoir's upper limit to avoid overflow or system stress.

Dosing volume constraint: The volume of fertilizer solution that can be injected into the reservoir per day was capped at 2 liters across all nutrient components:

$$u_{\mathrm{i}}(k) \leq 2 \quad \forall i = 1, \ldots, 6 \quad (10)$$

This constraint reflected the dosing pump's physical capacity or standard operational protocols, ensuring the system delivered precise and manageable amounts of concentrated nutrient solution daily. Water dosing (represented by $u_7(k)$) was not restricted by the 2-liter dosing cap, as it represents bulk volume that may be delivered through a different pumping system.

4.3.4.5. Scenario 4: Alternate-day dosing strategy under volume-constrained nutrient solution supply

In this scenario, the physical constraints from Scenario 3 was integrated with a reduced dosing frequency to simulate a volume-constrained alternate-day fertigation strategy. This scenario reflected a practical operational context where fertigation systems have limited reservoir capacity

and dosing hardware. Dosing is scheduled every other day—e.g., due to automation schedules, resource conservation, or labor limitations. Thus, in addition to the constraints mentioned in scenario 3, dosing was only allowed on odd-numbered days. This was enforced by explicitly constraining the dosing vector to zero on all even-numbered days:

$$u(k) = 0, \qquad for\ k = 2, 4, 6, \dots, 30 \qquad (11)$$

This constraint reduced the number of dosing opportunities by half (similar to Scenario 2), requiring the model to optimize nutrient delivery less frequently while still meeting plant demand and maintaining concentrations within target bounds.

4.3.5. Performance metrics

To assess the efficacy of each nutrient dosing strategy, a suite of quantitative performance metrics encompassing agronomic (water and fertilizer use efficiency), environmental (toxicity and deficiency index) and economic dimensions were employed.

Water use efficiency (WUE) was calculated as the ratio of total fresh biomass produced to the cumulative volume of water applied throughout the growing cycle (Eq. 12). This metric reflects the ability of each dosing scheme to translate water input into harvestable yield, thereby serving as a proxy for irrigation efficiency.

$$\text{WUE} = \frac{Fresh\ biomass\ (gr)}{Applied\ water\ (L)} \qquad (12)$$

Fertilizer use efficiency (FUE), defined as the total biomass produced per unit mass of fertilizer applied (Eq. 13), was used to evaluate the effectiveness of nutrient delivery strategies. Higher FUE values indicate more efficient conversion of fertilizer input into crop output, reducing waste and environmental loading.

$$\text{FUE} = \frac{Fresh\ biomass\ (gr)}{Applied\ fertilizer\ (gr)} \quad (13)$$

To quantify the adequacy of nutrient availability over time, we introduced two dimensionless composite indicators: the toxicity index (TI) and the deficiency index (DI). These indices reflected the cumulative deviation of nutrient concentrations from their agronomically acceptable bounds throughout the crop cycle provided in Table 9. Together, these indices provided an integrated measure of nutrient imbalance, which can lead to growth inhibition, yield reduction, or physiological stress.

$$\text{TI} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\max\left(0, \frac{C_{i,t} - C_i^{max}}{C_i^{max}}\right) \quad (14)$$

$$\text{DI} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\max\left(0, \frac{C_i^{min} - C_{i,t}}{C_i^{min}}\right) \quad (15)$$

where $C_{i,t}$ is the actual concentration of nutrient i at time step t, C_i^{max} is the upper threshold (toxicity limit) of nutrient i, C_i^{min} is the lower threshold (deficiency limit) of nutrient i, T is the total number of time steps during the growing period (T = 30), and N is the number of nutrients considered (N = 6).

In addition to resource use efficiency and nutritional balance, the economic cost of nutrient solution management was estimated per kilogram of lettuce produced. This included both water and fertilizer inputs. Water cost was calculated assuming high-purity reverse osmosis water, priced at €3.10 per cubic meter (Gil et al., 2024) and converted to USD using a fixed exchange rate of 1.10 (as listed on April 10, 2025). Fertilizer costs were based on a market price of $5.00 per pound for each of the six fertilizer types advertised by Green Leaf Aquarium (USA) and converted into

cost per gram. These metrics enabled a comprehensive evaluation of trade-offs between input efficiency, crop performance, and operational cost under each scenario.

4.4. Results and discussion

4.4.1. Temporal nutrient dynamics

Figure 21 presents the resulting nutrient concentration profiles over a 30-day growing cycle, covering six essential macronutrients: N, P, K, Ca, Mg, and S. Each subplot includes shaded areas denoting optimal concentration ranges (Table 9), within which nutrient supply is presumed to meet plant requirements without inducing toxicity or deficiency.

The implemented MPC strategy for the baseline scenario resulted in the complete depletion of all nutrients to their respective lower recommended thresholds on the harvesting day, except for Ca. Notably, the optimal initial nutrient concentrations were not aligned with the upper recommended thresholds, except for P, highlighting a more efficient starting formulation. Thus, in the absence of further dosing, the optimal initial concentrations for the nutrient solution were determined to be 172 mg/L for N, 50 mg/L for P, 26 mg/L for K, 162 mg/L for Ca, 26 mg/L for Mg, and 68 mg/L for S, prepared in 13.9 liters of water per lettuce head. To achieve this composition, 1,699.9 grams of fertilizers were dissolved in 970.8 liters of reverse osmosis water. The formulation consisted of 863.6 g calcium nitrate, 251.9 g magnesium sulfate, 211.1 g monopotassium phosphate, 260.5 g potassium nitrate, and 112.5 g potassium sulfate, with no contribution from magnesium nitrate. The formulation and corresponding fertilizer tank configuration using six selected stock solutions are illustrated in Table 8.

Element-specific analysis revealed differential sensitivity of nutrients to the imposed fertigation strategies and constraints (Figure 21). N, characterized by its high mobility and relatively wide

sufficiency window, was consistently maintained within optimal bounds across all dynamic scenarios. In contrast, Mg and S, which exhibit narrower tolerance ranges and steeper uptake kinetics, were more prone to deviations, particularly in Scenario 4, resulting in no violation (no deficiency or toxicity).

Notably, the implementation of Scenarios 1 through 4 substantially altered the initial P concentrations compared to the baseline, while concentrations of other nutrients remained largely unaffected at the outset. Alternate-day fertigation strategies (Scenarios 2 and 4) induced greater temporal variability, particularly for P, during the final week of the crop cycle. This was attributed to elevated late-stage P demand and the narrow range between P's upper and lower threshold values. Consequently, these scenarios appeared more susceptible to P imbalances—either deficiency or toxicity during the last growth stage.

Interestingly, all dynamic scenarios concluded with Ca concentrations approaching the lower sufficiency threshold, in contrast to the baseline scenario. Conversely, N concentrations at harvest consistently approached the upper threshold, suggesting preferential regulation by the MPC algorithm. K and Mg dynamics closely mirrored those observed under baseline conditions, indicating relative stability under varying dosing schedules.

Scenario 4, combining alternate-day dosing with strict volumetric constraints, exhibited the greatest amplitude in nutrient concentration fluctuations. Pronounced oscillations, particularly in P, Mg, and S, indicated overcompensation followed by rare under-dosing events. These fluctuations likely arose from the MPC framework, which encountered challenges in finding solutions when balancing tight volume limitations against nutrient sufficiency requirements under low-frequency control updates.

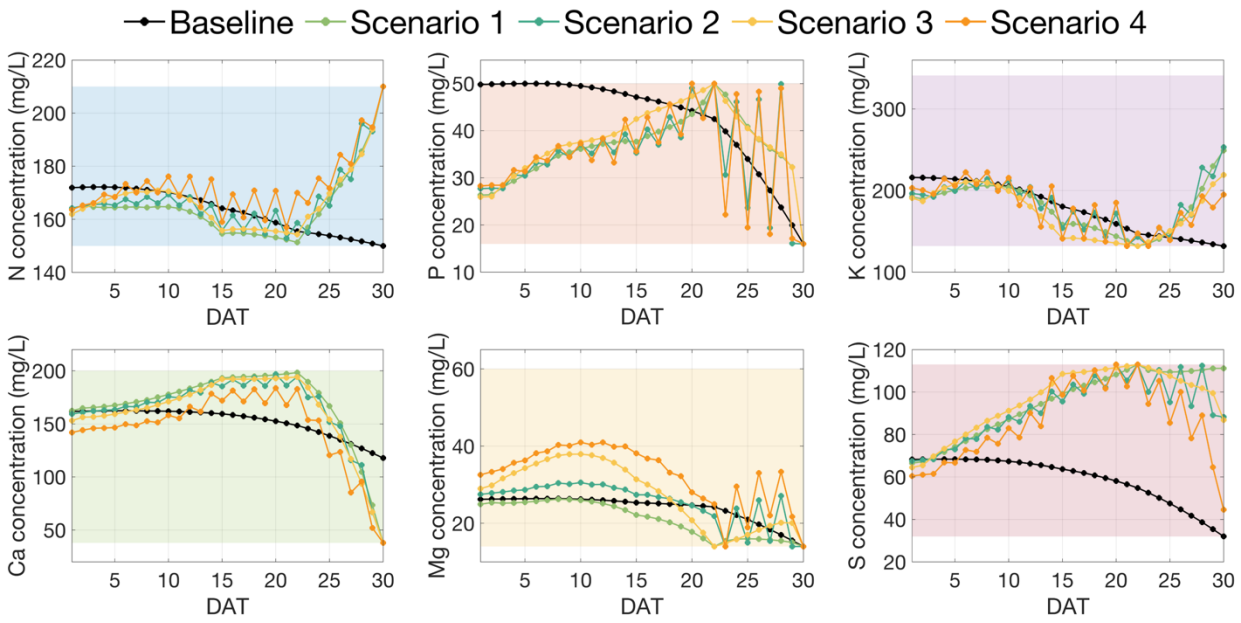


Figure 21. Temporal dynamics of nutrient concentrations in the tank under different fertigation scenarios. Simulated daily concentrations (mg/L) of six essential nutrients—nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and sulfur (S)—throughout a 30-day growing cycle (days after transplanting, DAT) of lettuce under five fertigation strategies: Baseline, Scenario 1, Scenario 2, Scenario 3, and Scenario 4.

4.4.2. Optimizing resource use under precision nutrient management scenarios

In the baseline scenario, results from the MPC framework demonstrated that allocating approximately 13.9 liters of water per lettuce head (Figure 22), combined with a nutrient formulation comprising 12.3 grams of calcium nitrate, 3.6 grams of magnesium sulfate, 3.0 grams of potassium phosphate, 3.7 grams of potassium nitrate, and 1.6 grams of potassium sulfate (totaling 24.3 grams of fertilizer; Figure 23), was sufficient to complete the lettuce growth cycle without inducing either nutrient toxicity or deficiency. Notably, magnesium nitrate was identified as a redundant input in this configuration, indicating that optimal crop performance can be achieved using a reduced set of five fertilizer types.

4.4.2.1. Water dosing

The evapotranspiration rate of 2.5 liters per lettuce head was observed. However, the baseline approach identified approximately 11.4 liters per head as the residual nutrient solution by the end of the growing cycle. While potentially reusable, this surplus introduces additional considerations related to wastewater treatment and associated costs, particularly in closed-loop hydroponic systems.

In contrast, Scenario 1 demonstrated the feasibility of producing lettuce with substantially improved water use efficiency. By starting the growing cycle with an initial water reservoir of 1.4 liters per head and incrementally replenishing it throughout the growth period (Figure 22), total water usage was reduced to 3.4 liters per head. In this case, the maximum water volume required during the cycle thus defines the optimal nutrient reservoir size under MPC utilization.

However, transitioning to an alternate-day dosing strategy with unrestricted access to water and nutrients resulted in a higher initial water requirement, which was approximately 60% greater than that observed in Scenario 1 (Figure 22). A similar trend was observed in Scenarios 3 and 4 (Figure 22), where more restrictive dosing frequencies and nutrient concentrations further shaped the dynamics of water and fertilizer utilization. Figure 22 shows that the baseline scenario relies on a static single-shot input (~970 L) at day 1, with a slow decline in nutrient solution volume and no further water adjustments, suggesting inefficiencies in both control and response. In contrast, optimized Scenarios 1–4 exhibit active and responsive water management with frequent small-volume additions. These findings highlight the substantial impact of farmer-selected nutrient management strategies on optimizing resource use efficiency.

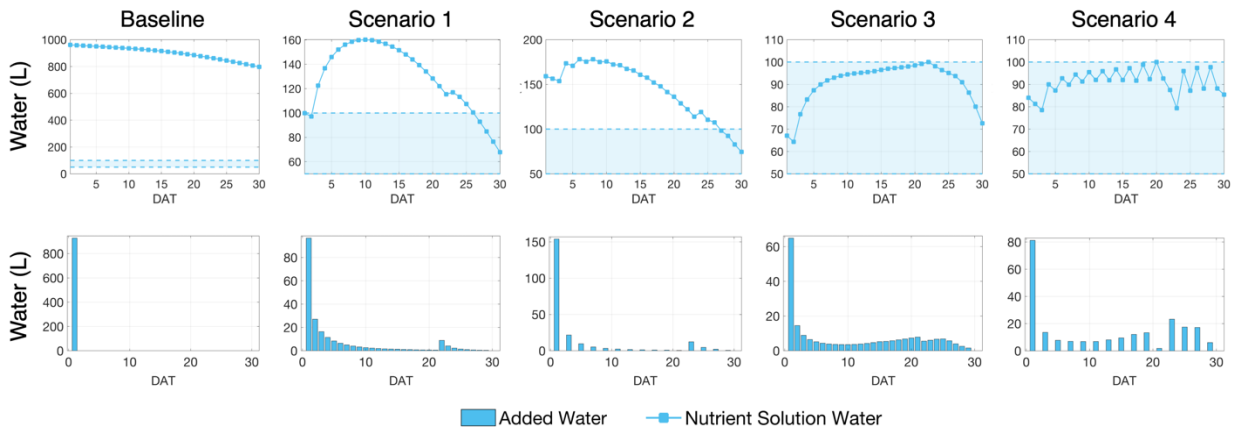


Figure 22. Dynamic water management under various nutrient delivery scenarios. Water dynamics are shown across five strategies over a 30-day cultivation. Top panels: Temporal evolution of total water volume within the system. Bottom panels: Daily volume of water added to maintain nutrient solution levels

4.4.2.2. Fertilizer dosing

Fertilizer redundancy patterns varied across the scenarios. While magnesium nitrate was dispensable in the baseline configuration, potassium-based fertilizers (e.g., potassium nitrate and potassium phosphate) emerged as redundant in Scenarios 1 and 2 (Figure 23). In contrast, under stricter nutrient input constraints in Scenarios 4 and 5, all fertilizers were utilized, suggesting a tighter margin for optimization. Across all cases, calcium nitrate consistently represented the highest share of applied nutrients, underscoring its pivotal role in lettuce cultivation (Figure 23).

Baseline management follows a static, one-time application strategy with extremely high initial input, particularly Calcium Nitrate (~864 g), and no further additions. In contrast, Scenarios 1–4 adopt dynamically adjusted dosing strategies, distributing fertilizer applications in smaller quantities throughout the crop cycle. Scenario 1 applies frequent, low-dose inputs across all nutrients, whereas Scenario 2 combines early-stage pulses with mid-to-late cycle boosts,

suggesting stage-specific nutrient demand targeting. Scenario 3 shows a gradual ramp-up of consistent low-level inputs, while Scenario 4 reflects a more adaptive pattern with variable inputs, especially after Day 20.

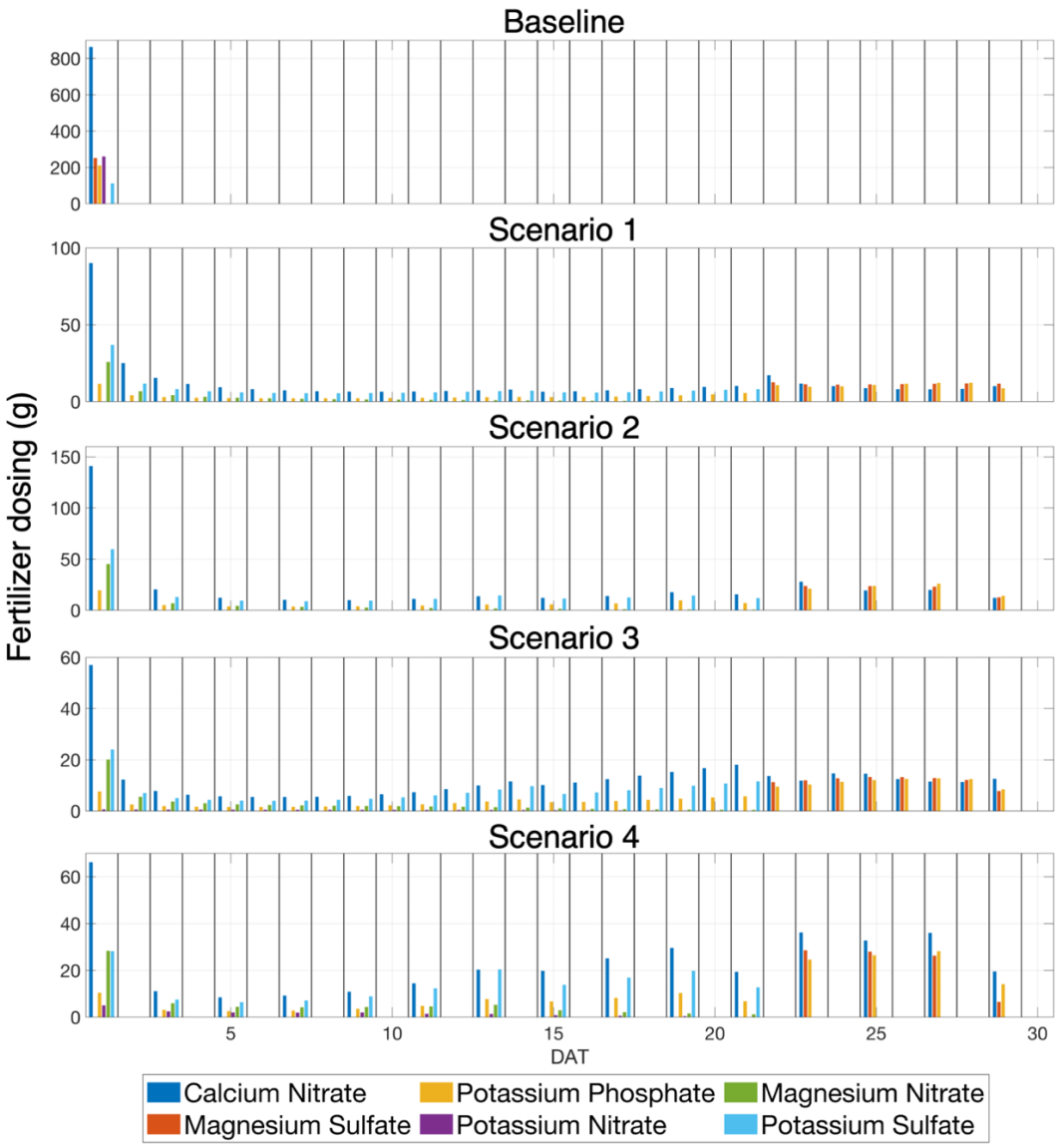


Figure 23. Temporal fertilizer dosing profiles across various nutrient delivery scenarios. Fertilizer dosing amounts (grams) are shown for six key macronutrient sources: Calcium Nitrate, Potassium Phosphate, Magnesium Nitrate, Magnesium Sulfate, Potassium Nitrate, and Potassium Sulfate—over a 30-day growing cycle.

4.4.3. Water and fertilizer consumption and efficiency

The baseline scenario, representative of conventional static fertigation practices, resulted in the highest total water (970.8 L) and fertilizer input (1699.6 g) across the 30-day cultivation period. Notably, this input strategy yielded the lowest WUE at 5.9 g/L and FUE at 6.2 g/g, indicating inefficient resource utilization and potential over-application losses. The baseline system also generated 797.5 L of wastewater, more than three times that of the following highest scenario.

In contrast, Scenarios 1–4, each applying dynamic, optimized fertigation algorithms, achieved WUE values between 30.5 and 34 g/L, an increase of nearly 6-fold compared to the baseline. Similarly, FUE values stabilized around 12.4–12.5 g/g, effectively doubling the efficiency of fertilizer conversion into biomass (Table 10). These improvements underscore the capability of dynamic nutrient delivery to align input timing with plant physiological demand, reducing wastage and leaching.

Remarkably, all scenarios—including the baseline—achieved zero average nutrient deficiency and toxicity (0 mg/L), confirming that MPC-optimized dosing strategies did not compromise crop nutrient status despite remarkably reduced input levels (Table 10). This indicates that the baseline strategy applied inputs far in excess of actual crop needs and that dynamic, model-based management can maintain optimal crop performance at far lower environmental and economic costs.

Table 10. Performance of different nutrient and water optimization strategies compared to a baseline scenario. Metrics include total water and fertilizer consumption, wastewater generation, nutrient deficiency and toxicity index, resource use efficiency, and associated costs.

Optimization strategy	Total consumed water (L)	Total wastewater (L)	Total fertilizer used (g)	DI (-)	TI (-)	WUE (g/L)	FUE (g/g)	Water cost (USD/kg)	Fertilizer cost (USD/kg)
Baseline	970.8	797.5	1699.6	0	0	5.9	6.2	0.32	1.78
Scenario 1	241.1	67.8	838.7	0	0	34	12.5	0.08	0.88
Scenario 2	247.8	74.5	846.2	0	0	32.6	12.4	0.08	0.89
Scenario 3	245.9	62.6	839.1	0	0	33	12.5	0.08	0.88
Scenario 4	258.7	85.4	847.2	0	0	30.5	12.4	0.08	0.89

4.4.4. Economic implications

Economic analysis further supported the advantages of the optimized approaches. The total cost of nutrient solution (combined water and fertilizer) in the baseline was 2.1 USD/kg, compared to just 0.96–0.97 USD/kg in optimized Scenarios 1–4, a 55% reduction in total input cost. This reduction is mainly driven by reduced fertilizer usage, with fertilizer costs halving from 1.78 USD/kg in the baseline to ~0.88–0.89 USD/kg in the other optimized treatments. Additionally, due to decreased total irrigation volumes, water costs dropped from 0.32 USD/kg in the baseline to 0.08 USD/kg in all other optimized scenarios.

4.4.5. Study limitations and future direction

The application of MPC in this study demonstrated the ability to dynamically and precisely manage water and six selected fertilizer inputs to maintain target concentrations of macronutrients (N, P, K, Ca, Mg, and S) across all five scenarios, even under strict operational constraints. Despite

this promising performance, several limitations warrant careful consideration to guide future implementations.

First, the result of MPC is sensitive to the accuracy of evapotranspiration (ET) and nutrient uptake models, which are strongly influenced by plant biomass and environmental conditions. In this study, ET was assumed to follow a fixed relationship with biomass, yet in real-world applications; ET is driven by complex interactions among temperature, vapor pressure deficit, wind speed, CO_2 concentration, and light intensity, whether from solar radiation, artificial lighting, or a combination thereof in controlled environments. These interactions not only influence water demand but also modulate biomass accumulation daily. Integrating machine learning models to predict ET and biomass accumulation using environmental data and computer vision tools may improve model fidelity and enhance real-time MPC performance.

Moreover, the uptake rates used in this study were deterministic and based on literature values that may not fully capture the inherent variability introduced by nutrient synergies and antagonisms in the hydroponic root zone. The dynamic interplay of ions in solution was not considered particularly under fluctuating environmental conditions. Nutrient absorption rates can vary substantially due to physiological stages, environmental stressors, and interactions within the root-zone solution. For example, uptake data reported by Albornoz & Lieth (2016) suggests that ammonium (NH_4) and nitrate (NO_3) exhibit the highest uncertainties in uptake, particularly in later growth stages. These nitrogen species are also more sensitive to environmental and physiological changes than other macronutrients, such as K and Mg, which displayed more consistent uptake patterns. This work aggregates nitrogen into a single form, potentially oversimplifying its behavior. Future optimization frameworks should consider the individual dynamics of NO_3 and NH_4 to enhance

dosing precision and reduce the risk of toxicity or deficiency. Also, salt precipitation within irrigation pipelines and potential clogging of emitters were not accounted for in this analysis.

While MPC adjusted dosing based on predicted uptake, it did not incorporate feedback from real-time nutrient solution analysis. Regular sampling and chemical analysis of the nutrients in hydroponic solution could allow for closed-loop feedback control, reducing drift and correcting for modeling errors, thereby improving robustness. Validation of the proposed simulations under real-world conditions is essential to evaluate the effectiveness of the recommended control ranges and optimization strategies.

The current cost function in the MPC formulation accounted for fertilizer costs and penalties for exceeding toxicity or deficiency thresholds. However, other relevant cost drivers were not considered, such as labor for dosing events or the final-day volume of unused nutrient solution. Expanding the objective function to include these parameters may allow tailoring of the optimization for economic or environmental performance, depending on the grower's priorities.

The modular nature of the MPC framework allows for extension to additional nutrients and incorporation of redundancy in fertilizer selection. Future studies should explore whether increased redundancy, through multiple sources of the same nutrient and with different nutrient combinations, either improves or compromises system efficiency, cost-effectiveness, and environmental sustainability. Furthermore, while this study focused on lettuce in a hydroponic system, the approach is extensible to other crops in CEA; it may be applicable to open-field systems with some adaptation. Successful implementation in diverse contexts will require robust modeling of evapotranspiration, crop-specific nutrient uptake, water-nutrient dynamics in the growing medium, and most importantly, feedback.

The present study utilized a single water source like reverse osmosis purified water to simulate nutrient and irrigation optimization. However, the proposed OptiDose framework is inherently adaptable and can be extended to accommodate multiple water sources with varying physicochemical properties. This feature allows for the development of more comprehensive and context-specific fertigation strategies. Future research should explore integrating water sources with diverse quality profiles, alongside their associated costs per cubic meter, to refine the optimization model further and enhance its applicability under real-world constraints.

Lastly, the OptiDose framework presented here provides a decision-support tool for determining dosing schedules that ensure consistency in nutrient supply. OptiDose offers a valuable platform for precision nutrient management in CEA systems by enabling systematic testing of hydroponic recipes within defined concentration thresholds.

4.5. Conclusions

Transitioning from conventional EC-based control to individual nutrient management is pivotal in achieving closed-loop, zero-waste hydroponic systems. This study presented a model predictive control (MPC)-based framework named OptiDose, developed to optimize water and macronutrient (N, P, K, Ca, Mg, and S) dosing in hydroponic lettuce production. By simulating five distinct farm management scenarios, OptiDose demonstrated its ability to dynamically regulate nutrient concentrations within recommended thresholds while enhancing resource use efficiency. Individualized dosing strategies increased water use efficiency from 5.9 to 34 g/L and fertilizer efficiency from 6.2 to 12.5 g/g. Additionally, the associated costs of water and fertilizer inputs were significantly reduced from 0.32 to 0.08 USD/kg and from 1.78 to 0.88 USD/kg, respectively. Across all scenarios, the MPC framework effectively minimized the objective function without inducing nutrient deficiencies or toxicities. These findings confirm that OptiDose enables precise

and adaptive fertigation control and can be flexibly applied to diverse hydroponic production systems.

4.6. References

Abioye, A. E., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., Mohammed, O. O., Otuoze, A. O., Oleolo, I. O., & Mayowa, A. (2023). Model based predictive control strategy for water saving drip irrigation. *Smart Agricultural Technology*, *4*, 100179. https://doi.org/10.1016/j.atech.2023.100179

Afram, A., & Janabi-Sharifi, F. (2014). Theory and applications of HVAC control systems – A review of model predictive control (MPC). *Building and Environment*, *72*, 343–355. https://doi.org/10.1016/j.buildenv.2013.11.016

Albornoz, F., & Lieth, J. H. (2016). Daily macronutrient uptake patterns in relation to plant age in hydroponic lettuce. *Journal of Plant Nutrition*, *39*(10), 1357–1364. https://doi.org/10.1080/01904167.2015.1109110

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., & Diehl, M. (2019). CasADi: A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, *11*(1), 1–36. https://doi.org/10.1007/s12532-018-0139-4

Bagnerini, P., Gaggero, M., & Ghio, M. (2023). Model Predictive Control for the Scheduling of Seedings in an Adaptive Vertical Farm. *2023 62nd IEEE Conference on Decision and Control (CDC)*, 4303–4308. https://doi.org/10.1109/CDC49753.2023.10384015

Bailey, B. J., Haggett, B. G. D., Hunter, A., Albery, W. J., & Svanberg, L. R. (1988). Monitoring nutrient film solutions using ion-selective electrodes. *Journal of Agricultural Engineering Research*, *40*(2), 129–142. https://doi.org/10.1016/0021-8634(88)90110-2

Bersani, C., Fossa, M., Priarone, A., Sacile, R., & Zero, E. (2021). Model Predictive Control versus Traditional Relay Control in a High Energy Efficiency Greenhouse. *Energies*, *14*(11), 3353. https://doi.org/10.3390/en14113353

Blasco, X., Martínez, M., Herrero, J. M., Ramos, C., & Sanchis, J. (2007). Model-based predictive control of greenhouse climate for reducing energy and water consumption. *Computers and Electronics in Agriculture*, *55*(1), 49–70. https://doi.org/10.1016/j.compag.2006.12.001

Bwambale, E., Abagale, F. K., & Anornu, G. K. (2023). Data-driven model predictive control for precision irrigation management. *Smart Agricultural Technology*, *3*, 100074. https://doi.org/10.1016/j.atech.2022.100074

Carotti, L., Graamans, L., Puksic, F., Butturini, M., Meinen, E., Heuvelink, E., & Stanghellini, C. (2021). Plant Factories Are Heating Up: Hunting for the Best Combination of Light Intensity, Air Temperature and Root-Zone Temperature in Lettuce Production. *Frontiers in Plant Science*, *11*, 592171. https://doi.org/10.3389/fpls.2020.592171

Catota-Ocapana, P., Minaya-Andino, C., Astudillo, P., & Pichoasamin, D. (2025). Smart Control Models Used for Nutrient Management in Hydroponic Crops: A Systematic Review. *IEEE Access*, *13*, 13070–13087. https://doi.org/10.1109/ACCESS.2025.3526171

Chen, W.-H., Mattson, N. S., & You, F. (2022). Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. *Applied Energy*, *320*, 119334. https://doi.org/10.1016/j.apenergy.2022.119334

Cho, W. J., Kim, H.-J., Jung, D. H., Kang, C. I., Choi, G.-L., & Son, J.-E. (2017). An Embedded System for Automated Hydroponic Nutrient Solution Management. *Transactions of the ASABE*, *60*(4), 1083–1096. https://doi.org/10.13031/trans.12163

Christofides, P. D., Scattolini, R., Muñoz De La Peña, D., & Liu, J. (2013). Distributed model predictive control: A tutorial review and future research directions. *Computers & Chemical Engineering*, *51*, 21–41. https://doi.org/10.1016/j.compchemeng.2012.05.011

Dela Vega, J. A., Gonzaga, J. A., & Gan Lim, L. A. (2021). Fuzzy-based automated nutrient solution control for a hydroponic tower system. *IOP Conference Series: Materials Science and Engineering*, *1109*(1), 012064. https://doi.org/10.1088/1757-899X/1109/1/012064

Ding, Y., Wang, L., Li, Y., & Li, D. (2018). Model predictive control and its application in agriculture: A review. *Computers and Electronics in Agriculture*, *151*, 104–117. https://doi.org/10.1016/j.compag.2018.06.004

Domingues, D. S., Takahashi, H. W., Camara, C. A. P., & Nixdorf, S. L. (2012). Automated system developed to control pH and concentration of nutrient solution evaluated in hydroponic lettuce production. *Computers and Electronics in Agriculture*, *84*, 53–61. https://doi.org/10.1016/j.compag.2012.02.006

García-Mañas, F., Rodríguez, F., Berenguel, M., & Maestre, J. M. (2024). Multi-Scenario Model Predictive Control for Greenhouse Crop Production Considering Market Price Uncertainty. *IEEE Transactions on Automation Science and Engineering*, *21*(3), 2936–2948. https://doi.org/10.1109/TASE.2023.3271896

Giannothanasis, E., Spanoudaki, E., Kinnas, S., Ntatsi, G., Voogt, W., & Savvas, D. (2024). Development and validation of an innovative algorithm for sodium accumulation management in

closed-loop soilless culture systems. *Agricultural Water Management*, *301*, 108968. https://doi.org/10.1016/j.agwat.2024.108968

Gil, J. D., González, R. A., Sánchez-Molina, J. A., Berenguel, M., & Rodríguez, F. (2024). Reverse osmosis desalination for greenhouse irrigation: Experimental characterization and economic evaluation based on energy hubs. *Desalination*, *574*, 117281. https://doi.org/10.1016/j.desal.2023.117281

Grewal, H. S., Maheshwari, B., & Parks, S. E. (2011). Water and nutrient use efficiency of a low-cost hydroponic greenhouse for a cucumber crop: An Australian case study. *Agricultural Water Management*, *98*(5), 841–846. https://doi.org/10.1016/j.agwat.2010.12.010

Hamidane, H., El Faiz, S., Rkik, I., El Khayat, M., Guerbaoui, M., Ed-Dahhak, A., & Lachhab, A. (2024). Constrained temperature and relative humidity predictive control: Agricultural greenhouse case of study. *Information Processing in Agriculture*, *11*(3), 409–420. https://doi.org/10.1016/j.inpa.2023.04.003

Hu, G., & You, F. (2022a). Model predictive control for greenhouse condition adjustment and crop production prediction. In *Computer Aided Chemical Engineering* (Vol. 51, pp. 1051–1056). Elsevier. https://doi.org/10.1016/B978-0-323-95879-0.50176-4

Hu, G., & You, F. (2022b). 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. https://doi.org/10.1016/j.rser.2022.112790

Jung, D. H., Kim, H. J., Choi, G. L., Ahn, T. I., Son, J. E., & Sudduth, K. A. (2015). Automated Lettuce Nutrient Solution Management Using an Array of Ion-Selective Electrodes. *Transactions of the ASABE*, 1309–1319. https://doi.org/10.13031/trans.58.11228

Jung, D.-H., Kim, H.-J., Kim, J. Y., Lee, T. S., & Park, S. H. (2020). Model Predictive Control via Output Feedback Neural Network for Improved Multi-Window Greenhouse Ventilation Control. *Sensors*, *20*(6), 1756. https://doi.org/10.3390/s20061756

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Kim, J., Kim, H.-J., Gang, M.-S., Kim, D.-W., Cho, W.-J., & Jang, J. K. (2023). Closed Hydroponic Nutrient Solution Management Using Multiple Water Sources. *Journal of Biosystems Engineering*, *48*(2), 215–224. https://doi.org/10.1007/s42853-023-00182-0

Langenfeld, N. J., Pinto, D. F., Faust, J. E., Heins, R., & Bugbee, B. (2022). Principles of Nutrient and Water Management for Indoor Agriculture. *Sustainability*, *14*(16), 10204. https://doi.org/10.3390/su141610204

Lin, D., Zhang, L., & Xia, X. (2020). Hierarchical model predictive control of greenhouse climate to reduce energy cost. *2020 2nd International Conference on Industrial Artificial Intelligence (IAI)*, 1–6. https://doi.org/10.1109/IAI50351.2020.9262227

Lin, D., Zhang, L., & Xia, X. (2021). Model predictive control of a Venlo-type greenhouse system considering electrical energy, water and carbon dioxide consumption. *Applied Energy*, *298*, 117163. https://doi.org/10.1016/j.apenergy.2021.117163

Mahmood, F., Govindan, R., & Al-Ansari, T. (2025). Efficient energy management and temperature control of a high-tech greenhouse using an improved data-driven model predictive control. *Energy Conversion and Management: X*, *26*, 100939. https://doi.org/10.1016/j.ecmx.2025.100939

Mahmood, F., Govindan, R., Bermak, A., Yang, D., & Al-Ansari, T. (2023a). Data-driven robust model predictive control for greenhouse temperature control and energy utilisation assessment. *Applied Energy*, *343*, 121190. https://doi.org/10.1016/j.apenergy.2023.121190

Mahmood, F., Govindan, R., Bermak, A., Yang, D., & Al-Ansari, T. (2023b). Greenhouse temperature regulation in the presence of uncertainties using data-driven robust model predictive control. In *Computer Aided Chemical Engineering* (Vol. 52, pp. 1591–1596). Elsevier. https://doi.org/10.1016/B978-0-443-15274-0.50253-5

Mashumah, S., Rivai, M., & Irfansyah, A. N. (2018). Nutrient Film Technique based Hydroponic System Using Fuzzy Logic Control. *2018 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 387–390. https://doi.org/10.1109/ISITIA.2018.8711201

Massa, D., Incrocci, L., Maggini, R., Bibbiani, C., Carmassi, G., Malorgio, F., & Pardossi, A. (2011). Simulation of crop water and mineral relations in greenhouse soilless culture. *Environmental Modelling & Software*, *26*(6), 711–722. https://doi.org/10.1016/j.envsoft.2011.01.004

Mattson, N. S., & Peters, C. (2014). *A Recipe for Hydroponic Success*. Inside Grower.

Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, *50*(12), 2967–2986. https://doi.org/10.1016/j.automatica.2014.10.128

Mohamed, B. T., Ahmed, A. M., Ahmed, A. A., Omar, Y. K., Makram, A. M., Fouad, K. K., Soliman, A. A., & Abo-Elmagd, A. M. (2022). Smart Hydroponic System Using Fuzzy Logic Control. *2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, 189–194. https://doi.org/10.1109/MIUCC55081.2022.9781654

Nakayama, S., Takada, T., Kimura, R., & Ohsumi, M. (2024). Model Predictive Control of Humidity Deficit and Temperature in Winter Greenhouses: Subspace Weather-Based Modelling and Sampling Period Effects. *Machines*, *12*(1), 56. https://doi.org/10.3390/machines12010056

Nizam, N. A. I. M., Chung, S. K., Yew, H. T., & Kiring, A. M. J. (2024). Hydroponic IoT-Based Control Using Adaptive Fuzzy, PID, and Consensus Control Methods for Nutrient Concentration Dosing. *Journal of Physics: Conference Series*, *2928*(1), 012001. https://doi.org/10.1088/1742-6596/2928/1/012001

Nurmahaludin, Cahyono, G. R., & Riadi, J. (2020). Nutrient Concentration Control System in Hydroponic Plants Based on Fuzzy Logic. *2020 International Conference on Applied Science and Technology (iCAST)*, 141–146. https://doi.org/10.1109/iCAST51016.2020.9557617

Padmanabha, M., Beckenbach, L., & Streif, S. (2020). Model Predictive Control of a Food Production Unit: A Case Study for Lettuce Production. *IFAC-PapersOnLine*, *53*(2), 15771–15776. https://doi.org/10.1016/j.ifacol.2020.12.204

Pandey, K., Singh, K. G., & Singh, A. (2023). Multi-Sensors based smart nutrient reuse management system for closed soilless culture under protected cultivation. *Computers and Electronics in Agriculture*, *204*, 107495. https://doi.org/10.1016/j.compag.2022.107495

Puno, J. C. V., Haban, J. J. I., Alejandrino, J. D., Bandala, A. A., & Dadios, E. P. (2020). Design of A Nutrient Film Technique Hydroponics System with Fuzzy Logic Control. *2020 IEEE*

REGION 10 CONFERENCE (TENCON), 403–408. https://doi.org/10.1109/TENCON50793.2020.9293749

Qi, D., Zhang, K., Zhao, C., Li, A., Song, B., & Li, A. (2024). Optimized model predictive control for solar assisted earth air heat exchanger system in greenhouse. *Renewable Energy*, *228*, 120582. https://doi.org/10.1016/j.renene.2024.120582

Qin, S. J., & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, *11*(7), 733–764. https://doi.org/10.1016/S0967-0661(02)00186-7

Rahman, M. K. I. A., Buyamin, S., Abidin, M. S. Z., & Mokji, M. M. (2018). Development of Automatic Mixing Process for Fertigation System in Rock Melon Cultivation. *International Journal of Electrical and Computer Engineering (IJECE)*, *8*(3), 1913. https://doi.org/10.11591/ijece.v8i3.pp1913-1919

Raviv, M., Lieth, J. H., & Bar-Tal, A. (2019). Significance of Soilless Culture in Agriculture. In *Soilless Culture* (pp. 3–14). Elsevier. https://doi.org/10.1016/B978-0-444-63696-6.00001-3

Ren, Z., Dong, Y., & Lin, D. (2022). Hierarchical model predictive control for managing agricultural greenhouse systems. *2022 4th International Conference on Industrial Artificial Intelligence (IAI)*, 1–5. https://doi.org/10.1109/IAI55780.2022.9976658

Savvas, D., Giannothanasis, E., Ntanasi, T., Karavidas, I., Drakatos, S., Panagiotakis, I., Neocleous, D., & Ntatsi, G. (2023). Improvement and validation of a decision support system to maintain optimal nutrient levels in crops grown in closed-loop soilless systems. *Agricultural Water Management*, *285*, 108373. https://doi.org/10.1016/j.agwat.2023.108373

Steidle Neto, A. J., Zolnier, S., & De Carvalho Lopes, D. (2014). Development and evaluation of an automated system for fertigation control in soilless tomato production. *Computers and Electronics in Agriculture*, *103*, 17–25. https://doi.org/10.1016/j.compag.2014.02.001

Trejo-Téllez, I., & Gómez-Merino, F. C. (2012). Nutrient Solutions for Hydroponic Systems. In T. Asao (Ed.), *Hydroponics—A Standard Methodology for Plant Biological Researches*. InTech. https://doi.org/10.5772/37578

Ullah, I., Fayaz, M., Aman, M., & Kim, D. (2022). Toward Autonomous Farming—A Novel Scheme Based on Learning to Prediction and Optimization for Smart Greenhouse Environment Control. *IEEE Internet of Things Journal*, *9*(24), 25300–25323. https://doi.org/10.1109/JIOT.2022.3196053

Ya-Gang Wang, Zhi-Gang Shi, & Wen-Jian Cai. (2001). PID autotuner and its application in HVAC systems. *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, 2192–2196 vol.3. https://doi.org/10.1109/ACC.2001.946075

Zhang, R., Zhou, J., Feng, L., & Ma, W. (2019). Study of variable liquid fertilization control system based on fuzzy PID control. *Proceedings of the Conference on Research in Adaptive and Convergent Systems*, 67–70. https://doi.org/10.1145/3338840.3355642

Chapter 5: Conclusions and future recommendations

5.1. Concluding remarks

Feeding a growing global population amid escalating resource constraints demands more sustainable and efficient food production systems. When properly managed, hydroponic farming offers a promising solution for the reliable cultivation of vegetables and leafy greens by decoupling production from external environmental variability and optimizing resource use. This dissertation aimed to harness the transformative potential of controlled environment hydroponic production to address the growing vulnerabilities of global food systems under increasing climate variability and extremes. By adopting a multi-dimensional nexus framework, this work systematically evaluated the trade-offs and synergies between greenhouse hydroponic and conventional open-field agriculture, with particular attention concerning water, nutrient, energy, and salinity dynamics. Furthermore, it introduced and validated data-driven frameworks to support the transition toward fully closed-loop hydroponic systems, highlighting the critical role of reliable sensor data, scalable computational tools, and modern machine-learning techniques for fault detection and diagnosis. Also, the findings underscore the promise of precision agriculture in closed systems, where advanced control algorithms and robust plant–water modeling can enable tailored, resource-efficient growing strategies.

Objective 1 (Chapter 2) evaluates the transition from conventional open-field lettuce production to closed-loop hydroponic systems in greenhouses through a water-salinity-nutrient-energy nexus framework. While open-field systems face persistent challenges with soil salinity and nutrient leaching, especially in arid regions, hydroponics significantly reduces these issues and achieves up to 90% water savings. However, these benefits come with increased energy demands for climate

control, depending on location. This study underscores the need for climate-specific irrigation strategies to balance water use efficiency, salinity control, and nutrient retention in open-field lettuce production. Optimized furrow irrigation schedules can reduce water withdrawals by up to 22 mm across seven Köppen climate zones by maximizing green water capture, though at the cost of higher nutrient leaching. The introduction of the irrigation salt score (ISS) enables cross-country benchmarking of salinity resilience, revealing that nations such as the UK and France are well-buffered against saline irrigation. In contrast, countries such as India, and Egypt exhibit high sensitivity due to limited rainfall and elevated evaporative demand. Drip irrigation proves superior in mitigating salinity buildup and nutrient leaching, with Egypt showing a considerable reduction in nutrient losses under drip systems compared to furrow irrigation. In contrast, furrow irrigation offers greater salt leaching potential, which arises from the precise regulation of water and solute fluxes in the root zone, governed by the interplay between irrigation strategy, atmospheric evaporative demand, and limited precipitation. However, even under open-field conditions with drip irrigation and an EC of 0.5 dS/m in Egypt, substantial nutrient leaching persists (~900 g/m^2) alongside an average water consumption of 107 L per kilogram of lettuce produced. In contrast, closed-loop hydroponic systems within greenhouses effectively eliminate leaching and reduce water use to approximately 77 L/kg, highlighting their superior efficiency in resource management. Additionally, closed-loop hydroponic systems in greenhouses present a viable long-term solution, with rooftop-harvested precipitation capable of fully meeting irrigation needs during key months in countries such as Bangladesh, the UK, Germany, and Japan. Greenhouses with climate designs tailored to lettuce production can substantially reduce atmospheric evaporative demand. When integrated with closed-loop hydroponic systems that recirculate nutrient solutions, these systems prevent nutrient discharge into groundwater or surface water bodies and avoid

detrimental impacts on soil. This combination supports environmentally sustainable production with minimal resource loss. However, the long-term sustainability of such systems is closely tied to their energy requirements for climate control. In Cwc climatic zones, the combined annual minimum energy demand, representing the lowest estimated requirement for operation, reaches 1,854 MJ/m^2, consisting of 1,150 MJ/m^2 for heating and 704 MJ/m^2 for cooling.

Despite having significant water savings in greenhouse crop production with hydroponic systems, a complete closed-loop system is almost impossible without precision in sensing nutrient composition and acidity in the solutions. Therefore, Objective 2 (Chapter 3) aims to investigate the potential of using machine learning models for fault detection and diagnosis (FDD) of EC and pH sensors in hydroponic systems, where sensor malfunctions can compromise nutrient accuracy and resource efficiency for precision hydroponic production. Among the five tested models—KNN, SVM, RF, ANN, and LSTM, RF consistently outperformed the others in accuracy. RF models delivered high fault detection accuracy, achieving 93.7% for the EC sensor and 96.5% for the pH sensor. In detailed sensor-based diagnostics, RF accurately detected EC sensor faults with 97.1% accuracy for bias, 95.1% for drift, and 100% for both spike and stuck errors. For pH sensors, the model achieved 85.3% accuracy in detecting drift and 92.1% for identifying normal operating conditions. Meanwhile, in sensor-independent detection—where EC was estimated based on climate and dosing variables—ANN models achieved the highest accuracy, correctly classifying 93.2% of readings as normal or faulty.

Precision dosing and control are critical for sustainable hydroponic production. Chapter 4 (Objective 3) introduces a model predictive control (MPC)-based framework designed to optimize the key individualized macro-nutrient (N, P, K, Ca, Mg, and S) dosing and adding makeup water water in hydroponic lettuce systems. Simulation of five distinct farm scenarios demonstrated that

the individualized dosing strategy not only maintained nutrient concentrations within agronomic target ranges but also substantially enhanced resource use efficiency. Water use efficiency increased sixfold, from 5.9 g/L under one-shot dosing to 34.0 g/L with daily dosing, while fertilizer use efficiency doubled from 6.2 to 12.5 g of biomass per gram of fertilizer applied. Furthermore, economic savings were evident: water input costs declined from 0.32 to 0.08 USD/kg of lettuce, and fertilizer costs dropped from 1.78 to 0.88 USD/kg. These results were achieved without triggering nutrient deficiency or toxicity, as the MPC successfully minimized the objective function that balances nutrient deviation penalties, cost, and system constraints. These results underscore MPC's potential to deliver precise, adaptive fertigation management and support sustainable intensification across diverse hydroponic systems.

The integration of Objectives 1, 2, and 3 collectively offers key insights into the transition toward closed-loop hydroponic systems and advances the technological framework necessary to achieve zero-discharge production systems.

5.2. Broader potential impacts

While this dissertation focuses primarily on lettuce production, the tools and methods developed across the three projects are broadly applicable to a wide range of crops cultivated in controlled environment agriculture (CEA) systems, including greenhouses and indoor vertical farms. Global scale nexus outcomes will be critical information for various stakeholders regarding comparative analysis and overview of open-field lettuce production compared to protected hydroponic production. Integrating advanced machine learning models and high-resolution data acquisition enables more precise monitoring and control of the critical aerial environment (e.g., air temperature, relative humidity, air speed, CO_2 concentration, solar radiation, etc.) and root-zone parameters (e.g., root-zone temperature, dissolved oxygen, and individual nutrient concentrations).

Since most of these variables are highly controllable in CEA systems, the proposed approaches have strong potential to improve operational efficiency, reduce resource waste, and enable crop-specific optimization strategies. These innovations lay the groundwork for scalable, data-driven management systems that support the transition to sustainable, closed-loop hydroponic food production systems. By providing actionable tools and frameworks, this research benefits growers through enhanced resource efficiency, supports researchers and students by offering replicable methodologies and open-access insights, and improves California's agricultural resilience and the global advancement of CEA. California's wide range of Köppen-Geiger climate zones—including Csa, Csb, BWh, BSk, Cfa, Dsa, Dsb, and ET—makes it an ideal proving ground for validating these approaches and scaling CEA technologies across diverse environmental and agronomic settings worldwide. FDD algorithms were developed for EC and pH sensors, but a similar approach could apply to any other sensors for monitoring nutrient solution temperature, dissolved oxygen (DO), ion concentration, etc..

5.3. Limitations and future direction

The scope of this dissertation was constrained by limited resources, which in turn imposed certain methodological and practical limitations. The limitations of each study conducted for the respective objectives of this dissertation are discussed in detail within the corresponding chapters. Overall, the assumptions made throughout the research warrant further validation, and the proposed methodologies should be applied with caution, considering their inherent limitations. A key limitation in deriving actionable insights from the results of this dissertation lies in their dependence on-farm management practices. The applicability and effectiveness of the proposed strategies are inherently tied to the specific operational conditions and management protocols employed in both open-field agriculture and CEA. As such, variability in implementation may

influence the generalizability and real-world impact of the findings. For example, the assumptions in the nexus framework presented in Chapter 1 were based on a select set of management practices to evaluate the comparative performance of open-field and greenhouse production systems. These practices represent only a subset of the wide range of strategies employed by farmers, which are often shaped by local knowledge, experience, and resource availability. As such, future studies should aim to expand this framework to incorporate a broader spectrum of management scenarios and farmer-specific preferences to enhance its applicability and robustness. Also, the assumption of a closed-concept greenhouse may overestimate the energy demand because air-side economizers could significantly reduce energy consumption, especially in cold and dry regions. Future studies should model the incorporation of air-side economizers to model greenhouse energy consumption.

This study demonstrates the strong potential of machine learning models, particularly RF, for real-time fault detection of EC and pH sensors in hydroponic systems, with applicability across various CEA settings. The proposed framework is compatible with edge devices and supports hardware-agnostic deployment, enhancing system interoperability. Future research should address variations in management practices, nutrient solution mixing dynamics, and environmental conditions to improve model robustness and practical adoption. Incorporating additional variables (e.g., buffer capacity), testing across diverse cultivars and real-world scenarios, and exploring hybrid or ensemble modeling approaches are recommended to enhance generalizability and reliability in complex agricultural systems.

One of the key advantages of applying MPC for individual nutrient management in closed-loop hydroponic systems lies in its capacity to incorporate real-time feedback from nutrient measurements. Traditionally, this feedback is obtained through laboratory analysis of the nutrient

solution, which, although accurate, can be time-consuming, costly, and impractical for frequent monitoring. As an alternative, ion-selective electrodes (ISEs) offer the potential for real-time, continuous monitoring of specific ionic concentrations, enabling dynamic control of nutrient dosing. Currently, commercially available ISEs are primarily developed for elements such as N (NO_3 and NH_4), P, K, Ca, Na, and Cl (Cho et al., 2017; Gang et al., 2025; Jung et al., 2015; Kim et al., 2023). However, further development and validation of sensors for Mg and S are recommended to achieve comprehensive nutrient profiling. Despite their potential, ISEs must be employed with caution due to common issues such as signal drift and ion interference, which may compromise measurement accuracy and, consequently, the effectiveness of the dosing strategy. Therefore, future research should focus on integrating arrays of ISEs with the MPC framework proposed in this dissertation to enhance the system's robustness and optimize nutrient delivery with greater precision. Integrating AI-driven fault detection and diagnosis for ion sensors with the MPC-based dosing algorithm is recommended for further testing, aiming to achieve zero nutrient discharge in hydroponic systems. Moreover, this work highlights the potential of MPC for greenhouse systems and resource optimization. Integrating MPC as an advanced control strategy could substantially enhance microclimate management and overall energy efficiency (Chen et al., 2022; Jung et al., 2020). Specifically, MPC offers a promising avenue for real-time, data-driven optimization of energy use and other controllable resources in CEA.

Additionally, the application of machine learning models in this dissertation primarily focused on selected system components like EC and pH sensors and macronutrients. Future research should extend these models to cover various sensors and actuators commonly used in CEA. For instance, integrating machine learning for FDD with ion-selective electrodes for real-time nutrient monitoring and climate sensors (e.g., temperature, relative humidity, and CO_2 concentration) could

enable more precise environmental control and decision-making. These enhancements will be critical for achieving closed-loop, autonomous greenhouse management systems operating efficiently and resilience.

5.4. References

Chen, W.-H., Mattson, N. S., & You, F. (2022). Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. *Applied Energy*, *320*, 119334. https://doi.org/10.1016/j.apenergy.2022.119334

Cho, W. J., Kim, H.-J., Jung, D. H., Kang, C. I., Choi, G.-L., & Son, J.-E. (2017). An Embedded System for Automated Hydroponic Nutrient Solution Management. *Transactions of the ASABE*, *60*(4), 1083–1096. https://doi.org/10.13031/trans.12163

Gang, M.-S., Kim, H.-J., Ahn, T. I., Cho, W.-J., Lee, S.-H., Lee, J. Y., & Hwang, J.-E. (2025). Development of nutrient management system based on ion-equivalent concentration ratio and ion monitoring to maintain ionic balance in closed hydroponic solution. *Computers and Electronics in Agriculture*, *237*, 110588. https://doi.org/10.1016/j.compag.2025.110588

Jung, D. H., Kim, H. J., Choi, G. L., Ahn, T. I., Son, J. E., & Sudduth, K. A. (2015). Automated Lettuce Nutrient Solution Management Using an Array of Ion-Selective Electrodes. *Transactions of the ASABE*, 1309–1319. https://doi.org/10.13031/trans.58.11228

Jung, D.-H., Kim, H.-J., Kim, J. Y., Lee, T. S., & Park, S. H. (2020). Model Predictive Control via Output Feedback Neural Network for Improved Multi-Window Greenhouse Ventilation Control. *Sensors*, *20*(6), 1756. https://doi.org/10.3390/s20061756

Kim, J., Kim, H.-J., Gang, M.-S., Kim, D.-W., Cho, W.-J., & Jang, J. K. (2023). Closed Hydroponic Nutrient Solution Management Using Multiple Water Sources. *Journal of Biosystems Engineering*, *48*(2), 215–224. https://doi.org/10.1007/s42853-023-00182-0

Appendices

All computational analyses and simulations presented in Chapters 2, 3, and 4 were performed using a high-performance workstation configured with an Intel® Xeon® w5-2465X processor (16 cores, 32 threads, 3.1 GHz), 128 GB of RAM, and a 64-bit Windows 11 Pro operating system (Build 10.0.22621.2506). Most of the analyses were conducted using MATLAB for model development and optimization, Python for data processing and statistical analysis, and ArcGIS for geospatial data visualization and spatial analysis. This integrated computational setup enabled efficient processing of large datasets, execution of memory-intensive simulations, and generation of high-resolution spatial outputs critical to the research.

Appendix A. Chapter 2 supplementary material

List of acronyms, abbreviations, symbols, and subscripts

Symbol	Description
CEA	Controlled environment agriculture
NFT	Nutrient film technique
GW	Green water
BW	Blue water
CWR	Crop water requirement

ET_o	Reference evapotranspiration
ET_c	Crop evapotranspiration
FAO	Food and agriculture organization
R_n	Net radiation
G	Soil heat flux density
e_s	Saturation vapor pressure
e_a	Actual vapor pressure
Δ	Slope of the saturation vapor pressure curve
γ	Psychrometric constant
u_2	Wind speed at 2 m height
θ	Soil moisture
Z_r	Depth of roots
P	Precipitation
I	Irrigation
DP	Deep percolation
TDS	Total dissolved solids
$C_{N,r}$	TDS of nutrients found in the root zone volume of water
$C_{N,I}$	Nutrient concentration of the applied water via irrigation
EC	Electrical conductivity
k_e	Conversion factor of EC to TDS
K_s	Salinity stress coefficient
K_y	Response of yield to salinity stress
EC_e	Electrical conductivity of the saturated extract
$EC_{e,thr}$	Threshold level for lettuce salinity tolerance
b	Yield reduction per percentage increase in EC_e
TAW	Total available water
p	Depletion fraction

ISS	Irrigation salt-score
LAI	Leaf area index
$ET_{o,g}$	Reference evapotranspiration in greenhouses
$T_{mean,g}$, $T_{max,g}$, and $T_{min,g}$	Mean, maximum, and minimum air temperature
τ	Transmissivity of greenhouse
R_a	Extraterrestrial radiation
$K_{c,g}$	Greenhouse crop coefficient
HBW	Harvested blue water
GREENHEAT	Greenhouse energy model
$Q_{c,h}$	Heating/cooling energy requirement
Q_s	Heat gain from solar radiation
Q_{sl}	Heat gain from supplemental lighting
Q_{CO2}	Heat gain from CO_2 generators
Q_m	Heat gain from motor of recirculating fans
Q_t	Heat transfer by conduction and convection
Q_i	Heat transfer due to air exchange
Q_g	Heat transfer through the greenhouse floor
Q_p	Heat transfer along the perimeter
Q_r	Heat loss through longwave radiation
Q_{ET}	Heat loss in the process of plant evapotranspiration
SWU	Specific water use
SEU	Specific energy use
BCR	Benefit-cost ratio

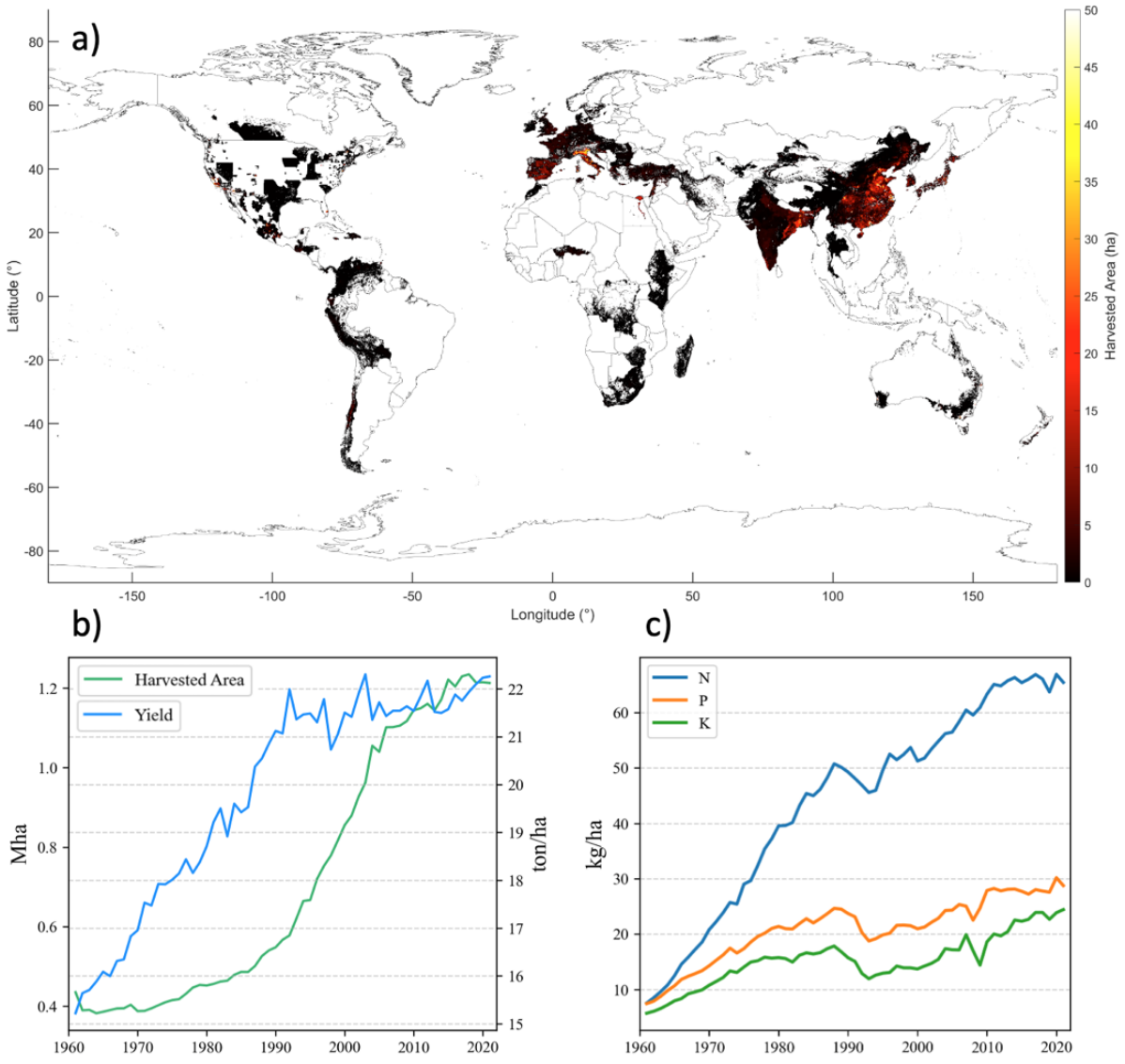


Figure 24. Cultivated area distribution (a) (Monfreda et al., 2008) and trends in lettuce yield, harvested area (b), and global nitrogen, phosphorous, and potassium fertilizer application (c).

Table 11. Lettuce parameters used in this study related to evapotranspiration, evaporation, and salinity.

Parameters	Values	Unit	Climate classification	Source
$K_{c, ini}$	0.7	-	All	Allen et al. (1998)
$K_{c, mid}$	1	-		
$K_{c, end}$	0.95	-		
K_{cb}	0.15	-		
K_e	0.8	-		
p	0.3	-		
$Z_{r, max}$	0.4	m		
$EC_{e,thr}$	1.5	dS/m		Rhoades et al., (1992)
b	12	%/(dS/m)		
K_y	1.05	-		
Initial period	25	day	BWh, BWk, BSh, BSk	Allen et al. (1998)
	27	day	Af, Am, Aw, Csa, Csb, Csc, Cwa, Cwb, Cwc, Cfa, Cfb, Cfc, Dsa, Dsb, Dsc, Dsd, Dwa, Dwb, Dwc, Dwd, Dfa, Dfb, Dfc, Dfd, ET, EF	
Development period	35	day	BWh, BWk, BSh, BSk	
	38	day	Af, Am, Aw, Csa, Csb, Csc, Cwa, Cwb, Cwc, Cfa, Cfb, Cfc, Dsa, Dsb, Dsc, Dsd, Dwa, Dwb, Dwc, Dwd, Dfa, Dfb, Dfc, Dfd, ET, EF	
Mid-season period	30	day	All	

Late-season period	25	day		
Growing period	100	day	BWh, BWk, BSh, BSk	
	105	day	Af, Am, Aw, Csa, Csb, Csc, Cwa, Cwb, Cwc, Cfa, Cfb, Cfc, Dsa, Dsb, Dsc, Dsd, Dwa, Dwb, Dwc, Dwd, Dfa, Dfb, Dfc, Dfd, ET, EF	
Planting date	01-Nov-2015	-	BWh, BWk, BSh, BSk	
	01-Apr-2015	-	Dsa, Dsb, Dsc, Dsd, Dwa, Dwb, Dwc, Dwd, Dfa, Dfb, Dfc, Dfd, ET, EF	
	01-Jan-2016	-	Af, Am, Aw, Csa, Csb, Csc, Cwa, Cwb, Cwc, Cfa, Cfb, Cfc	

Table 12. Comparison of irrigation and fertigation strategies for salinity simulation in open-field lettuce production.

Irrigation technique	**Irrigation Efficiency**	**Water salinity (dS/m)**	**Description**
Furrow	60%	0.5	Irrigation applied when soil moisture reaches water stress threshold
		1.0	
		1.5	
		2.0	
Drip	90%	0.5	Soil moisture consistently maintained at or near field capacity
		1.0	
		1.5	
		2.0	

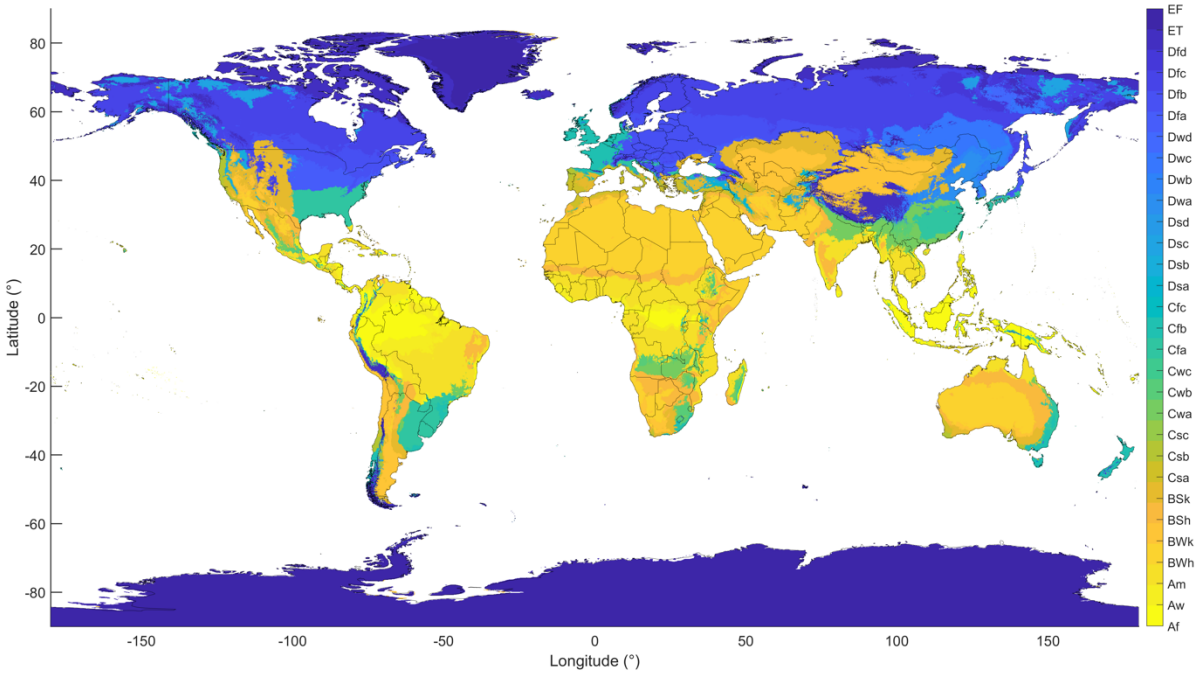


Figure 25. The global distribution of Köppen-Geiger climate classifications indicates 30 distinct climate zones worldwide (Beck et al., 2018).

Table 13. Lettuce cultivation areas across the diverse Köppen-Geiger climate classes worldwide.

Climate class	**Description**	**Harvested area (ha)**	**Harvested area (%)**
Af	Tropical, Rainforest	1,703	0.2%
Aw	Tropical, Savanna	12,416	1.5%
Am	Tropical, Monsoon	69,498	8.1%
BWh	Arid, Desert, Hot	51,023	6.0%
BWk	Arid, Desert, Cold	5,433	0.6%
BSh	Arid, Steppe, Hot	26,775	3.1%
BSk	Arid, Steppe, Cold	69,928	8.2%
Csa	Temperate, Dry summer, Hot summer	66,006	7.7%
Csb	Temperate, Dry summer, Warm summer	59,409	6.9%
Csc	Temperate, Dry summer, Cold summer	7	0.0%

Cwa	Temperate, Dry winter, Hot summer	133,571	15.6%
Cwb	Temperate, Dry winter, Warm summer	11,274	1.3%
Cwc	Temperate, Dry winter, Cold summer	9	0.0%
Cfa	Temperate, Without dry season, Hot summer	166,984	19.5%
Cfb	Temperate, Without dry season, Warm summer	42,564	5.0%
Cfc	Temperate, Without dry season, Cold summer	35	0.0%
Dsa	Cold, Dry summer, Hot summer	998	0.1%
Dsb	Cold, Dry summer, Warm summer	3,068	0.4%
Dsc	Cold, Dry summer, Cold summer	136	0.0%
Dsd	Cold, Dry summer, Very cold winter	-	0.0%
Dwa	Cold, Dry winter, Hot summer	87,034	10.2%
Dwb	Cold, Dry winter, Warm summer	7,962	0.9%
Dwc	Cold, Dry winter, Cold summer	919	0.1%
Dwd	Cold, Dry winter, Very cold winter	-	0.0%
Dfa	Cold, Without dry season, Hot summer	8,424	1.0%
Dfb	Cold, Without dry season, Warm summer	25,431	3.0%
Dfc	Cold, Without dry season, Cold summer	2,119	0.2%
Dfd	Cold, Without dry season, Very cold winter	-	0.0%
ET	Polar, Tundra	2,818	0.3%
EF	Polar, Frost	1	0.0%
Total	-	855,545	100%

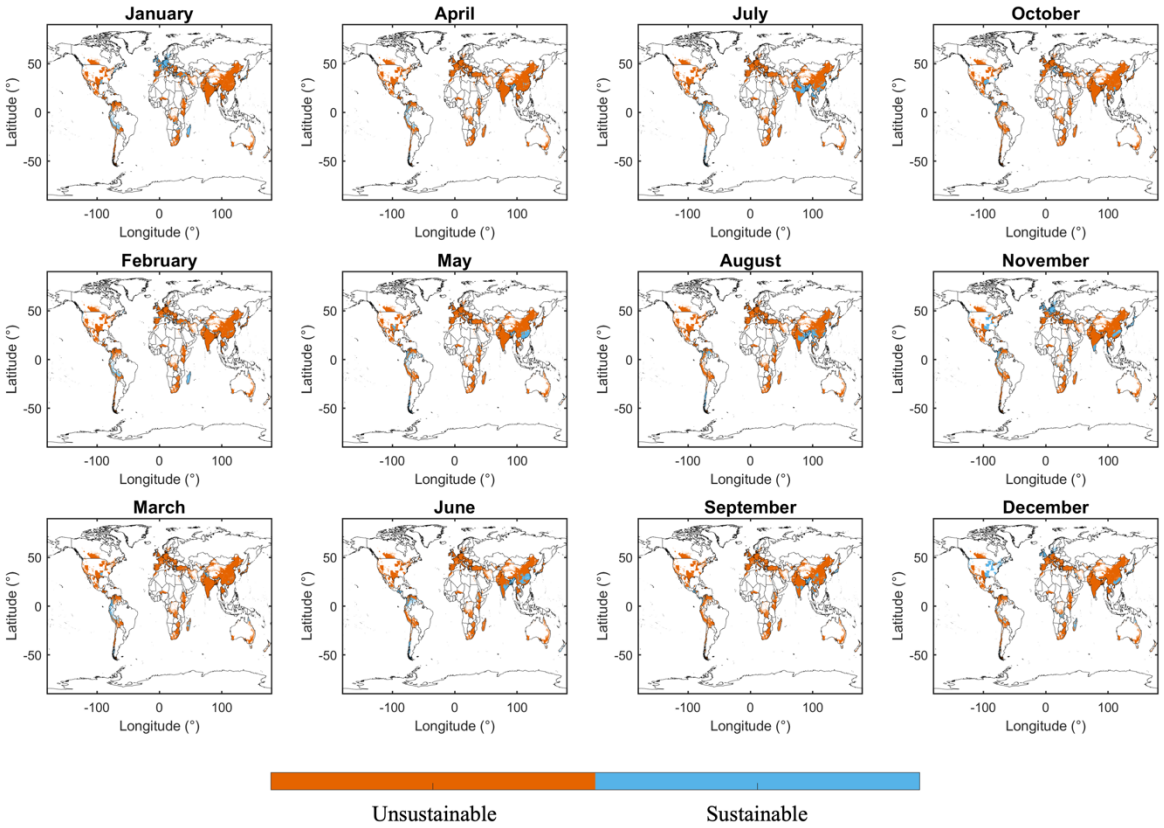


Figure 26. Comparing sustainable regions for lettuce greenhouse production by utilizing harvested blue water for hydroponics.

Table 14. List of the variables used in order to perform the analysis in this study.

Variable	Spatial resolution	Reference
Lettuce cultivated area	0.083°×0.083°	Monfreda et al. (2008)
Köppen-Geiger climate classes	0.083°×0.083°	Beck et al. (2018)
Initial soil EC	0.083°×0.083°	Shangguan et al., (2014)
Soil characteristics	0.05° × 0.05°	Köchy et al., (2015)
Extraterrestrial radiation (daily)	0.1° × 0.1°	Allen et al. (1998)
Maximum air temperature (daily)	0.1° × 0.1°	Copernicus Climate Change Service, (2019)
Minimum air temperature (daily)		
Air dew temperature (daily)		

Wind speed (daily)		
Cloud cover (daily)		
Precipitation (daily)		
Elevation	0.5° × 0.5°	JISAO, (2024)
Air temperature (hourly)	0.1° × 0.1°	Copernicus Climate Change Service, (2018)
Wind speed v (hourly)		
Wind speed u (hourly)		
Cloud cover (hourly)		
Relative humidity (hourly)		
Soil temperature level 2 (hourly)		

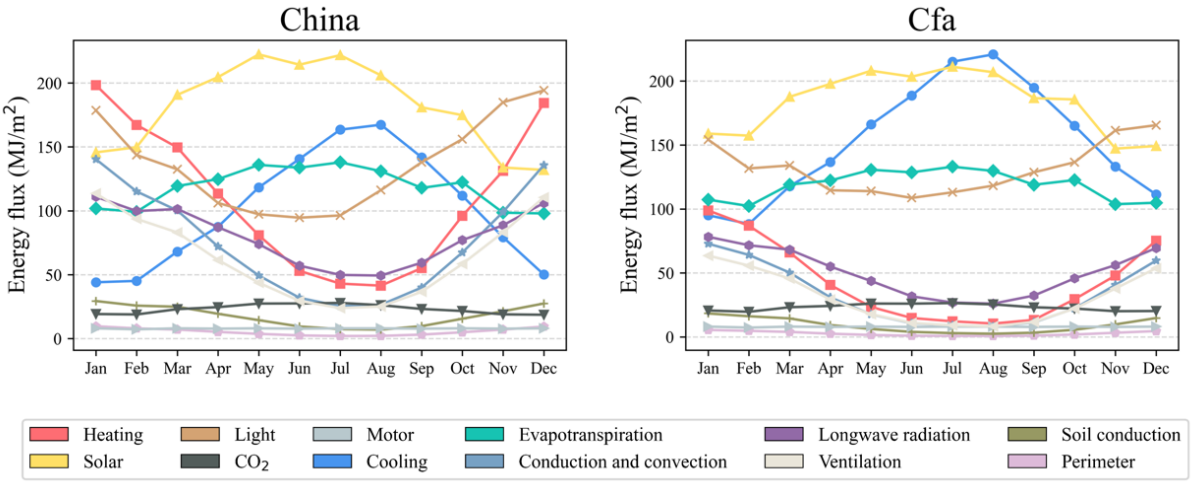


Figure 27. Monthly energy fluxes fluctuation in China (largest lettuce producer country), and in Cfa (temperate, without dry season, hot summer, the largest lettuce cultivated climate class).

Key meteorological parameters for FAO56 Penman-Monteith reference evapotranspiration calculation following Allen et al. (1998)

Slope of the Saturation Vapor Pressure Curve (Δ)

This parameter quantifies the rate of change of saturation vapor pressure with respect to the average air temperature. It is expressed as:

$$\Delta = \frac{4098 * 0.6108 \exp\left(\frac{17.27 * T_{ave}}{T_{ave} + 237.3}\right)}{(T_{ave} + 237.3)^2} \quad (S2.1)$$

where T_{ave} is the average air temperature (°C) measured at a height of 2 meters above the ground.

Net Radiation at the Reference Crop Surface (R_n)

Net radiation is the balance between net shortwave radiation (R_{ns}) and net longwave radiation (R_{nl}):

$$R_n = R_{ns} - R_{nl} \quad (S2.2)$$

Net Shortwave Radiation (R_{ns}): This is the portion of incoming solar radiation (R_{ns}) that is not reflected by the surface. The fraction of radiation reflected is known as albedo. For a reference grass crop, albedo is 0.23, giving:

$$R_{ns} = R_s * (1 - \text{albedo}) \quad (S2.3)$$

Net Longwave Radiation (R_{nl}): This represents the difference between incoming and outgoing longwave radiation, governed by the Stefan-Boltzmann law:

$$R_{nl} = \sigma * \left(\frac{T_{max,K}^2 + T_{min,K}^2}{2}\right) * \left(0.34 - 0.14 * \sqrt{e_a}\right) * \left(1.35 * \frac{R_s}{R_{so}} - 0.35\right) \quad (S2.4)$$

where:

σ represents the Stefan-Boltzmann constant (4.903×10^{-9} MJ/K^4/m^2/day)

$T_{max,K}$ and $T_{min,K}$ are the maximum and minimum temperatures in Kelvin

e_a is the actual vapor pressure

R_s denotes the measured solar radiation (MJ/m^2/day)

R_{so} is the calculated clear-sky radiation (MJ m^2/day)

The calculation of R_{so} depends on extraterrestrial solar radiation (R_a, MJ/m^2/day) and elevation (elev, m).

$$R_{so} = R_a * (0.75 + 0.00002 * \text{elev}) \quad \text{(S2.5)}$$

Extraterrestrial Radiation (R_a)

Extraterrestrial radiation represents the incoming solar radiation from celestial sources. It is calculated as:

$$R_a = \frac{24(60)}{\pi} G_{sc} d_r [\omega_s \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \sin(\omega_s)] \quad \text{(S2.6)}$$

$$d_r = 1 + 0.033 \cos\left(\frac{2\pi}{365} J\right) \quad \text{(S2.7)}$$

$$\delta = 0.409 \sin\left(\frac{2\pi}{365} J - 1.39\right) \quad \text{(S2.8)}$$

$$\omega_s = \arccos[-\tan(\varphi) \tan(\delta)] \quad \text{(S2.9)}$$

where:

G_{sc} refers to solar constant = 0.0820 MJ/m^2/min

d_r is inverse relative distance Earth-Sun (factor in adjusting the incoming solar radiation to account for the elliptical orbit of the Earth)

ω_s is sunset hour angle (rad)

φ represents latitude (rad)

δ is solar decimation (rad)

J is the Julian day of the year.

Psychrometric Constant (γ)

The psychrometric constant (γ) relates temperature and vapor pressure to atmospheric conditions:

$$\gamma = \frac{c_p * Pr}{\varepsilon * \lambda} \quad (S2.10)$$

c_p represents the specific heat capacity at constant pressure (MJ/kg/°C), with a value of 1.013×10^{-3}. λ denotes the latent heat of vaporization (MJ/kg), set at 2.45. Meanwhile, ε is the molecular weight ratio between water vapor and dry air, equating to 0.622. Atmospheric pressure (Pr, kPa) is the pressure caused by the weight of the atmosphere and varies with elevation (elev, m).

$$Pr = 101.3 * \left(\frac{293 - 0.0065 * elev}{293}\right)^{5.26} \quad (S2.11)$$

Wind Speed at 2 m (u_2)

Wind speed data is typically measured at 10 m. To estimate wind speed at 2 m, the following equation is used:

$$u_2 = u_z \frac{4.87}{\ln(67.8z - 5.42)} = u_{10} \frac{4.87}{\ln(67.8 * 10 - 5.42)} \quad (S2.12)$$

where u_{10} is wind speed measured at 10 m (m/s).

Saturation and Actual Vapor Pressure (e_s and e_a)

The saturation vapor pressure is derived from the daily maximum and minimum temperatures:

$$e_s = \frac{e^0(T_{max}) + e^0(T_{min})}{2} \quad (S2.13)$$

$$e^0(T_{max}) = 0.6108 \exp\left(\frac{17.27T_{max}}{T_{max} + 237.2}\right) \quad (S2.14)$$

$$e^0(T_{min}) = 0.6108 \exp\left(\frac{17.27T_{min}}{T_{min} + 237.2}\right) \quad (S2.15)$$

The actual vapor pressure is estimated using the dewpoint temperature:

$$e_a = e^0(T_{dew}) = 0.6108 \exp\left(\frac{17.27T_{dew}}{T_{dew} + 237.2}\right) \quad (S2.16)$$

Sub-models of the greenhouse energy balance following Ahamed et al. (2018)

The net heat gain from solar radiation (Q_s) can be quantified using the expression:

$$Q_s = \alpha_s S \quad (S2.17)$$

where α_s represented the absorptivity of the greenhouse surfaces, and S denoted the incident solar radiation using Eq. S2.18:

$$S = \sum \tau_i A_i I_i = \tau_{sr} A_{sr} I_{sr} + \tau_{sw} A_{sw} I_{sw} + \tau_{nr} A_{nr} I_{nr} + \tau_{nw} A_{nw} I_{nw} + \tau_{ew} A_{ew} I_{ew}$$
$$+ \tau_{ww} A_{ww} I_{ww} + \tau_{er} A_{er} I_{er} + \tau_{wr} A_{wr} I_{wr} \quad (S2.18)$$

The heat contribution from supplemental lighting can be calculated using the following equation:

$$Q_{sl} = WF_{hc} F_a A_f \quad (S2.18)$$

The heat generated by the motors of recirculating fans can be estimated using the following equation:

$$Q_m = N_f \frac{P_m}{E_m} F_{um} F_{lm} \quad (S2.19)$$

The heat produced by CO_2 generators can be calculated using the following equation.

$$Q_{CO_2} = 0.278 \times NHV \times MFR \times \frac{A_f}{PR} \quad (S2.20)$$

The heat transfer occurring via conduction and convection in both transparent and opaque enclosures can be determined using the following approach:

$$Q_t = (U_t A_t + U_n A_n) \times (T_i - T_o) \quad (S2.21)$$

The heat loss from greenhouses due to infiltration can be estimated using the following equation:

$$Q_i = 0.33NV(T_i - T_o) \quad (S2.22)$$

The heat transfer through the floor can be calculated using the following formula:

$$Q_g = \frac{k_s}{H} A_f (T_i - T_s) \quad (S2.23)$$

The heat loss along the perimeter can be estimated using the following formula:

$$Q_p = F_p P (T_i - T_o) \quad (S2.24)$$

The radiative heat loss transmitted from a greenhouse can be represented by the following expression.

$$Q_r = \sigma \varepsilon_c A_c F_c \left(T_i^4 - T_c^4\right) + \sigma \varepsilon_i \tau_l A_f F_{sk} (T_i^4 - T_{sk}^4) \quad (S2.25)$$

The heat consumed by plants during the process of evapotranspiration can be estimated using the following method.

$$Q_{ET} = M_T L_v \quad (S2.26)$$

Detailed descriptions of the variables used in equations (17) to (26) are provided in Ahamed et al. (2018).

Table 15. Relative changes (%) in water use, nutrient leaching, and salinity stress comparing El Niño (2015–2016) to the 5-year average (2019–2023).

Results	**Cultivation method**	**Irrigation system**	**Irrigation EC**	**Relative change (%)**
Blue water	Open field	Furrow	0.5, 1.0, 1.5, 2.0	-0.3
		Drip		-0.3
Green water		Furrow	0.5, 1.0, 1.5, 2.0	1.7
		Drip		1.8
Nutrient leaching		Furrow	0.5	-31.2
			1.0	-32.0
			1.5	-31.9
			2.0	-31.4
		Drip	0.5	-45.9
			1.0	-47.8
			1.5	48.8
			2.0	-49.4
Salinity stress coefficient		Furrow	0.5	0.5
			1.0	-4.5
			1.5	-5.9
			2.0	-7.0
		Drip	0.5	-2.4
			1.0	-5.9
			1.5	-7.8
			2.0	-8.8
Harvested blue water	Greenhouse	Hydroponic	-	-4.4
Precipitation	-	-	-	7.5

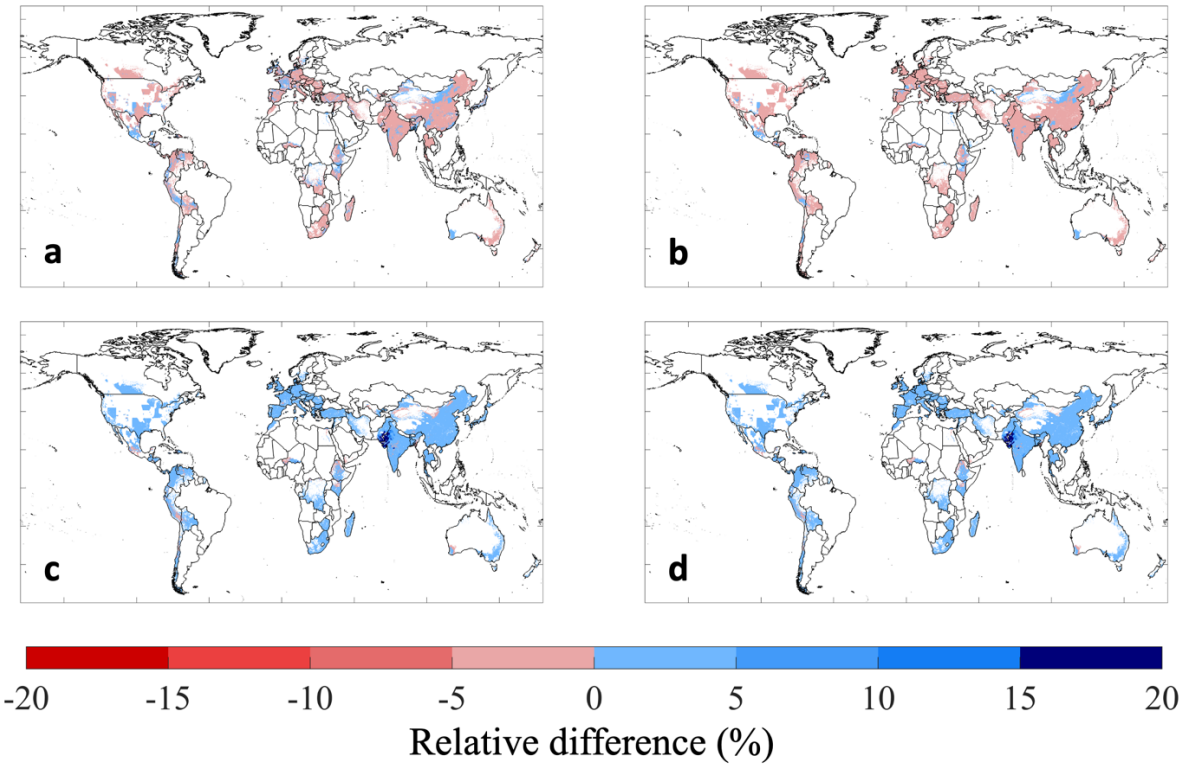


Figure 28. Global relative differences (%) in blue and green water use between the El Niño period (2015–2016) and the 5-year average (2019–2023) under different irrigation systems.

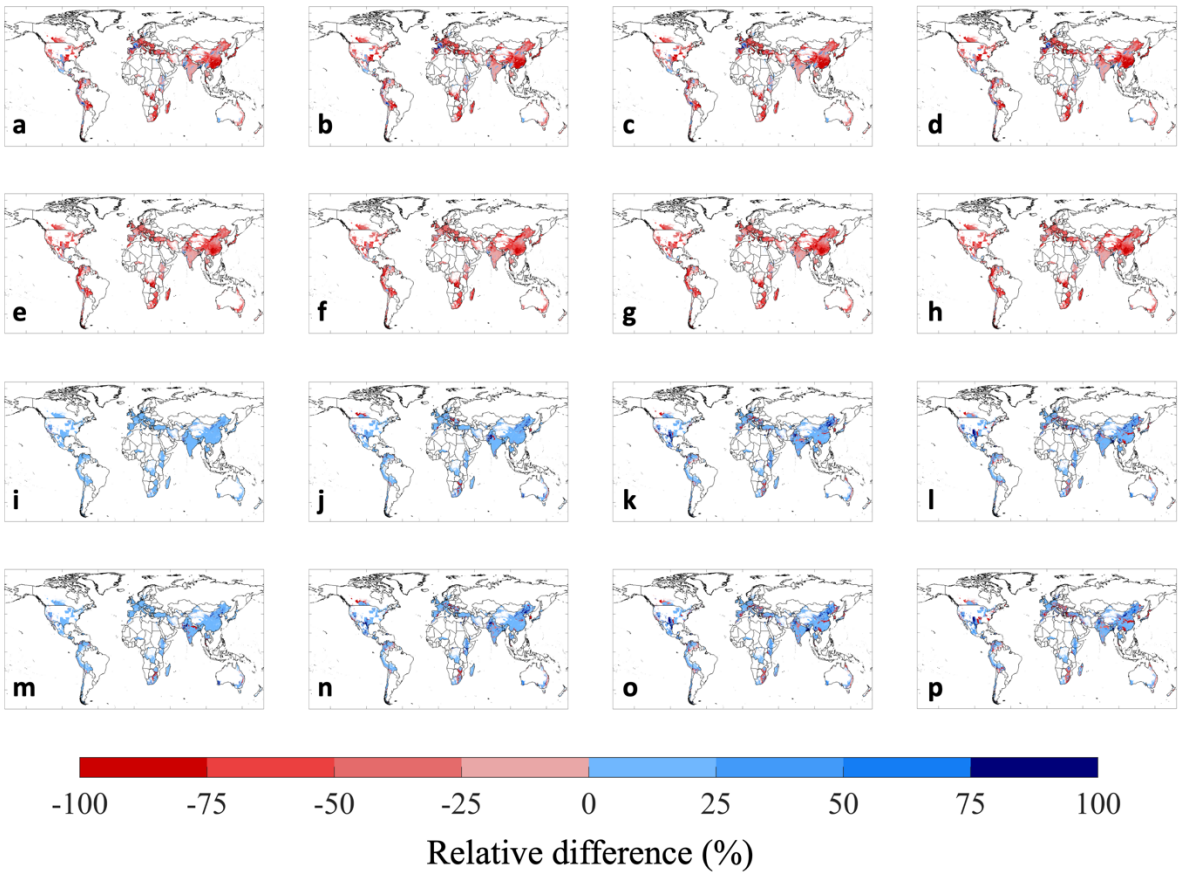


Figure 29. Relative differences (%) in nutrient leaching and salinity stress coefficient between the El Niño period (2015–2016) and the 5-year average (2019–2023) under different irrigation systems and water salinities.

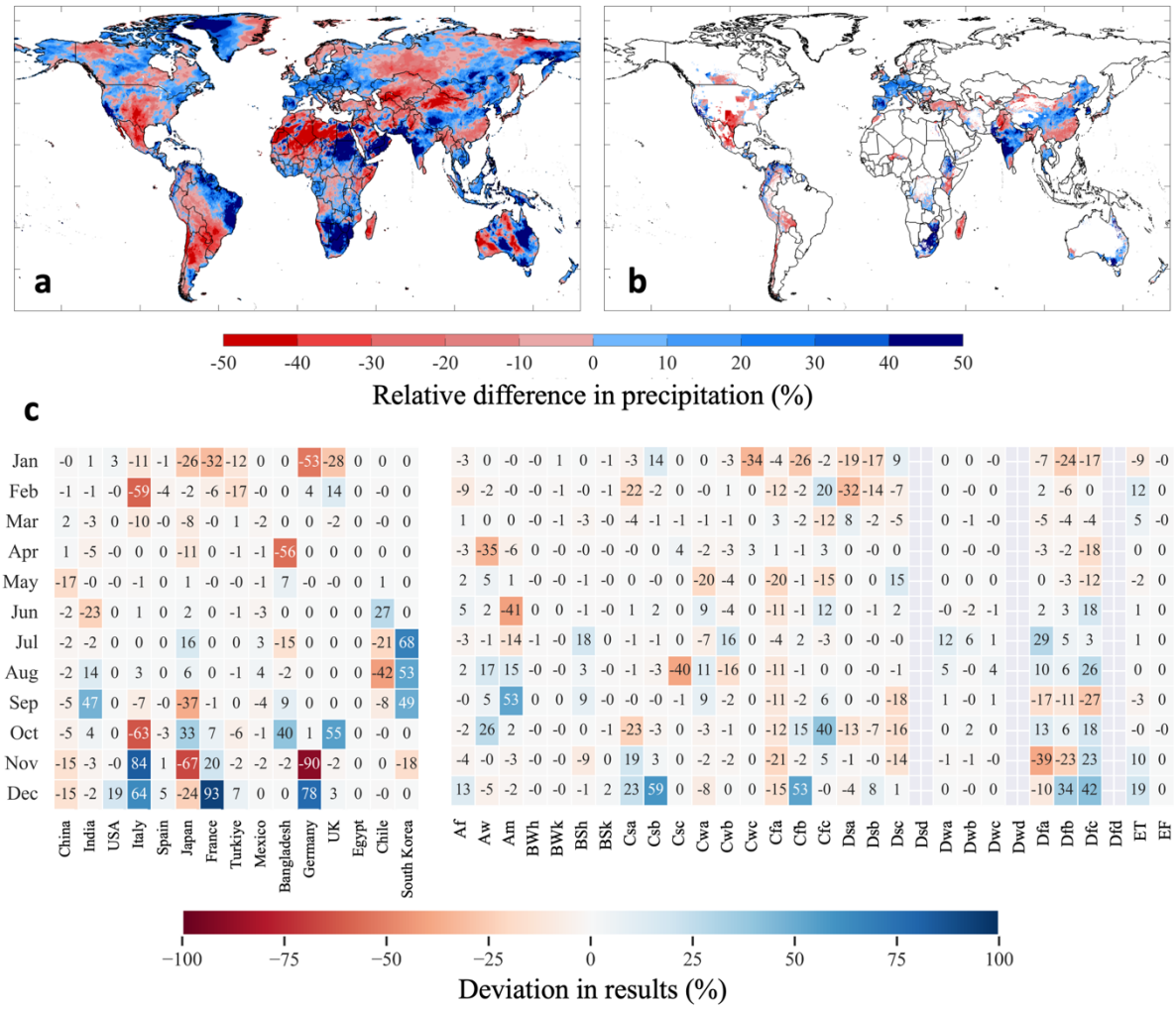


Figure 30. Relative deviation in precipitation (%) between the El Niño period (2015–2016) and the 5-year average (2019–2023).

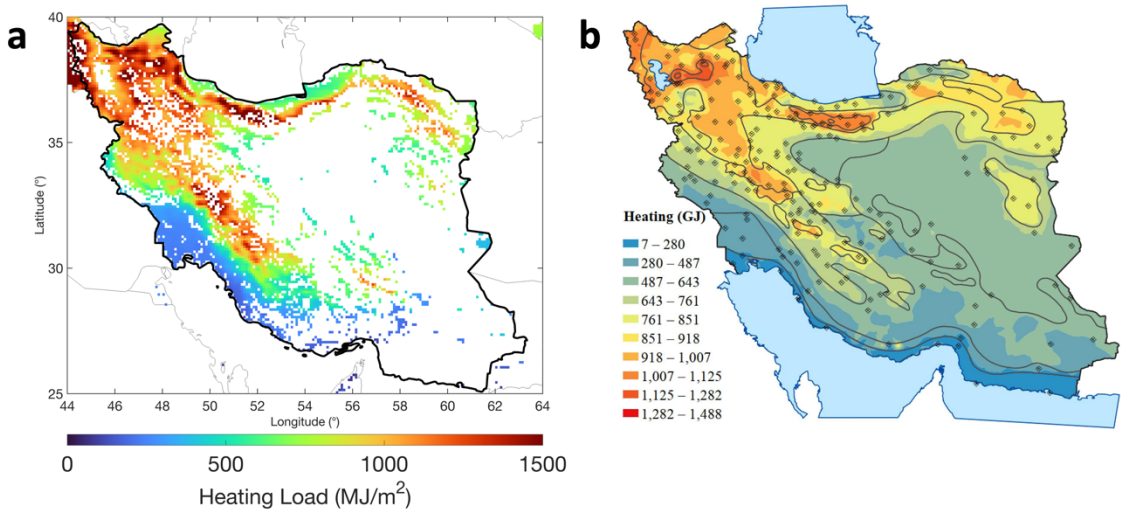


Figure 31. Spatial distribution of annual heating energy demand across Iran. (a) Estimated annual heating load (MJ/m2) based on our high-resolution thermal energy balance model; (b) corresponding annual heating demand (GJ) estimated by Nasrollahi et al. (2021).

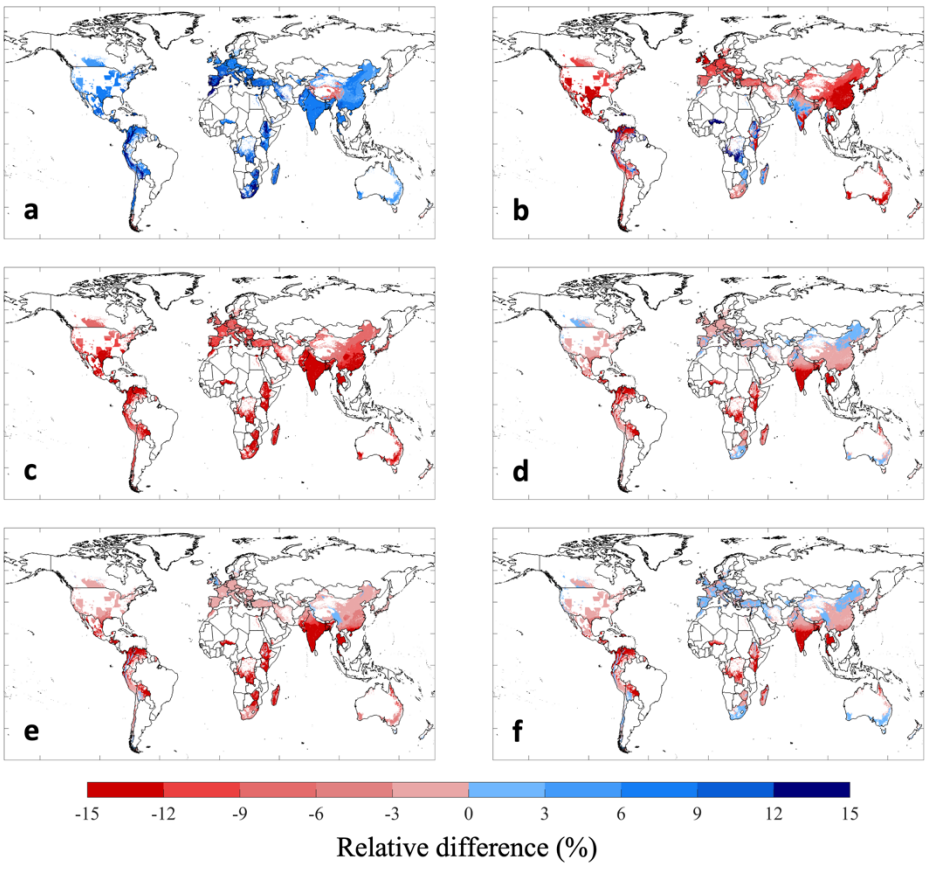


Figure 32. Global impact of narrowing greenhouse temperature setpoints on energy demand. Relative difference (%) in (a) cooling, (b) heating energy requirements, (c) longwave radiation, (d) conduction and convection, (e) greenhouse floor, and (f) infiltration.

Table 16. Impact of narrowing greenhouse temperature setpoints on energy requirements across global climate zones.

Classification	Cooling energy requirement			Heating energy requirement			Total energy requirement		
	20°C - 28°C (MJ/m^2)	22°C - 26°C (MJ/m^2)	Relative difference (%)	20°C - 28°C (MJ/m^2)	22°C - 26°C (MJ/m^2)	Relative difference (%)	20°C - 28°C (MJ/m^2)	22°C - 26°C (MJ/m^2)	Relative difference (%)
Af	2123	2328	10	79	82	4	2202	2410	9
Aw	2328	2525	8	64	63	-2	2392	2587	8
Am	2505	2710	8	96	98	2	2601	2808	8
BWh	2469	2665	8	199	188	-5	2668	2853	7

BWk	1269	1365	8	1119	1010	-10	2388	2375	-1
BSh	2400	2594	8	200	188	-6	2600	2782	7
BSk	1272	1367	8	1077	968	-10	2349	2336	-1
Csa	1407	1551	10	678	620	-9	2086	2171	4
Csb	1285	1403	9	768	686	-11	2053	2089	2
Csc	710	748	5	1468	1306	-11	2178	2054	-6
Cwa	2096	2265	8	352	317	-10	2448	2582	5
Cwb	1444	1600	11	637	586	-8	2081	2186	5
Cwc	713	795	11	1141	1014	-11	1854	1809	-2
Cfa	1832	1954	7	519	438	-16	2352	2393	2
Cfb	971	1046	8	1018	896	-12	1990	1942	-2
Cfc	555	548	-1	1593	1392	-13	2147	1941	-10
Dsa	1137	1250	10	1140	1045	-8	2277	2295	1
Dsb	881	954	8	1455	1323	-9	2336	2277	-3
Dsc	605	624	3	1883	1699	-10	2489	2323	-7
Dsd	-	-	-	-	-	-	-	-	-
Dwa	1335	1405	5	1130	998	-12	2465	2403	-3
Dwb	945	991	5	1528	1372	-10	2473	2363	-4
Dwc	698	721	3	1897	1720	-9	2595	2440	-6
Dwd	-	-	-	-	-	-	-	-	-
Dfa	1126	1206	7	1135	1014	-11	2261	2219	-2
Dfb	824	878	6	1486	1343	-10	2311	2221	-4
Dfc	631	652	3	1888	1716	-9	2519	2368	-6
Dfd	-	-	-	-	-	-	-	-	-
ET	409	407	0	2278	2072	-9	2687	2479	-8
EF	102	83	-19	3867	3642	-6	3969	3724	-6

Appendix B. Chapter 3 supplementary material

Table 17. Hyperparameter tuning results for five machine learning models in classifying EC and pH sensors for fault detection and diagnosis.

Classification Method	Hyperparameters and range	Optimal values			
		EC, detection	EC, diagnosis	pH, detection	pH, diagnosis
SVM	C: [0.1, 1, 10, 100, 1000] Kernel: [linear, poly, rbf, sigmoid] Gamma: [scale, auto]	1000, rbf, scale	1000, rbf, scale	1000, rbf, auto	1000, rbf, auto
KNN	n_neighbors: [3, 5, 7, 9, 11] weights: [uniform, distance] metric: [euclidean, manhattan]	3, distance, manhattan	9, distance, manhattan	11, distance, manhattan	3, distance, manhattan
ANN	Hidden Layers: {[24, 12, 6], [36, 18, 6], [32, 32, 32], [32, 16, 8], [64, 64, 64], [64, 32, 16], [128, 128, 128], [128, 64, 32], [256, 256, 256], [256, 128, 64],	Hidden Layers: [128, 64, 32], Epochs: 1000, Batch Size: 32, Dropout Rate: 0.0, Learning Rate: 0.0001	Hidden Layers: [128, 128, 128], Epochs: 1000, Batch Size: 64, Dropout Rate: 0.0, Learning Rate: 0.0001	Hidden Layers: [256, 128, 64], Epochs: 1000, Batch Size: 64, Dropout Rate: 0.0,	Hidden Layers: [128, 128, 128], Epochs: 500, Batch Size: 64, Dropout Rate: 0.0, Learning Rate: 0.0001

	[512, 512, 512], [512, 256, 128]}, Epochs: [100, 200, 500, 1000], Batch Size: [32, 50, 64] Dropout Rate: [0.0, 0.2, 0.5] Learning Rate: [0.0001, 0.001, 0.01]			Learning Rate: 0.0001	
RF	n-estimator Range: [100, 200, 500, 1000], Max Features Range: ['log2', 'sqrt'], Min Samples Split Range: [2, 3, 5, 10]. Min Sample Leaf Range: [1, 2, 4]	n-estimator Range: 100, Max Features Range: 'log2', Min Samples Split Range: 3, Min Sample Leaf Range: 1	n-estimator Range: 100, Max Features Range: 'log2', Min Samples Split Range: 3, Min Sample Leaf Range: 1	n-estimator Range: 100, Max Features Range: 'sqrt', Min Samples Split Range: 5, Min Sample Leaf Range: 1	n-estimator Range: 100, Max Features Range: 'sqrt', Min Samples Split Range: 10, Min Sample Leaf Range: 2
LSTM*	Hidden Layers: {[24, 12, 6, 6], [36, 18, 6, 6], [32, 32, 32, 32], [32, 16, 8, 8], [64, 64, 64, 64], [64, 32, 16, 16],	Hidden Layers: [32, 32, 32], Epochs: 1000, Batch Size: 64,	Hidden Layers: [64,64,64], Epochs: 500, Batch Size: 32, Dropout Rate: 0.0, Learning Rate: 0.0001	Hidden Layers: [32, 32, 32], Epochs: 1000, Batch Size: 64, Dropout Rate: 0.0,	Hidden Layers: [128,128,128], Epochs: 1000, Batch Size: 64,

	[128, 128, 128, 128], [128, 64, 32, 32], [256, 256, 256, 256], [256, 128, 64, 64], [512, 512, 512, 512], [512, 256, 128, 128]} Batch Size: [32, 50, 64] Dropout Rate: [0.0, 0.2, 0.5] Learning Rate: [0.0001, 0.001, 0.01]	Dropout Rate: 0.0, Learning Rate: 0.0001		Learning Rate: 0.0001	Dropout Rate: 0.0, Learning Rate: 0.0001

*Note: The LSTM method contains an extra layer of long short-term memory architecture.

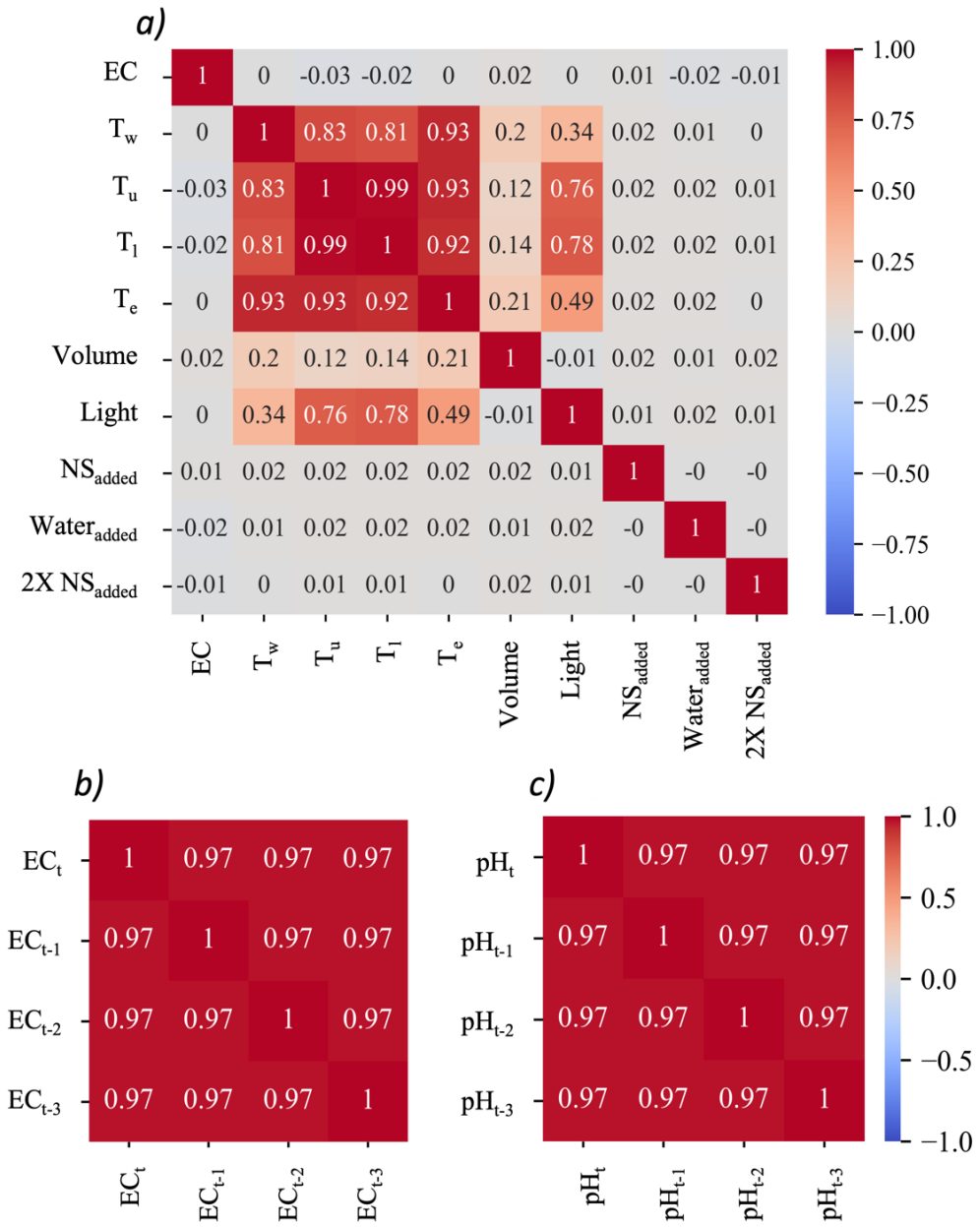


Figure 33. Pearson correlation coefficients between climate variables and dosing pulses with EC (a) and between EC and pH with three previous observations (b, c).

Table 18. Performance metrics (precision, recall, f1-score, and overall accuracy) in predicting EC and pH across different sensor fault detection models (KNN, SVM, RF, ANN, LSTM).

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
KNN	fault	0.919	0.853	0.885	0.890	0.987	0.897	0.940	0.942
	normal	0.863	0.925	0.893		0.906	0.988	0.945	
SVM	fault	0.797	0.769	0.783	0.786	0.986	0.891	0.936	0.939
	normal	0.777	0.804	0.790		0.901	0.987	0.942	
RF	fault	0.912	0.967	0.939	0.937	0.946	0.986	0.966	0.965
	normal	0.965	0.907	0.935		0.985	0.944	0.964	
ANN	fault	0.858	0.988	0.919	0.912	0.908	0.990	0.947	0.944
	normal	0.986	0.837	0.905		0.989	0.900	0.942	
LSTM	fault	0.855	0.992	0.919	0.912	0.894	0.984	0.937	0.936
	normal	0.990	0.832	0.904		0.982	0.883	0.930	

Table 19. Performance metrics (precision, recall, and f1-score) in predicting EC and pH across different sensor fault diagnosis models: (a) KNN, (b) SVM, (c) RF, (d) ANN, and (e) LSTM.

(a)

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
kNN	bias	0.667	0.71	0.688	0.777	0.788	0.842	0.814	0.845
	drift	0.711	0.703	0.707		0.728	0.765	0.746	
	normal	0.696	0.677	0.686		0.882	0.643	0.744	
	precision	0.883	0.811	0.846		0.906	0.788	0.843	
	spike	0.977	0.992	0.985		0.945	1.000	0.972	
	stuck	0.960	1.000	0.979		0.817	1.000	0.899	

(b)

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
SVM	bias	0.548	0.488	0.516	0.762	0.737	1	0.849	0.859
	drift	0.726	0.652	0.687		0.928	0.634	0.754	
	normal	0.557	0.673	0.609		0.863	0.664	0.75	
	precision	0.887	0.801	0.842		0.944	0.856	0.898	
	spike	0.989	0.988	0.988		0.981	1	0.991	
	stuck	0.909	1	0.952		0.791	1	0.883	

(c)

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
RF	bias	0.906	0.971	0.938	0.930	0.522	0.489	0.505	0.819
	drift	0.956	0.927	0.941		0.902	0.853	0.877	
	normal	0.869	0.827	0.847		0.948	0.921	0.934	
	precision	0.971	0.935	0.952		0.925	0.878	0.901	
	spike	1.000	1.000	1.000		1.000	0.998	0.999	
	stuck	0.873	0.91	0.891		0.635	0.762	0.693	

(d)

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
ANN	bias	0.983	0.962	0.972	0.879	0.785	0.459	0.579	0.817
	drift	0.906	0.952	0.928		0.866	0.773	0.817	
	normal	0.951	0.502	0.657		0.959	0.891	0.924	
	precision	0.972	0.883	0.925		0.833	0.794	0.813	
	spike	0.979	0.984	0.982		0.884	0.976	0.928	

	stuck	0.654	1.000	0.791		0.649	0.997	0.786	

(e)

		EC				pH			
		precision	**recall**	**f1-score**	**Overall accuracy**	**precision**	**recall**	**f1-score**	**Overall accuracy**
LSTM	bias	0.991	0.954	0.972	0.877	0.730	0.465	0.568	0.801
	drift	0.903	0.948	0.925		0.878	0.700	0.779	
	normal	0.934	0.511	0.660		0.960	0.895	0.926	
	precision	0.978	0.874	0.923		0.812	0.788	0.800	
	spike	0.973	0.983	0.978		0.871	0.938	0.903	
	stuck	0.651	1.000	0.789		0.628	0.997	0.771	

Table 20. Pseudocode of the approach for sensor fault detection and diagnosis.

```
1: START
# Calibration of sensors
2: CALIBRATE pH sensor with two points: 4.0 and 10.0
3: CALIBRATE EC sensor with two points: 1 and 50 dS/m
# Initial setup
4: MEASURE EC and pH using Vernier sensors
5: MEASURE Temperature (T) and Relative Humidity (RH) using DHT22 sensors
6: MEASURE Water Level using HC-SR04 sensor
7: SETUP Arduino Mega and Uno for data recording
8: SET data recording interval to 5 seconds
# Data collection
8: COLLECT data continuously
9: STORE EC, pH, T, RH, and water level in memory
# Fault generation and classification
10: DEFINE nominal accuracy:
11:    - EC: ±3% - pH: ±0.2
12: DEFINE fault types:
```

13: - Drift: $s(t) = v(t) + \eta + d(t)$, $d(t) = d(t-1) + \vartheta$

14: - Bias: $s(t) = v(t) + \eta + \vartheta$

15: - Precision Degradation: $s(t) = v(t) + \eta + \gamma$, $\gamma \sim N(0, \delta_\gamma^2)$, $\delta_\gamma^2 \gg \delta_\eta^2$

16: - Spike: $s(t) = v(t) + \eta + \vartheta(t)$

17: - Stuck: $s(t) = \vartheta$ (sensors' initial readings)

18: FOR each observation in synthetic faults:

19: CALCULATE absolute deviation:

20: deviation = $|s(t) - v(t)|$

21: IF deviation ≤ nominal accuracy:

22: ASSIGN observation as "normal"

23: ELSE:

24: ASSIGN observation as "faulty"

 # Feature analysis

25: CALCULATE Pearson Correlation Coefficient (ρ) for EC and influencing variables

26: IDENTIFY significant variables for fault detection

 # Machine Learning Model Training

27: APPLY data pre-processing:

28: - AVERAGE data to 5-minute intervals

29: - HANDLE data imbalance using SMOTE

30: - SCALE variables

31: - SPLIT data into training (70%) and testing (30%) sets

32: - IMPLEMENT 4-fold cross

33: TRAIN models using historical data:

34: - SVM, (SVR)

35: - KNN

36: - NN

37: - RF

38: - LSTM

39: PERFORM hyperparameter tuning with grid search

 # Fault Detection and Diagnosis (FDD)

39: FOR each model:

40: EVALUATE model performance using:

41: - Confusion Matrix

42: - Accuracy
43: - Precision
44: - Recall
45: - F1-score
46: COMPARE predicted EC with manufacturer's nominal accuracy (±3%)
47: DETERMINE sensor-independent fault detection accuracy
48: IMPLEMENT 4-fold cross
 # Output results
49: DISPLAY performance metrics
50: GENERATE confusion matrices and visualizations for fault detection and system performance
51: END

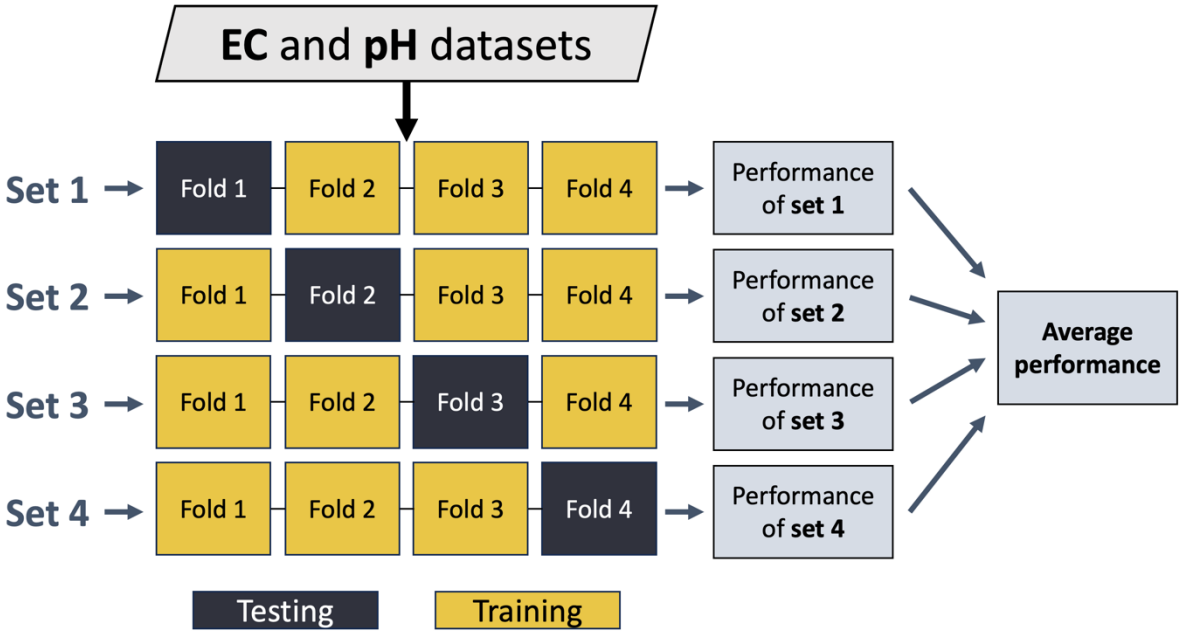


Figure 34. Schematic representation of k-fold cross-validation for evaluating model performance using EC and pH datasets.

Appendix C. Chapter 4 supplementary material

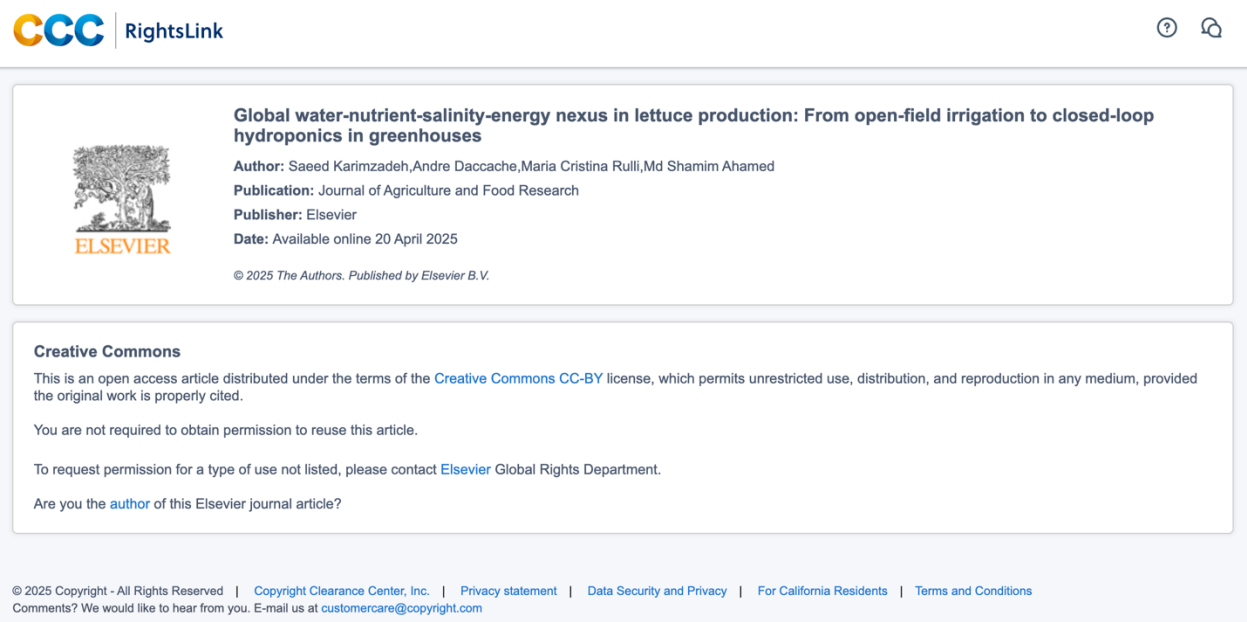
Table 21. Schedule of water and nutrient solution replenishment events in the nutrient solution tank during the experimental period.

Entry number	Date	Water added (L)	Nutrient solution added (L)	Notes
1	Feb 29, 2024	10	0	Water only
2	Mar 4, 2024	10	0	Water only
3	Mar 6, 2024	0	10	Standard nutrient solution
4	Mar 8, 2024	10	10	Combined water and nutrient solution
5	Mar 10, 2024	0	10	Standard nutrient solution
6	Mar 12, 2024	0	10	Calcium nitrate only
7	Mar 14, 2024	0	10	Standard nutrient solution
8	Mar 15, 2024	10	0	Water only
9	Mar 16, 2024	10	0	Water only
10	Mar 17, 2024	10	0	Water only
11	Mar 20, 2024	10	10	Combined water and nutrient solution

Copyright permissions

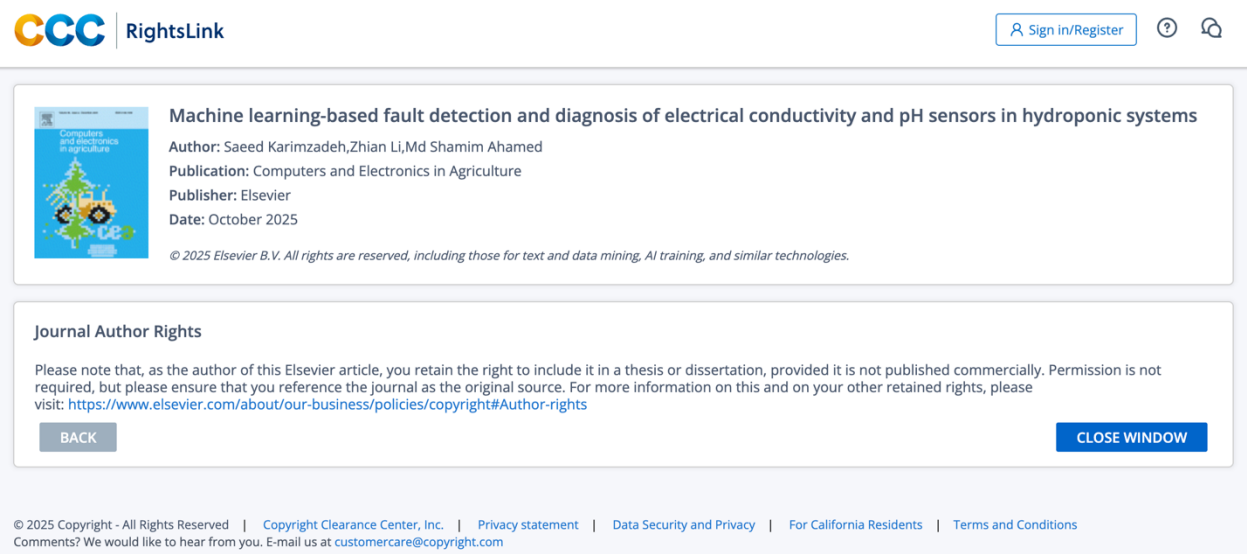
This dissertation includes content from two peer-reviewed journal articles authored by the candidate and published by Elsevier. The copyright and reuse permissions for these articles are as follows:

Chapter 2 is based on the article titled "*Global water-nutrient-salinity-energy nexus in lettuce production: From open-field irrigation to closed-loop hydroponics in greenhouses*" published in the *Journal of Agriculture and Food Research* (Elsevier). This article is published under a Creative Commons CC-BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Therefore, inclusion of this article in the dissertation complies with the publisher's open access policy.



Chapter 3 includes content from the article "*Machine learning-based fault detection and diagnosis of electrical conductivity and pH sensors in hydroponic systems*" accepted for publication in *Computers and Electronics in Agriculture* (Elsevier). According to Elsevier's journal author rights,

authors retain the right to include their articles in a thesis or dissertation, provided it is not published commercially and the original source is properly cited. No further permission is required for such inclusion.



All reused content has been properly cited, and the inclusion of these articles adheres to the relevant copyright and reuse policies as outlined by Elsevier.