



Development of indoor precision vegetable production
system and its mathematical models

PhD Dissertation

by

Nezha Kharraz

Gödöllő
2026

Doctoral school and program denomination:

Doctoral School of Engineering Sciences

Doctoral Program of Mechanical Engineering

Science: Engineering

Head of School: Prof. Dr. László Bozó, MHAS
Institute of Environmental Sciences
Hungarian University of Agriculture and Life Sciences, Budapest,
Hungary

Program Leader: Prof. Dr. Gábor Kalácska, DSc
Institute of Technology
Hungarian University of Agriculture and Life Sciences, Gödöllő,
Hungary

Supervisor: Dr. Szabó István, PhD
Institute of Technology
Hungarian University of Agriculture and Life Sciences, Gödöllő,
Hungary

.....
Affirmation of supervisor

.....
Affirmation of head of doctoral program

Table of Contents

NOMENCLATURE AND ABBREVIATIONS	7
1. INTRODUCTION, OBJECTIVES	12
1.1. Introduction.....	12
1.2. Objectives.....	13
2. LITERATURE REVIEW	14
2.1. Smart Indoor Plant Cultivation History : From Monitoring Technology to Automated Agriculture	14
2.2. IoT Architectures in Precision Agriculture.....	17
2.2.1. The Growing System and Sensor Network	17
2.2.2. Sensor Types and Monitoring Parameters	18
2.2.3. LoRa and LoRaWAN Architecture and Deployment	19
2.2.4. Agricultural WSN Implementations	20
2.2.5. Sensor Network Topologies.....	21
2.2.6. Artificial intelligence in growing systems	22
2.3. Edge Computing Paradigms.....	22
2.4. Cloud Computing Applications	24
2.5. Plant Growth Modeling Approaches	26
2.5.1. Classical Growth Models (Logistic, Richards, and Gompertz equations)	26
2.5.2. Stochastic Model Equation	27
2.5.3. Empirical Modeling for Plant Growth Dynamics.....	28
2.5.4. Malthusian Exponential Growth Equation.....	28
2.5.5. Logit Function for Regression Modeling	29
2.5.6. Limitations of Single-Dimensional and Disconnected Modeling Approaches	29
2.6. Interdependence Modeling in Plant Growth	30
2.7. Statistical and Machine Learning Methods for Predictive Analysis	30
2.7.1. Time Series Analysis.....	30
2.7.2. Multivariate Linear Regression for Resource Contributions.....	31
2.7.3. Random Forest Model	32
2.7.4. Support Vector Machines (SVM)	32
2.7.5. Deep Neural Networks (DNN).....	33
2.7.6. Gradient Boosting.....	34
2.8. Evolution of Plant Growth Models in Agricultural Applications	35
2.9. Performance Metrics and Indicators	36
2.9.1. Photosynthetic Efficiency	36
2.9.2. Root Mean Square Error.....	36
2.9.3. Coefficient of Determination.....	36
2.9.4. Biomass Conversion Metrics.....	37
2.9.5. Performance Metrics	37
2.10. Growth Metrics.....	38
2.10.1. Plant Height	38
2.10.2. Biomass Yield	39
2.10.3. Leaf Area.....	39
2.11. Data Processing Methodologies	40
2.11.1. Data Transmission Protocols	40
2.11.2. Data Acquisition Working Protocol.....	41
2.11.3. Data training	42
2.12. Smart Farming Initiatives: EU and Hungary	43

2.12.1.	EU smart farming initiatives (European Commission, 2023).....	43
2.12.2.	Hungary smart farming initiatives.....	44
2.13.	Critical research gaps.....	45
3.	MATERIALS AND METHODS.....	47
3.1.	Experimental Setup for Plant Growth Evaluation.....	47
3.1.1.	Environmental Control and Growth Conditions.....	47
3.1.2.	IoT-Based Data Collection and Storage of IOT Devices.....	51
3.1.3.	Experimental Conditions.....	51
3.2.	IoT System Architecture and Hardware.....	55
3.2.1.	WiFi-Based MQTT Communication and Wi-Fi Device Communication with Ethernet-Backhauled MQTT.....	55
3.2.2.	Time Responsiveness Analys.....	55
3.3.	Mathematical Modeling of Plant Growth.....	58
3.3.1.	Multidimensional Logistic Growth state vector $X(t)$	58
3.3.2.	Proposed Interdependence Modeling in the 4D Model.....	59
3.4.	Machine Learning and Data Processing Analytics.....	60
3.4.1.	Hybrid Model: Plant Growth and Resource Efficiency.....	60
3.4.2.	Dataset Structure and Model Inputs.....	61
3.4.3.	Cloud-Based Data Processing.....	61
3.4.4.	Data Collection.....	62
3.5.	Novel Growth Performance Metrics.....	64
3.5.1.	Integrated Agricultural Efficiency Metric (IAEM).....	64
3.5.2.	The Growth Efficiency ratio GER.....	66
3.5.3.	Plant Growth Index (PGI).....	67
3.5.4.	Relative Growth Rate (RGR) Metric.....	68
3.5.5.	Light Use Efficiency (LUE) Metrics.....	70
3.5.6.	System-Level Energy Efficiency.....	71
3.5.7.	Ressource Use Efficiency.....	71
3.6.	Image Analysis.....	72
3.7.	Feedback Control and Simulation.....	74
4.	RESULTS AND DISCUSSION.....	75
4.1.	Growth Efficiency Ratio GER.....	75
4.1.1.	Combined Resource Use vs GER.....	75
4.1.2.	GER Heatmap.....	76
4.2.	Plant Growth Traits Characteristics.....	78
4.2.1.	Biomass yield.....	78
4.2.2.	Plant height.....	80
4.2.3.	Leaf Area.....	81
4.3.	Machine Learning for Plant Growth Index (PGI).....	83
4.4.	IoT and Real-Time Monitoring.....	85
4.4.1.	Multivariate Regression Analysis of Growth Drivers.....	85
4.4.2.	Predictive Models Performance.....	85
4.4.3.	Resource Optimization with IoT-Based Monitoring.....	86
4.5.	Integrated Agricultural Efficiency Metric (IAEM).....	88
4.5.1.	Data Latency Metric (DL).....	88
4.5.2.	Resource Efficiency (RE).....	88
4.5.3.	Real-Time Alert Accuracy (RA).....	88
4.5.4.	Cloud Data Availability (CDA).....	89
4.5.5.	Pairwise Comparison algorithm using AHP for weights.....	89

4.6.	IoT Framework Performance.....	90
4.6.1.	Analysis of Light intensity	90
4.6.2.	Development of the Growth Response Model	91
4.6.3.	RGR Model Validation.....	92
4.7.	Optimal Configuration	93
4.7.1.	Light Use Efficiency (LUE) Metrics	93
4.7.2.	Energy Efficiency Analysis.....	94
4.8.	Comparative Growth Performance of Treatment Groups	95
4.8.1.	Chlorophyll Evolution	95
4.8.2.	Biomass Accumulation.....	96
4.8.3.	Height Progression.....	96
4.8.4.	Leaf Area Development	97
4.8.5.	Allometric Relationship Between Height (H) and Biomass (B)	97
4.8.6.	Sensitivity of Growth Rate to Nutrient Concentration	99
4.8.7.	Sensitivity of Growth Rate to Light	100
4.8.8.	Sensitivity of Growth Rate to CO ₂	102
4.9.	Advanced Growth Four-Dimensional Model Validation.....	103
5.	NEW SCIENTIFIC RESULTS.....	109
5.1.	Multi-Dimensional Logistic Growth Model	109
5.2.	Allometric Relationship Between Height (H) and Biomass (B).....	110
5.3.	Relative Growth Rate Under Controlled Environmental Conditions.....	110
5.4.	Quantitative Models for Light Use Efficiency vs. Leaf Area Index (LAI).....	111
5.4.1.	Light Use Efficiency at Canopy Level (LUPE)	111
5.4.2.	Light Use Efficiency at Lamp Level (LULE).....	111
5.5.	Plant Growth Index (PGI) with Machine-Learned Weights.....	111
5.6.	Mathematical Model and Control Logic for Nutrient Optimization	112
6.	CONCLUSION AND SUGGESTIONS	113
7.	SUMMARY	114
8.	ÖSSZEFOGLALÁS (SUMMARY IN HUNGARIAN)	117
9.	APPENDICES.....	119
A1:	Bibliography.....	119
A2:	Publications related to the dissertation.....	128
A3:	Comparative Analysis of IoT Systems in Smart Agriculture	130
A4:	UML Model of the system	133
A5:	IoT System Architecture for Lighting Control	134
A6:	Time Responsiveness Analysis of MQTT Communication Protocols	134
A7:	Weight Determination.	136
A8:	Python Script For Machine Learning Model (Linear Regression):	136
A9:	Calculation of Light Use Efficiency Metrics.....	139
A10:	Image Analysis Methodology for Plant Phenotyping.....	140
A11:	Image Processing for Plant Detection and Lighting Simulation.....	141
A12:	Image Processing for Plant Detection and Lighting Simulation.....	143
A13:	Real-Time Plant Growth Feedback Control System	143

A14: Data Used To Compute GER Over 30 Days.	144
A15: Biomass Yield And GER Across Light Conditions.....	144
A16: Biomass Yield And GER Across Water Conditions	144
A17: Biomass Yield And GER Across Nutrient Conditions	144
A18: The Effect Of Fertilizer Input And Watering Levels On Plant Growth.....	145
A19: The Effect Of Light Consumption And Watering Levels On Plant Growth.	145
A20: Data for Leaf Area Growth Vs Water Levels.....	146
A21: Data for Leaf Area Vs Nutrient	146
A22 : Data for Leaf Area Vs Light	147
A23: User Input Data Acquisition	148
A24: Weight Calculation for PGI Parameters	148
A25: Raw Data For Computing The Plant Growth Index (PGI)	149
A26: Computation of the Plant Growth Index (PGI)	150
A27: Influence Of Environmental Factors On Crop Growth Using Multivariate Regression.....	150
A28:Experimental Environmental Conditions and Corresponding Plant Growth Measurements...	151
A29:Comparison between experimented crop yield and predicted values obtained using SVM, Gradient Boosting, Deep Neural Network, and Time Series models for individual plant samples..	151
A30:Measured Growth Before and After IoT Implementation	151
A31: System Optimization Metrics Before And After IoT Implementation	152
A32: Real-Time Alert Accuracy Improvements.....	152
A33: Experimental vs. simulated RGR values (Treatment Group A).....	152
A34: Experimental vs. simulated RGR values (Treatment Group B).....	152
A35: Experimental vs. simulated RGR values (Treatment Group C)	153
A36: MATLAB Code for Stepwise Regression Analysis of RGR (Treatment Groups A, B, C)	153
A37: Summary Metrics Across Configurations	154
A38: Treatment Groups.....	155
A39: Monitoring API Alerts	158
10. ACKNOWLEDGEMENT	161

NOMENCLATURE AND ABBREVIATIONS

Symbol	Description	Unit
$A, A(t)$	Leaf area	cm^2
a, b, c	Polynomial coefficients (GER model)	Unitless
b_{model}	Bias term (SVM, DNN)	Unitless
B	Effective bandwidth	$\text{bits} \cdot \text{s}^{-1}$ (bps)
B_{trees}	Number of trees (Random Forest)	Unitless
$B, B(t)$	Biomass	g
$C, C(t)$	Chlorophyll content	SPAD or $\mu\text{g} \cdot \text{cm}^{-2}$
$C_{\text{footprint}}$	CO ₂ reduction	%
$\text{CO}_2, \text{CO}_2(t)$	Carbon dioxide concentration	ppm
$\frac{dX(t)}{dt}$	Growth rate of trait X	unit of $X \cdot \text{time}^{-1}$
e	Base of natural logarithm	Unitless
E	Error / loss function	Unitless
E_{savings}	Energy savings	%
$f(X)$	Classifier function (SVM)	Unitless
$f_t(x)$	Ensemble model at iteration t(GB)	Unitless
$f_b(x')$	Prediction of b-th tree	Unitless
GR	Growth rate	$\text{cm} \cdot \text{day}^{-1}$ or $\text{cm} \cdot \text{week}^{-1}$
$h_t(x)$	Weak learner at iteration t	Unitless
$H, H(t)$	Plant height	cm
K	Carrying capacity	same as trait (cm, g, etc.)
K_{nutrient}	Potassium concentration	$\text{mg} \cdot \text{kg}^{-1}$
K_L	Light saturation constant	$\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
K_N, K_J	Half-saturation constants (nutrient/light)	$\text{mg} \cdot \text{kg}^{-1} / \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
L, I, J	Light intensity	$\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
ℓ	Loss function (GB)	Unitless
n	Number of samples	Unitless
N	Nutrient input (e.g., nitrogen)	$\text{mg} \cdot \text{kg}^{-1}$ or $\text{g} \cdot \text{day}^{-1}$
$N_{\text{available}}$	Available nutrients	g
N_{used}	Absorbed nutrients	g
P	Phosphorus concentration	$\text{mg} \cdot \text{kg}^{-1}$
P_{max}	Maximum photosynthetic rate	$\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
P_{success}	Transmission success probability	Unitless
r	Intrinsic growth rate	day^{-1}
R	Respiration rate	$\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
S	Sustainability factor	Unitless
S_{msg}	Message size	bits (b)
SR	Success rate	%
t	Time	s, h, or d
T	Temperature	$^{\circ}\text{C}$
T_p	Propagation/processing delay	ms

Δt	Time step	d
w	Weight vector (SVM)	Unitless
w_1, w_2, w_3	PGI weights	Unitless
W	Water input	$L \cdot day^{-1}$
x	Input feature vector	context-dependent
X(t)	State variable	context-dependent
X_0, P_0	Initial value	same as variable
Y	Output / label	context-dependent

Greek symbols

Symbol	Usage (Context)	Unit
α	1. Weighting coefficient in sustainability factor S	Unitless
	2. Scaling constant in biomass model $B = \alpha L^\beta C^\gamma$	Unitless
	3. Light efficiency coefficient in growth rate (GR)	$\text{cm} \cdot \text{m}^2 \cdot \text{s} \cdot \mu\text{mol}^{-1}$
	4. Learning rate in Gradient Boosting	Unitless
	5. Asymptotic maximum volume in Gompertz model	cm^3
	6. Contrast parameter in image processing	Unitless
β	1. Weighting coefficient in dynamic carrying capacity	Unitless
	2. Nutrient efficiency coefficient in growth rate (GR)	$\text{cm} \cdot \text{ppm}^{-1}$
	3. Displacement parameter in Gompertz model	Unitless
	4. Exponent in biomass model $B = \alpha L^\beta C^\gamma$	Unitless
	5. Regression coefficient in multivariate models	Unitless
	6. Time delay components (β_1, β_2) in IoT latency model Ψ	ms
	7. Brightness parameter in image processing	Unitless
γ	1. Growth rate coefficient in Gompertz model	day^{-1}
	2. Exponent for chlorophyll influence in biomass model	Unitless
Δ	Change or increment $(\Delta I, \Delta t)$	Same as variable
ε	1. Error term in regression models	Context-dependent
	2. Assimilation efficiency in growth rate	Unitless
	3. Adjustment / noise factor	Unitless
η	Biological noise term in growth rate calculations	Unitless
κ	Conversion factor (energy efficiency, $\mu\text{mol} \cdot \text{J}^{-1}$)	$\mu\text{mol} \cdot \text{J}^{-1}$
ℓ	Loss function (Gradient Boosting)	Unitless
μ	Drift term in stochastic differential equations (SDEs)	Context-dependent
σ	1. Diffusion term in SDEs	Context-dependent
	2. Standard deviation	Context-dependent
σ_{pooled}	Pooled standard deviation	Context-dependent
ϕ	Photosynthetic rate (function of light, temperature, chlorophyll)	$\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$
Ψ	Total end-to-end system delay	ms
Ψ_{eff}	Effective delay (with retransmissions)	ms
ω	Weights in IAEM metric	Unitless
Σ	Summation operator	N/A
$\odot$	Element-wise multiplication	N/A

Abbreviations

AHP: Analytic Hierarchy Process
AI: Artificial Intelligence
API: Application Programming Interface
ARM: Advanced RISC Machine
CAP: Common Agricultural Policy
CBOR: Concise Binary Object Representation
CEAC: Controlled Environment Agriculture Center
CI: Confidence Interval
CO₂: Carbon Dioxide
CoAP: Constrained Application Protocol
CRU: Combined Ressource Use
CYE: Crop Yield Efficiency
DNN: Deep Neural Networks
DSS: Decision Support System
4D: Four-Dimensional
EC: Electrical Conductivity
ETL: Extract, Transform, Load
FLC: Fuzzy Logic Control
FMIS: Farm Management Information System
FN: False Negatives
FP: False Positives
GER: Growth Efficiency Ratio
GPIO: General Purpose Input/Output
HSV: Hue, Saturation, Value (color space)
HTTP: Hypertext Transfer Protocol
IAEM: Integrated Agricultural Efficiency Metric
IDS: Intrusion Detection System
IoT: Internet of Things
IQR: Interquartile Range
JSON: JavaScript Object Notation
LAB: Lightness, A, B (perceptually uniform color space)
LAI: Leaf Area Index
LED: Light Emitting Diode
LoRaWAN: Long Range Wide Area Network
LoS: Line-of-Sight
LUE: Light Use Efficiency
LUEL: Light Use Efficiency at Lamp Level
LUEP: Light Use Efficiency at Plant Canopy Level
MAC: Medium Access Control
MAE: Mean Absolute Error
MQTT: Message Queuing Telemetry Transport
MSE: Mean Squared Error

NLoS: Non-Line-of-Sight
NPP: Net Primary Productivity
NUE: Nutrient-Use Efficiency
PAR: Photosynthetically Active Radiation
PER: Packet Error Rate
PGI: Plant Growth Index
PPFD: Photosynthetic Photon Flux Density
ppm: parts per million
PWM: Pulse Width Modulation
QoS: Quality of Service
 R^2 : R-Squared (Coefficient of Determination)
RAM: Random Access Memory
REST: Representational State Transfer
RH: Relative Humidity
RGR: Relative Growth Rate
RMSE: Root Mean Square Error
RSSI: Received Signal Strength Indicator
RUE: Radiation Use Efficiency
SD: Standard Deviation
SDE: Stochastic Differential Equation
SNR: Signal-to-Noise Ratio
SPAD: Soil Plant Analysis Development (chlorophyll meter)
SVM: Support Vector Machines
TCP/IP: Transmission Control Protocol/Internet Protocol
TLS: Transport Layer Security
TN: True Negatives
TP: True Positives
TWSN: Terrestrial Wireless Sensor Network
UDP: User Datagram Protocol
VPD: Vapor Pressure Deficit
VWC: Volumetric Water Content
WSN: Wireless Sensor Network
WUSN: Wireless Underground Sensor Network
xCEA: Controlled Environment Agriculture

INTRODUCTION, OBJECTIVES

1. INTRODUCTION, OBJECTIVES

1.1. Introduction

Controlled Environment Agriculture (CEA) has become a focal point of agricultural innovation, especially through the use of soilless cultivation systems such as hydroponics and aeroponics. These indoor farming platforms enable precise control over environmental variables including temperature, relative humidity, CO₂ and O₂ concentrations, water pH, electrical conductivity, and nutrient exposure. Such conditions not only determine whether a plant grows but also influence how it grows. A core objective in the field of CEA is to understand the specific needs of different plant species across various growth stages. However, this thesis focuses exclusively on lettuce (*Lactuca sativa* L.) as a model organism to validate the proposed framework. By analyzing phenotypic traits such as color, size, texture, yield, growth rate, flavor, and nutrient density, researchers can prevent unwanted phenotypic expressions in crops exposed to varying environments.

To address the complexity of optimizing growth conditions, logistic growth models have emerged as valuable tools for modeling plant development in CEA systems. These models provide a data-driven framework for simulating and predicting growth trends, enabling more precise management of critical variables such as light, temperature, and nutrient levels. Researchers are increasingly integrating these models with real-time analytics to support sustainable agriculture, improve yield, and enhance resource efficiency. This study focuses on the use of experimental sensor data and empirical modeling techniques to identify interrelationships between environmental and physiological factors affecting plant growth. Through machine learning and predictive modeling approaches, the proposed framework aims to optimize resource allocation and improve cultivation performance in indoor farming systems.

Platforms integrating environmental sensors with cloud services for real-time decision-making have shown promising results in various applications, including an IoT-based hydroponic system for lettuce (*Lactuca sativa* L.) (Li et al., 2022), deep learning for spectral optimization in tomato cultivation (Park et al., 2020), and data fusion frameworks for optimizing irrigation and lighting (Wang et al., 2023). The integration of cloud services offers several advantages, including centralized data storage, remote monitoring, real-time analytics, scalability, and continuous synchronization between sensors, controllers, and predictive models. These capabilities improve system adaptability and enable faster responses to environmental variations in indoor farming environments. However, challenges persist, particularly in dynamic light management, where many systems still rely on fixed photoperiods or static light intensities without real-time adaptability.

Recent advancements in feedback control and machine learning have attempted to improve Light Use Efficiency (LUE), but often in isolation, targeting only a single parameter such as photoperiod or intensity. Furthermore, many machine learning applications in CEA lack physiological interpretability and sufficient integration with stochastic or empirical growth models, which can reduce model robustness and generalizability under varying environmental conditions.

INTRODUCTION, OBJECTIVES

This thesis investigates the integration of logistic growth modeling, machine learning, cloud-based IoT infrastructures, and real-time environmental sensing within a unified smart indoor farming framework. Particular attention is given to the interaction between environmental variables and plant physiological responses in order to support more adaptive and data-driven cultivation strategies for Controlled Environment Agriculture systems.

1.2. Objectives

This study employs controlled environment agriculture (CEA) in indoor systems integrated with IoT sensors and hybrid mathematical models (stochastic, empirical, and optimization-based) to analyze indoor plant growth under varying resource inputs. While lettuce (*Lactuca sativa* L.) is used as the primary model organism for validation of the developed framework and methodologies. The produced datasets and simulations will be used to validate predictive models and optimize resource efficiency.

The main objectives are to investigate the following:

- a) To evaluate the influence of a four-dimensional logistic growth model on the predictive accuracy of plant development under dynamic environmental conditions.
- b) To develop and validate a multivariate regression model describing the relationship between relative growth , light intensity, photoperiod, and plant maturity.
- c) To compare light use efficiency at both the canopy level (LUEP) and lamp level (LUEL) as a function of Leaf Area Index (LAI) in order to identify the optimal canopy density for energy-efficient production.
- d) To determine the optimal machine-learned weighting coefficients for a composite Plant Growth Index (PGI) used to quantify plant health and resource-use efficiency.
- e) To derive a multivariate regression model that identifies the nutrient concentration thresholds beyond which additional inputs provide zero or negative marginal growth returns, enabling precision dosing and waste reduction.
- f) To establish and validate a generalized empirical allometric relationship between plant height and biomass for lettuce (*Lactuca sativa* L.) plants.

2. LITERATURE REVIEW

This chapter reviews the current state of indoor plant growing, with a focus on the integration of advanced technologies such as artificial intelligence (AI) and the Internet of Things (IoT).

It examines the role of mathematical modeling in optimizing growth conditions, resource efficiency, and automation in controlled environments, particularly for small-scale and indoor farming systems. Recent advancements in predictive growth models, spectral light optimization, and adaptive environmental control are discussed, alongside their potential to enhance yield and reduce energy consumption. However, existing studies often lack unified, scalable frameworks that combine real-time monitoring with dynamic growth optimization; a gap this research directly addresses. This chapter also identifies key challenges in the literature, setting the foundation for this research's innovative computational and IoT-based approaches to advance precision indoor agriculture.

2.1. Smart Indoor Plant Cultivation History : From Monitoring Technology to Automated Agriculture

The history of smart technology in controlled environment agriculture began with foundational advances in monitoring and control equipment. The oldest documented example dates back to 1986, when researchers at the University of Arizona's CEAC developed one of the first computerized hydroponic monitoring systems, utilizing early sensors to track pH and electrical conductivity (EC) in nutrient solutions (Sipahioglu & Barringer, 2003). This was followed by the introduction of the first commercial automation system in 1995, when the Dutch company Priva launched its "NutriFlex" climate computer, enabling automated regulation of temperature, humidity, and irrigation in glasshouses (Marković et al., 2012).

The technological landscape was further transformed by the advent of IoT-enabled sensor networks and advanced actuation systems. Philips GreenPower's adaptive LED systems, for example, demonstrated dynamic spectral tuning for growth-stage optimization (van Iersel & Gianino, 2017). Climate automation also advanced significantly, with Priva's 2019 IntelliFlex system achieving precise ± 0.1 kPa control of vapor pressure deficit (VPD) (Noh & Lee, 2022). More recently, the integration of artificial intelligence has enhanced diagnostic capabilities, as seen in Bowery Farming's 2021 computer vision systems, which can predict nutrient deficiencies up to 14 days before visual symptoms emerge (Uzhinskiy, 2023).

The application of these evolving technologies has enabled a new generation of highly efficient commercial farms. AeroFarms' 2015 vertical farm, for instance, implemented real-time root-zone monitoring with ± 0.2 pH precision, demonstrating the practical viability of intensive data-driven cultivation (Kozai et al., 2015). Compared to traditional greenhouses, modern smart farms achieve 40–60% greater resource efficiency (Benke & Tomkins, 2017). Recent large-scale deployments, such as Infarm's 2022 modular farms and Crop One's 2023 fully autonomous facility, have reported dramatic yield increases of 200–300% through the application of comprehensive data-driven cultivation strategies (Butturini & Marcelis, 2019).

LITERATURE REVIEW

The following section outlines the chronological progression of initial advancements in Smart Farming Equipment :

1. 1986 – University of Arizona’s computer-controlled hydroponics (pH/EC monitoring) shown in Fig. 2.1 :

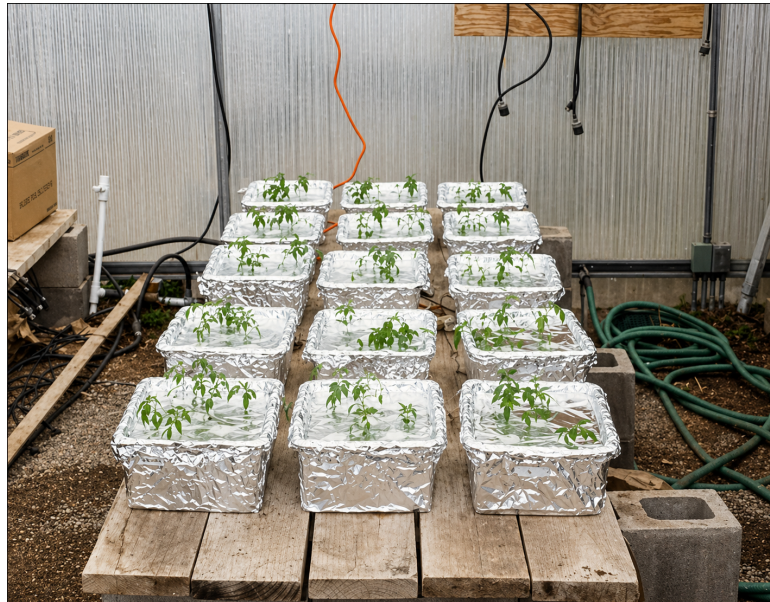


Fig. 2.1. Computer-controlled hydroponics 1986 (Al-Saif, 1986).

2. 1995 – Priva’s NutriFlex (first commercial climate computer for greenhouses) shown in Fig. 2.2 :

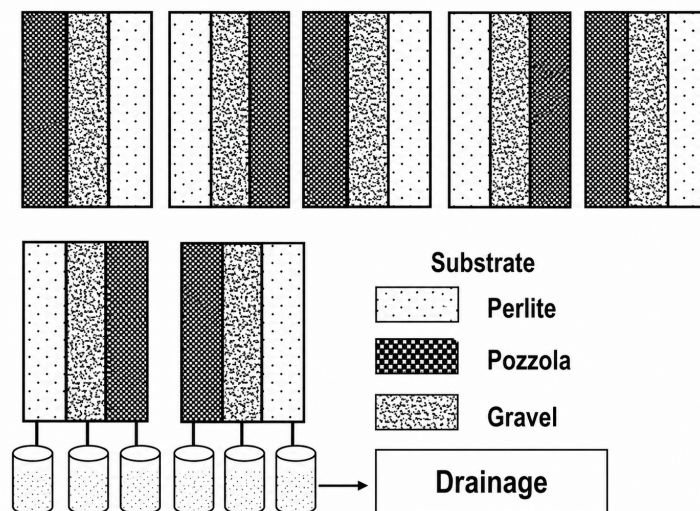


Fig. 2.2. computerised Priva Nutriflex result 1995 (Popovski & Popovska, 2012).

LITERATURE REVIEW

3. 2004 – First IoT-enabled greenhouse trials (Japan's "Plant Factory" project) shown in Fig. 2.3 :



Fig. 2.3. basic unit of CPPS at Chiba University in 2004 and commercialized in 2005 (Kozai, 2013).

4. 2014 – Introduction of spectral-tunable LEDs (NASA-funded research for space farming) shown in Fig. 2.4 :



Fig. 2.4. Growing Food in Space Using LED Grow Lights (Wagner, 2021).

The following section outlines the chronological progression of initial advancements in smart indoor farming:

5. 2015 – AeroFarms deploys IoT automation in vertical farming in Fig. 2.5 :



Fig. 2.5. AeroFarms' Newark-based vertical farm (Bloom, 2021).

6. Vertical farm 2023 shown in Fig. 2.6 :



Fig. 2.6. An automated vertical farm (IoT Worlds, 2022)

2.2. IoT Architectures in Precision Agriculture

2.2.1. The Growing System and Sensor Network

The implementation of a smart agriculture system involves the integration of multiple sensors and devices to monitor and control the crop environment in real time. As illustrated in Fig. 2.7, the system includes a network of soil moisture sensors, temperature and humidity sensors, and light sensors, all of which feed data to a centralized microcontroller. This data is then transmitted to a cloud-based platform via Wi-Fi or GSM modules, enabling remote monitoring and decision-making. The system architecture supports real-time automation, ensuring that critical parameters like soil moisture, air humidity, and light intensity are consistently regulated. This closed-loop setup not only improves efficiency and reduces manual labor but also enhances plant health and yield predictability through precision agriculture techniques.



Fig. 2.7. Soil moisture levels measured by a capacitive soil moisture sensor (Harry-Lees, 2020).

Wireless sensor network-based systems represent a promising approach to fully automate agriculture, as sensors provide real-time, critical data and help eliminate the high cost of unnecessary wiring. In agriculture, this technology improves greenhouse systems by efficiently transmitting real-time data via radio signals to operators. Building on this concept, a model titled "Plant Monitoring and Leaf Disease Detection with Classification using Machine Learning in MATLAB" was proposed for implementing an intelligent system for both monitoring and disease detection (Ramya et al., 2020). The system employs a combination of environmental sensors (temperature, humidity, and NPK) and a camera to collect data on plant health and environmental conditions as shown in Fig. 2.8. These data are transmitted through a NodeMCU and GSM module to a mobile device using the Internet. Leaf disease classification is handled using machine learning algorithms developed in MATLAB, providing timely and accurate diagnostics.

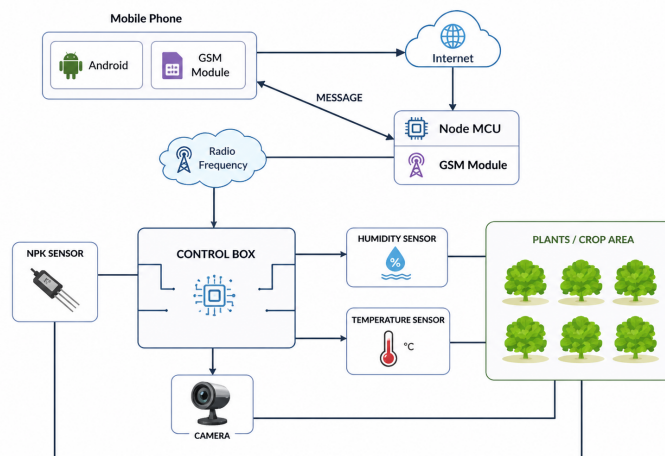


Fig. 2.8. Functional Diagram of the Plant Monitoring and Leaf Disease Detection System (Ramya et al., 2020)

2.2.2. Sensor Types and Monitoring Parameters

Fig. 2.9 presents the main sensors and control components used in an aeroponic monitoring system. These devices are responsible for measuring and regulating the environmental and nutrient

LITERATURE REVIEW

conditions required for optimal plant growth. The system integrates several sensing units, including a pH sensor, electrical conductivity (EC) sensor, light intensity sensor, humidity sensor, carbon dioxide (CO₂) sensor, and water level sensor. These sensors continuously monitor critical environmental variables such as nutrient solution quality, light exposure, humidity, and CO₂ concentration.

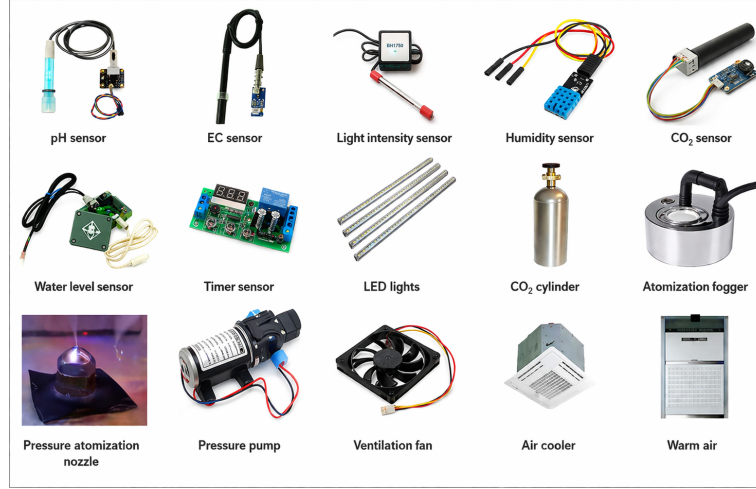


Fig. 2.9. Sensors and actuators used in aeroponic system (Sneha et al., 2023)

Together, these sensors and actuators form an integrated monitoring and control system that enables precise regulation of the aeroponic growing environment. By continuously measuring environmental parameters and adjusting system components accordingly, the system helps maintain optimal conditions for plant development and improves overall cultivation efficiency (Sneha et al., 2023).

2.2.3. LoRa and LoRaWAN Architecture and Deployment

LoRa, patented by Semtech Corporation and commercially introduced in 2015, defines the physical layer of communication using Chirp Spread Spectrum (CSS) modulation. Originally developed for military communications, CSS enables long-range, low-power transmission with strong resistance to noise and interference. Operating in unlicensed ISM sub-GHz bands, LoRa supports communication over several kilometers while allowing adjustable data rates and robustness through configurable spreading factors (SF7–SF12).

Built on top of LoRa, LoRaWAN defines the medium access control (MAC) layer and network architecture for IoT connectivity, as illustrated in Fig. 2.10. The architecture follows a star-of-stars topology composed of end devices communicating with gateways, which relay data to a network server and subsequently to an application server for cloud storage and processing. LoRaWAN uses an ALOHA-based MAC protocol to coordinate communications among multiple end nodes.

LoRaWAN supports three device classes: Class A (energy-efficient bidirectional communication), Class B (scheduled receive windows), and Class C (continuous listening with higher power consumption).

As shown in Fig. 2.10, LoRaWAN systems can be evaluated using both indoor and outdoor experimental environments through real-world testbeds and simulation platforms.

LITERATURE REVIEW

Implementations may involve development modules (e.g., TTGO, Heltec, MCCI) as well as commercial LoRa-enabled devices and sensors. Performance evaluations generally focus on key indicators including coverage, Time on Air (TOA), Packet Error Rate (PER), Received Signal Strength Indicator (RSSI), and Signal-to-Noise Ratio (SNR) (Bin et al., 2024; Lavanya et al., 2020).

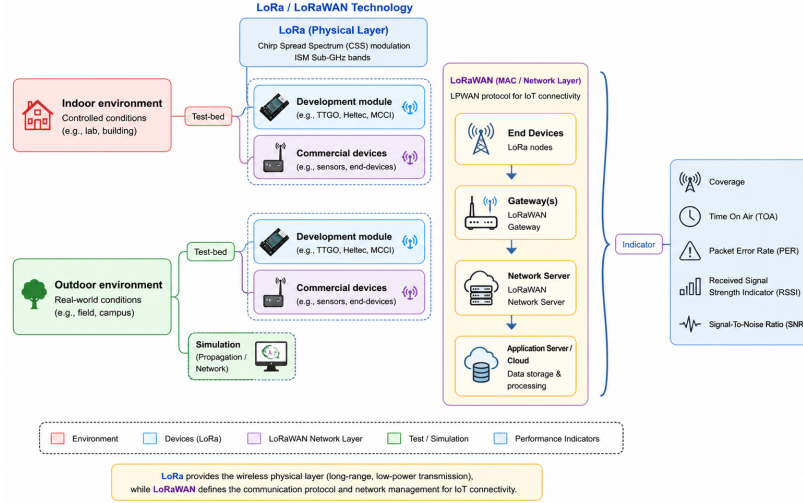


Fig. 2.10. Framework for performance analysis of LoRa and LoRaWAN (Bin et al., 2024)

2.2.4. Agricultural WSN Implementations

Wireless Sensor Networks (WSNs) have emerged as a key technology in precision agriculture due to their ability to monitor environmental conditions in real time and support automated decision-making. As illustrated in Fig. 2.11, agricultural WSNs typically consist of distributed sensor nodes, gateway nodes, and surface sinks connected to remote servers through the Internet for data collection, processing, and monitoring.

Two main categories of agricultural WSNs are commonly used: Terrestrial Wireless Sensor Networks (TWSNs), where nodes are deployed above ground, and Wireless Underground Sensor Networks (WUSNs), where sensor nodes are embedded in the soil to monitor parameters such as soil moisture and temperature. WUSNs support both underground-to-underground and underground-to-surface communication links through gateway nodes and surface sinks, enabling reliable data transmission for intelligent agricultural management. These systems also provide important features such as energy-efficient operation, fault tolerance, dynamic topology configuration, and remote monitoring via GSM/GPRS and Internet-based services (Ojha et al., 2015).

LITERATURE REVIEW

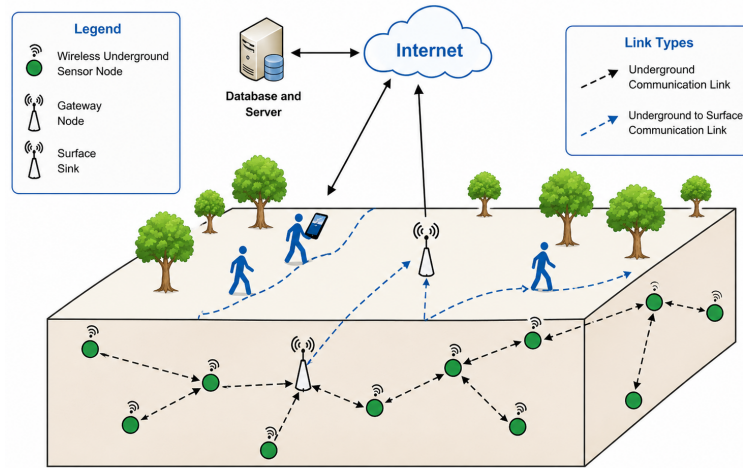


Fig. 2.11. A typical wireless sensor network deployed for agricultural applications (Ojha et al., 2015).

2.2.5. Sensor Network Topologies

The topology and communication protocols of sensor networks are critical to the performance of smart agriculture systems, particularly in terms of latency, reliability, and robustness under varying environmental conditions. Adaptive mechanisms that dynamically switch between LoRaWAN and IEEE 802.11ac protocols using IEEE 802.11ac for high-throughput multimedia transmission and LoRaWAN for low-data-rate sensor communication can significantly improve system performance (Goap et al., 2018). Experimental evaluations of packet error rate (PER) under line-of-sight (LoS) and non-line-of-sight (NLoS) conditions further demonstrate the impact of topology on protocol performance. In LoS environments over distances up to 1500 meters, IEEE 802.11ac achieves a lower PER than LoRaWAN, underscoring its suitability for high-quality transmission in low-interference scenarios. Conversely, LoRaWAN outperforms IEEE 802.11ac in NLoS conditions, where obstructions reduce the effective range of IEEE 802.11ac to approximately 95 meters. This robustness in complex propagation environments highlights LoRaWAN's adaptability for agricultural settings (Ramli et al., 2020).

Fig. 2.12 illustrates a hybrid IoT communication network integrating Wi-Fi and LoRa technologies for smart agriculture systems. This topology enables efficient data routing by assigning high-bandwidth transmissions to Wi-Fi and long-range, low-power communication to LoRa. Such a hybrid structure reduces latency for critical data, improves reliability under variable conditions, and enhances scalability in distributed agricultural environments (Lakhia et al., 2018).

LITERATURE REVIEW

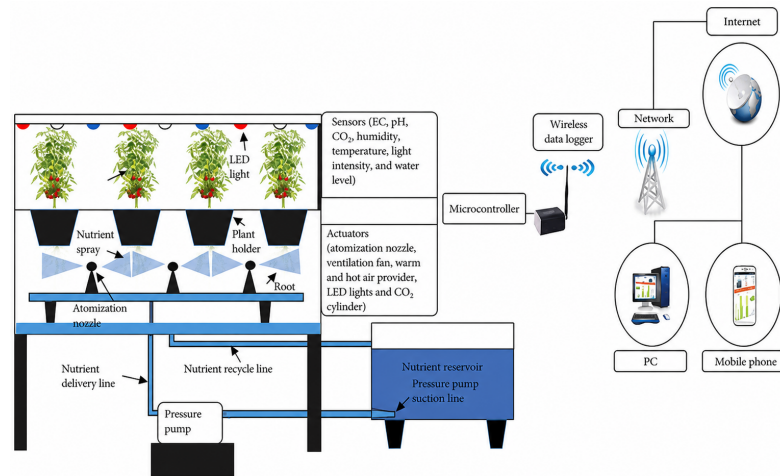


Fig.2.12.Integrated Network for Smart Hydroponic System Monitoring (Lakhiar et al., 2018).

2.2.6. Artificial intelligence in growing systems

Maintaining optimal nutrient solution parameters including temperature, pH, and electrical conductivity (EC) is critical for plant growth in growing systems. Additional variables such as atomization time/intervals, air temperature, humidity, light intensity, and CO₂ concentration further complicate system management, demanding significant human oversight, expertise, and energy. These parameters should be always monitored within precise ranges to ensure ideal growth conditions. Failure to do so can lead to plant stress, yield loss, or crop failure.

This gap makes growing systems prone to delayed responses, increasing risks of irreversible damage. Such complexities, coupled with the need for high-precision control, have limited widespread adoption not due to cost, but operational demands. To address these challenges, modern systems integrate IoT sensors and AI-driven automation, as illustrated in Fig. 2.13, where data flows from IoT devices through a gateway to a centralized platform, enabling real-time monitoring, fault detection, and automated adjustments to maintain ideal growing conditions while reducing reliance on manual intervention (Lakhiar et al., 2018).

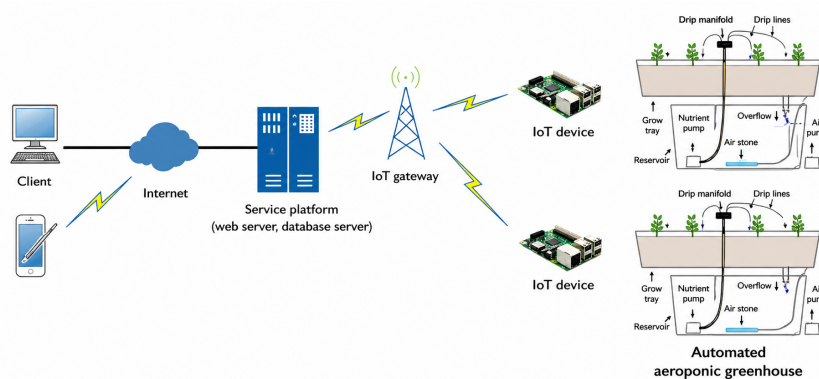


Fig. 2.13. IoT in Aeroponic systems (Lakhiar et al., 2018)

2.3. Edge Computing Paradigms

Recent advances in distributed computing have demonstrated the efficacy of fog-based systems for agricultural monitoring. As illustrated in Fig. 2.14, a hierarchical processing model (Guardo et

LITERATURE REVIEW

al., 2018) effectively partitions computational tasks between edge devices (Fog Aggregator Nodes) and gateways (Fog Collector Nodes), leveraging localized data filtering and hourly aggregation to reduce cloud transmissions by 90%. This architecture employs LoRa-based star topologies to balance long-range connectivity (2-15km) with low power consumption (<50kbps), while maintaining $O(N)$ time complexity for edge-side algorithms. The two-tier fog implementation proves particularly valuable for latency-tolerant applications like soil monitoring, where cluster-based analysis of field variability enables adaptive irrigation control without cloud dependency. Such frameworks provide empirical validation for edge computing's role in precision agriculture, particularly in resource-constrained environments where distributed intelligence enhances scalability and reduces actuation delays.

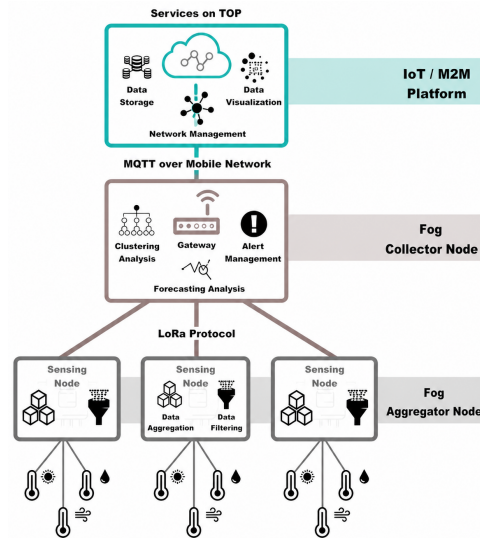


Fig.2.14. Fog computing for real-time actuation (Guardo et al., 2018)

Edge computing addresses critical latency and bandwidth challenges in agricultural IoT by processing sensor data closer to the source, reducing cloud server load while enabling real-time responses for irrigation and climate control (Zhang et al., 2020). The three-layer edge architecture (device-edge-cloud) optimizes resource allocation, with edge nodes handling time-sensitive tasks like nutrient dosing (<500ms) and cloud servers managing long-term growth analytics.

Fig. 2.15 below illustrates four edge computing paradigms for deep neural network (DNN) deployment in agricultural IoT: (a) Edge-based inference where raw sensor data is processed entirely on edge servers; (b) Device-based execution with pre-deployed models; (c) Collaborative edge-device partitioning for latency-sensitive tasks; and (d) Cloud-assisted processing for complex analytics. For hydroponic control systems, mode (c) optimally balances Arduino's limited compute resources with edge servers' capability to handle partial DNN layers, reducing latency by 30-50% compared to pure cloud approaches (Zhang et al., 2020).

LITERATURE REVIEW

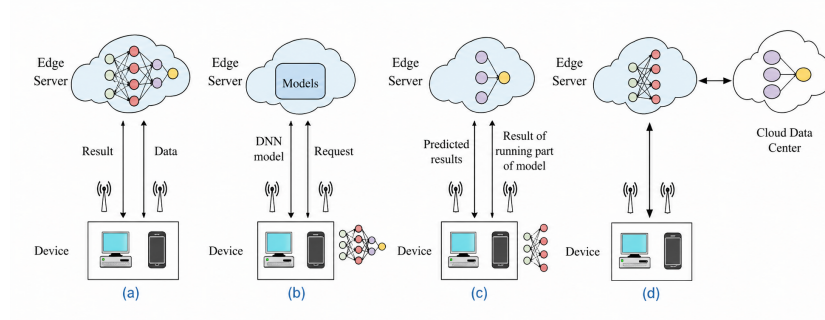


Fig.2.15. Edge-Cloud Computing Architectures for Agricultural IoT (Zhang et al., 2020)

2.4. Cloud Computing Applications

Cloud computing plays a pivotal role in modern smart agriculture by enabling scalable data storage, real-time analytics, and seamless integration with IoT systems. Platforms like AWS (Amazon Web Services) and Azure (Microsoft, 2023) support agricultural IoT applications by processing vast datasets from sensors, drones, and farm machinery, facilitating precision farming and resource optimization (Choudhary et al., 2025). For instance, cloud infrastructures synchronize field data such as soil moisture and weather patterns with AI-driven models to optimize irrigation, pest control, and yield prediction. Furthermore, cloud services enhance supply chain transparency and farm management systems (FMIS) by aggregating data from satellites, on-ground sensors, and market trends.

Fig. 2.16 illustrates the synergy between IoT devices (sensors, drones, farm machinery) and web services (cloud platforms, APIs, data analytics tools) in smart agriculture. IoT devices collect real-time field data (soil moisture, temperature, crop health), which is transmitted to web services for processing, storage, and actionable insights. Cloud-based web services enable scalable data management, remote monitoring, and integration with AI models, facilitating precision farming and decision-making.

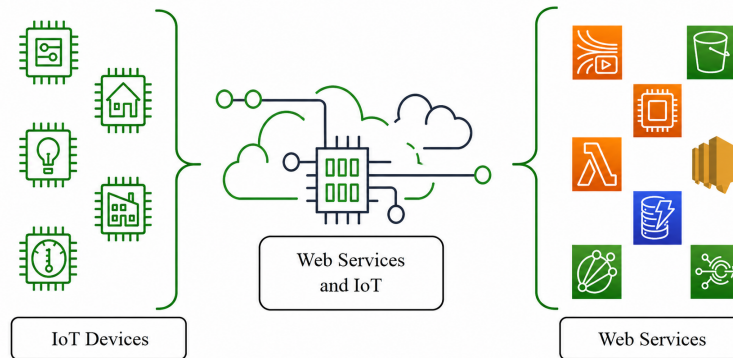


Fig. 2.16. Integration of IoT Devices with Web Services (Choudhary et al., 2025)

Table 2.1 highlights the distinct capabilities of four smart farming systems. The proposed Azure-based pipeline (2024) demonstrates advanced cloud integration using Azure IoT Hub and Blob Storage, combined with dynamic feedback control through MATLAB Simulink and predictive analytics (Gradient Boosting, SVM, and DNN). This architecture enables real-time monitoring, scalability, and secure data transmission suitable for large-scale agricultural deployments. In contrast, the ThingSpeak-based system (Muhammad et al., 2023) targets low-cost, small-scale

LITERATURE REVIEW

applications with basic data aggregation and no predictive analytics. The IoT smart farming system proposed by Ramakrishna et al. (2019) focuses primarily on automated irrigation with moderate scalability but lacks comprehensive sensing and advanced decision-making models. The cloud-based smart farming system by Bilal et al. (2023) integrates IoT and cloud services for small to medium farms but offers limited analytics and no real-time adaptive feedback. Overall, the proposed Azure-based pipeline provides a more comprehensive, intelligent, and scalable solution for precision agriculture.

Table 2.1 Comparison of smart farming approaches.

System	Our Azure-Based Pipeline (2024)	ThingSpeak-Based System (2023)	IoT Smart Farming (2019)	Cloud-Based Smart Farming (2023)
Cloud Platform	Azure IoT Hub, Blob Storage, Event Hub	ThingSpeak Cloud	Generic cloud platform	Cloud computing services
Data Processing	Advanced ETL pipeline with MATLAB Simulink	Basic aggregation	ML for irrigation	IoT sensors, centralized processing
Data Storage	Blob Storage with containers by sensor type	Centralized in ThingSpeak	General cloud storage	Real-time cloud storage
Sensors Used	Light, Moisture, Temp, N, P, DHT11 for K, CO2 sensors	temp/humidity	Soil moisture, temp, UV	Temp, humidity, soil moisture, pH
Hardware	Raspberry Pi for sensor integration	Raspberry Pi Model B	Zig bee-enabled sensors	Raspberry Pi with sensors
Real-Time Feedback	Dynamic adjustments via MATLAB Simulink	No feedback loop	Automated irrigation	Manual/automated control
Predictive Analytics	Gradient Boosting, SVM, Deep Neural Networks	No predictive analytics	ML for irrigation	Basic ML for optimization
Monitoring Tools	Power BI dashboards, real-time alerts	Web/mobile apps	Mobile apps, SMS alerts	Web/mobile monitoring
Decision Support	Predictive models, real-time alerts	Alerts based on thresholds	Automated decisions	Limited decision support
Security	Secure transfer with Azure authentication	Basic security	Secure communication	IoT security protocols

LITERATURE REVIEW

Scalability	Highly scalable for large-scale	Small-scale applications	Moderate scalability	Scalable for small/medium farms
Innovations	ETL pipeline, predictive analytics	Cost-effective, simple monitoring	IoT, ML for irrigation	Combines IoT, ML, cloud
Limitations	Might be Higher cost for larger scale	Limited analytics, basic system	Focuses on irrigation	Limited predictive analytics

2.5. Plant Growth Modeling Approaches

2.5.1. Classical Growth Models (Logistic, Richards, and Gompertz equations)

Logistic growth models play a crucial role in controlled environment agriculture by enabling data-driven approaches to optimize crop growth conditions (Tantalaki et al., 2019). Researchers are actively developing applications that integrate real-time data analytics with these models to improve agricultural practices for sustainable agriculture (Sahoo et al., 2024). The integration of logistic growth models allows precise control over variables such as light, temperature, and nutrients, enhancing resource management and increasing crop yields (Mana et al., 2024).

The logistic growth model, introduced by Verhulst (1838), has served as a foundational framework for understanding population dynamics under resource constraints. The model is defined by two key parameters: carrying capacity (K) and growth rate (r), where K represents the maximum sustainable population size, and r reflects the intrinsic growth rate (Lotka, 1925; Volterra, 1926). The standard differential form is expressed in Equation (1):

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) \quad (1)$$

Where N = population size, r = intrinsic growth rate, and K = carrying capacity.

The logistic growth equation in Equation (2) also provides the following analytical form:

$$X(t) = \frac{K}{1 + \frac{K-X_0}{X_0} e^{-rt}} \quad (2)$$

Where:

X_0 is the initial value at $t = 0$, and e^{-rt} is the exponential decay term determining the rate at which $X(t)$ approaches the carrying capacity K.

LITERATURE REVIEW

The Gompertz model, a sigmoid function widely used in plant growth analysis, provides a robust framework for characterizing growth patterns under controlled conditions. Two common parametrizations include:

Volume growth model in Equation (3):

$$y_1 = a_1 e^{-e^{b_1 - c_1 x}} \quad (3)$$

Where: y_1 = plant volume (cm^3), x = time (days), a_1 = asymptotic maximum volume (cm^3), b_1 = dimensionless displacement parameter, c_1 = growth rate coefficient (day^{-1})

Lateral area growth model in Equation (4):

$$y_2 = a_2 e^{-e^{b_2 - c_2 x}} \quad (4)$$

Where: y_2 = side area (cm^2), a_2 = maximum potential side area (cm^2), b_2 = dimensionless displacement parameter, c_2 = growth rate coefficient (day^{-1}), x = time (days)

Logistic models have been pivotal in agricultural research, particularly in analyzing crop growth under constrained environments. For instance, they have been used to simulate plant height, biomass accumulation, and nutrient dynamics. In recent years, modified logistic equations have been highlighted for their application in simulating growth dynamics in photosynthetic organisms, integrating environmental variables such as light intensity and nutrient availability (Takaichi & Kawano, 2016, 2023).

2.5.2. Stochastic Model Equation

The stochastic differential equation (SDE) captures both the deterministic trends in environmental conditions and the random fluctuations.

The general form in Equation (5) for each environmental factor $X(t)$ (such as light, water, or temperature) (Niinemets, 2016) of the stochastic equation is:

$$dX_t = \mu(X_t, t) dt + \sigma(X_t, t) dW_t \quad (5)$$

Where $\mu(X_t, t)$ represents the drift term modeling the deterministic trend, $\sigma(X_t, t)$ denotes the diffusion term accounting for random fluctuations, and dW_t corresponds to the Brownian motion. This equation provides a robust framework for simulating dynamic and uncertain environmental conditions.

In agricultural modeling, the general SDE framework can be adapted to represent environmental inputs including light, water, and temperature by tailoring the deterministic component (μ) and stochastic variability (σ) to the specific dynamics of each parameter. For light, μ may represent the planned exposure schedule, while σ accounts for fluctuations such as lighting failures. For water, μ corresponds to scheduled irrigation levels, whereas σ captures inconsistencies in absorption rates. Similarly, for temperature, μ represents expected heating or cooling trends, while σ models unexpected variations such as equipment malfunctions.

LITERATURE REVIEW

By solving the SDEs, multiple environmental scenarios can be simulated under both expected and random influences (Evers et al., 2018). Each generated dataset produces different outcomes based on the stochastic components, enabling probabilistic predictions of future environmental conditions (Balduzzi et al., 2017).

2.5.3. Empirical Modeling for Plant Growth Dynamics

Empirical models are widely used in agricultural and plant-growth studies to establish relationships between environmental variables and crop performance indicators. These models rely on experimentally recorded datasets of plant growth metrics to estimate coefficients $b_1, b_2, b_3, \dots, b_n$ associated with different environmental factors, thereby providing a functional relationship between system inputs and plant responses. Such models are commonly applied to estimate biomass production, plant height, and growth dynamics, and are often validated using simulation or experimental results.

The empirical model represented by Equation (6) is a commonly used mathematical formulation in plant-growth analysis (Jones et al., 2017):

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n + e \quad (6)$$

where Y is the target growth metric (as biomass yield in grams (g) or plant height in centimeters (cm)); $X_1, X_2, X_3, \dots, X_n$ are input variables such as light intensity ($\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$), water intake (L), nutrient concentration (ppm), and temperature ($^{\circ}\text{C}$), b_0 is the intercept (expressed in the same unit as Y); $b_1, b_2, b_3, \dots, b_n$ are regression coefficients whose units depend on the ratio between Y and each corresponding X_k (as $\text{g}/(\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1})$), and e is the error term representing unexplained variability (expressed in the same unit as Y).

Simulation-based empirical approaches are also commonly used to identify optimal environmental conditions and resource thresholds for plant growth. By evaluating multiple simulation scenarios, studies can determine the maximum effective levels of inputs such as water, light, and nutrients beyond which additional increases may lead to diminishing returns rather than efficiency improvements (Reich, 2014; Kolhar & Jagtap, 2023). Such simulation frameworks support resource-efficient cultivation strategies and optimization in controlled environment agriculture.

2.5.4. Malthusian Exponential Growth Equation

The Malthusian exponential growth model describes population (or biomass) dynamics under the assumption of unlimited resources and the absence of external constraints, where the rate of increase remains proportional to the current population size (Peralta et al., 2024). This model is expressed by the Differential Equation in (7):

$$\frac{dP(t)}{dt} = r \cdot P(t) \quad (7)$$

where $P(t)$ represents the population size or plant biomass at time t , and r is the intrinsic growth rate (time^{-1}).

LITERATURE REVIEW

The analytical solution of this equation yields exponential growth as shown in Equation (8):

$$P(t) = P_0 e^{rt} \quad (8)$$

with P_0 denotes the initial population or biomass value.

The Malthusian framework has been widely applied to model biological and agricultural growth processes, including plant development based on real-time sensor data (Bushuev et al., 2024). Although simple and computationally efficient, the model is primarily suitable for early growth stages or environments where resource limitations are minimal. In controlled agricultural systems involving environmental feedback and resource constraints, extended formulations such as logistic growth models are generally preferred.

2.5.5. Logit Function for Regression Modeling

The logit function is a mathematical transformation commonly used in statistical modeling to describe the relationship between a binary or proportional dependent variable and a set of input factors. It transforms probabilities from the bounded interval (0, 1) into an unbounded continuous scale, enabling the modeling of categorical or bounded outcomes through logistic regression methods. The logit function is expressed as Equation (9):

$$\text{logit}(P_t) = \ln \left(\frac{P_t}{1-P_t} \right) \quad (9)$$

where P_t represents the probability or proportion of the target outcome. This transformation establishes a nonlinear relationship between probabilities and predictor variables, facilitating the estimation of coefficients that quantify the influence of each factor on the modeled outcome (Domínguez-Almendros et al., 2011).

2.5.6. Limitations of Single-Dimensional and Disconnected Modeling Approaches

Early growth models such as the Verhulst Logistic Growth Model (1838) and the Lotka–Volterra models (1925, 1926) primarily address single-dimensional phenomena, including population growth and interspecies competition. These foundational approaches were not designed to represent the physiological integration observed in higher plants, where traits such as biomass, height, chlorophyll content, and leaf area exhibit nonlinear interdependencies (Poorter et al., 2011).

Similarly, stochastic models capture environmental variability but generally treat growth factors independently, without considering interactions among plant traits (Niinemets, 2016). Empirical regression models estimate isolated input–output relationships but are limited in representing dynamic feedback mechanisms between physiological variables such as chlorophyll content and leaf area (Jones et al., 2017). Optimization-based approaches can identify resource thresholds and optimal environmental conditions; however, they often assume independent trait responses and do not fully incorporate coupled carrying-capacity constraints observed in functional–structural plant

LITERATURE REVIEW

modeling studies (Evers et al., 2018). Likewise, the Malthusian exponential model does not account for resource limitations or trait covariation.

The application of logistic functions to plant growth modeling (Gregorczyk, 1991) also remains largely restricted to isolated growth parameters. Similar limitations have been reported in photosynthetic modeling, where single-trait approaches fail to capture light-response plasticity across multiple plant species (Niinemets, 2016). Although the framework proposed by Thornley and France (2005) provides greater flexibility, it remains limited in accurately representing complex trait interactions, particularly under varying nutrient conditions (Balduzzi et al., 2017). Furthermore, recent reviews indicate that even advanced machine-learning approaches may struggle to capture physiological trait covariation without explicit biological constraints (Jones et al., 2017; Kolhar & Jagtap, 2023).

In addition, the dynamic nature of carrying capacity (K_t) and growth rate (r), influenced by environmental variability, nutrient depletion, seasonal changes, and temperature fluctuations, remains insufficiently addressed in many existing plant-growth modeling approaches.

2.6. Interdependence Modeling in Plant Growth

Previous studies in plant allometry have demonstrated that plant traits such as biomass, height, leaf area, and chlorophyll content are not independent but exhibit interdependent relationships (Niklas, 1994; West et al., 1997).

These interactions can be represented as in Equation (10):

$$H(t) = g(B(t)) \quad (10)$$

where plant height is expressed as a function of biomass. Such relationships frequently follow allometric or power-law scaling patterns, in which one physiological trait scales as a power function of another. Biological scaling laws have shown that plant height commonly exhibits an allometric relationship with biomass (Enquist et al., 2007), often represented as in Equation (11):

$$H(t) = \alpha_H B(t)^{\beta_H} \quad (11)$$

where α_H and β_H are allometric scaling parameters describing the proportional relationship between plant height and biomass.

2.7. Statistical and Machine Learning Methods for Predictive Analysis

2.7.1. Time Series Analysis

Time-series regression models are widely used in agricultural and environmental studies to analyze plant growth dynamics and predict future development based on historical observations and environmental conditions. These approaches model plant growth as a function of time-dependent environmental variables, enabling the analysis of trends, temporal variability, and seasonal effects (Kumar, 2024).

LITERATURE REVIEW

A general form of the model is expressed in Equation (12):

$$\text{Growth} = a(t) \cdot \text{Light} + b(t) \cdot \text{Moisture} + c_1(t) \cdot \text{Nitrogen} + c_2(t) \cdot \text{Phosphorus} + c_3(t) \cdot \text{Potassium} + d(t) \cdot \text{Temperature} + \epsilon \quad (12)$$

Where: t : time (days), $a(t)$: time-varying coefficient representing the sensitivity of plant growth to light intensity $\left[\frac{\text{g}}{\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}} \right]$, $b(t)$: time-varying coefficient associated with moisture availability $\left[\frac{\text{g}}{\%} \right]$ (or $\frac{\text{g}}{\text{m}^3/\text{m}^3}$), $c_1(t)$, $c_2(t)$, $c_3(t)$: time-dependent coefficients representing the influence of nitrogen, phosphorus, and potassium, respectively $\left[\frac{\text{g}}{\text{mg/L}} \right]$, $d(t)$: time-varying coefficient capturing the effect of temperature $\left[\frac{\text{g}}{^\circ\text{C}} \right]$, $\epsilon(t)$: stochastic error term accounting for unobserved variability and measurement noise [g]

Time-series analysis enables the identification of temporal patterns, seasonal behavior, and dynamic relationships among environmental variables and plant responses. In agricultural IoT systems, these models are commonly combined with real-time sensor networks and cloud-based platforms for continuous environmental monitoring and predictive growth analysis (Wei et al., 2022).

2.7.2. Multivariate Linear Regression for Resource Contributions

To quantify the relationship between the independent variables (light, moisture, temperature, and nutrient levels) and the dependent variable (plant growth), multivariate linear regression was utilized (Delsole & Tippet, 2022). This statistical technique enables the estimation of the impact of multiple independent variables on a single dependent variable, providing a comprehensive understanding of their collective influence. The predicted plant growth is computed using Equation (13):

$$\text{Growth} = a \cdot \text{Light}' + b \cdot \text{Moisture}' + c_1 \cdot \text{Nitrogen}' + c_2 \cdot \text{Phosphorus}' + c_3 \cdot \text{Potassium}' + d \cdot \text{Temperature}' + \epsilon \quad (13)$$

where:

a, b, c_1, c_2, c_3, d : regression coefficients representing the contribution of each normalized environmental variable to plant growth [g].

$\text{Light}', \text{Moisture}', \text{Nitrogen}', \text{Phosphorus}', \text{Potassium}', \text{Temperature}'$: normalized (dimensionless) input variables.

ϵ : stochastic error term accounting for unexplained variability [g].

The regression model is generally solved by minimizing the sum of squared prediction errors as in Equation(14):

$$\text{Minimize } \sum_{i=1}^n (\text{Growth}_i - \widehat{\text{Growth}}_i)^2 \quad (14)$$

LITERATURE REVIEW

where Growth_i represents the observed plant growth value and $\widehat{\text{Growth}}_i$ denotes the predicted growth value for observation i .

Prior to regression analysis, environmental variables are commonly normalized to account for differences in units and numerical scales. One frequently used normalization method is z-score standardization according to Equation (15) :

$$X' = \frac{X - \mu}{\sigma} \quad (15)$$

where μ and σ denote the mean and standard deviation of the corresponding variable. This normalization ensures that all predictors contribute comparably to the regression analysis, improves numerical stability during model optimization, and allows for meaningful comparison of regression coefficients representing the relative influence of each environmental factor on plant growth.

2.7.3. Random Forest Model

The Random Forest model is widely used in agricultural prediction tasks to estimate crop yield and plant growth based on environmental variables such as light intensity, moisture content, temperature, and nutrient levels. This machine-learning approach constructs multiple decision trees using randomly selected subsets of the dataset and randomized feature selection at each split, thereby improving predictive accuracy and reducing overfitting (Pan et al., 2024).

The final prediction for a new observation x' is obtained by averaging the outputs of all decision trees, as expressed in Equation (16):

$$\hat{y}(x') = \frac{1}{B} \sum_{b=1}^B f_b(x') \quad (16)$$

where B is the number of trees and $f_b(x')$ is the prediction from the b -th tree.

2.7.4. Support Vector Machines (SVM)

Support Vector Machines (SVMs) are supervised machine-learning algorithms widely used for classification tasks, particularly in high-dimensional feature spaces (Jiang et al., 2021). The primary objective of SVMs is to determine an optimal hyperplane that separates data into distinct classes while maximizing the margin between the hyperplane and the nearest training samples.

Rather than directly mapping input variables X to binary outputs $Y \in \{-1, 1\}$, SVMs construct a real-valued decision function $f: X \rightarrow \mathbb{R}$, where the predicted class label is determined by the sign of $f(X)$. The margin, defined as $Y \cdot f(X)$, represents the confidence of classification.

The optimization process seeks to minimize the norm of the weight vector while ensuring correct classification of the training samples, as expressed in Equation (17):

$$Y_i(w \cdot X_i + b) \geq 1 \quad \forall i \quad (17)$$

LITERATURE REVIEW

where X_i denotes the input feature vector of sample i , $Y_i \in \{-1, 1\}$ represents the class label, w is the weight vector defining the separating hyperplane, and b is the bias term.

To address nonlinearly separable datasets, SVMs employ kernel functions such as the radial basis function (RBF) and polynomial kernels, which implicitly project input data into higher-dimensional feature spaces through kernel functions $K(x_i, x_j)$. Due to their robustness against overfitting and their effectiveness with limited training data, SVMs are widely applied in agricultural prediction and classification tasks, including crop-yield estimation and environmental condition analysis.

2.7.5. Deep Neural Networks (DNN)

Deep Neural Networks (DNNs) are a class of machine-learning algorithms designed to model complex nonlinear relationships between input features and output targets (LeCun et al., 2015).

A DNN consists of multiple layers, including an input layer, one or more hidden layers, and an output layer. Each layer is composed of interconnected neurons, where each connection is associated with a weight and bias term.

Fully connected feed-forward architectures are commonly trained using stochastic gradient descent optimization techniques to minimize a selected loss function. DNNs are widely applied to capture nonlinear relationships between environmental parameters and crop-growth responses, enabling accurate prediction and pattern recognition from large-scale agricultural datasets (LeCun et al., 2015).

Fig. 2.17 illustrates the principal computational processes involved in deep neural networks:

- (1) Forward-pass transformations from input to output layers.
- (2) Backpropagation of error derivatives through the network using the chain rule.

The figure also demonstrates how gradients ($\partial E / \partial w$) are computed layer by layer to optimize network weights during training. Sigmoid activation functions introduce nonlinearity within hidden layers, enabling the network to learn complex patterns and feature interactions.

LITERATURE REVIEW

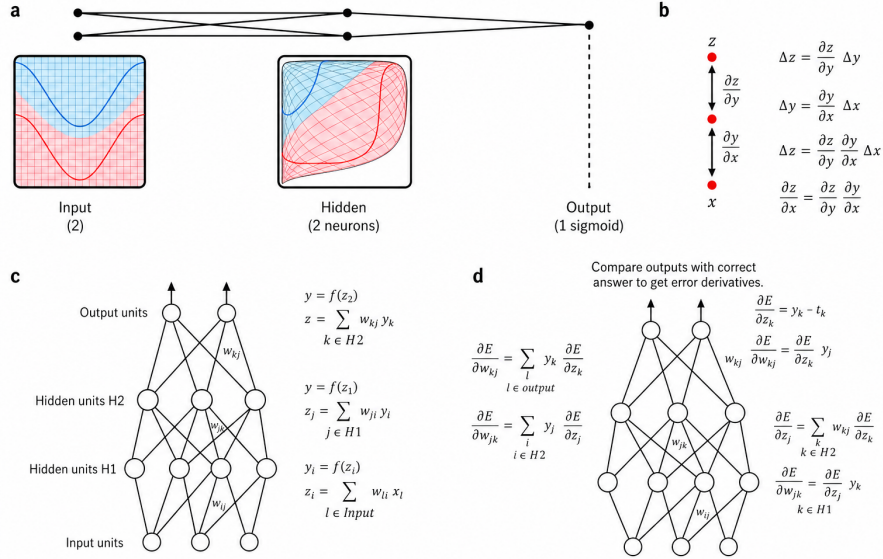


Fig. 2.17. Architecture of a multilayer neural network with backpropagation (LeCun et al., 2015).

2.7.6. Gradient Boosting

Gradient Boosting is an ensemble machine-learning technique widely used for both regression and classification tasks. The method operates by sequentially training weak learners, where each new learner attempts to correct the residual errors produced by the previous model, thereby progressively improving predictive performance (Friedman, 2000). The technique minimizes a predefined loss function $\ell(f)$, which measures the discrepancy between predicted and observed values. At each iteration, a new weak learner $h_t(x)$, scaled by a learning rate α , is added to the current ensemble $f_t(x)$ according to the update rule shown in Equation (18):

$$f_{t+1}(x) = f_t(x) + \alpha h_t(x) \quad (18)$$

where $f_t(x)$ denotes the ensemble prediction at iteration t , $h_t(x)$ represents the weak learner, α is the learning rate controlling the contribution of each learner, and t is the boosting iteration index.

Recent studies have investigated the application of Gradient Boosting and other machine-learning techniques to crop-yield prediction and precision agriculture under controlled-environment conditions. Research in this area increasingly focuses on integrating real-time IoT sensor data with adaptive machine-learning frameworks and cloud-based analytics to support data-driven agricultural management. ETL solutions such as Apache NiFi are commonly employed to facilitate real-time data ingestion, transformation, and processing from distributed IoT sensors.

Machine-learning approaches including Support Vector Machines (SVMs), Gradient Boosting, and Deep Neural Networks (DNNs) have been applied to high-resolution environmental datasets containing variables such as light intensity, soil moisture, temperature, and nutrient levels, together with corresponding plant-growth metrics.

LITERATURE REVIEW

2.8. Evolution of Plant Growth Models in Agricultural Applications

Table 2.2 summarizes the development of logistic growth models in agricultural research, from its foundational origins (Verhulst, 1838) to contemporary adaptations. The comparison focuses on how successive approaches have addressed or failed to address two persistent challenges:

- a. Modeling multidimensional plant traits.
- b. Accounting for dynamic environmental conditions.

Table 2.2. Overview of Logistic Growth Models and Their Limitations in Agricultural and Ecological Contexts

Study (Model)	Year	Limitation
Verhulst Logistic Growth Model (Verhulst, 1838)	1838	Focuses on single-dimensional traits; lacks dynamic updates to carrying capacity (K) and growth rate (r) under changing environmental conditions.
Lotka Logistic Models (Lotka, 1925)	1925	Simplistic representation of interspecies competition; does not account for multidimensional traits or resource depletion effects over time.
Volterra (Volterra, 1926)	1926	Early work on competition models, limited to basic logistic growth, lacks modern ecological or biological trait considerations.
Gregorczyk (Gregorczyk, 1991)	1991	Limited to single traits; does not incorporate advanced environmental factors or interdependent plant traits like chlorophyll or biomass.
Goto, Takaichi, and Kawano (Goto et al., 2015)	2015	Focused on fire-fighting foams and aquatic organisms, limiting its scope for plant growth modeling or other environmental contexts.
Takaichi and Kawano (Takaichi & Kawano, 2016)	2016	Focused on eco-toxicity evaluation for microbial organisms; lacks general applicability to multicellular organisms or dynamic environmental conditions.
Thornley and France (Thornley & France, 2005)	2005	Open-ended logistic function; lacks precision for incorporating multidimensional traits and interdependent variables such as nutrient and environmental constraints.
Thornley, Shepherd and France (Thornley et al., 2007)	2007	A logistic function model for plant growth but lacks deeper integration of environmental interactions and multidimensional trait modeling.
Gachoki et al. (Gachoki et al., 2022)	2022	Models plant growth using logistic and other curves; limited in handling interdimensional dependencies like plant height and leaf area in controlled environments.
Takaichi and Kawano (Takaichi & Kawano, 2023)	2023	Limited to photosynthetic organisms; struggles to integrate interdependent plant traits (height, biomass) into a unified multidimensional framework.

LITERATURE REVIEW

2.9. Performance Metrics and Indicators

2.9.1. Photosynthetic Efficiency

Photosynthetic efficiency metrics evaluate how effectively plants convert light energy into biomass:

- Photosynthetically Active Radiation (PAR): Measures light wavelengths that plants can utilize for photosynthesis. PAR is fundamental in determining light quality in growth environments.
- Photosynthetic Photon Flux Density (PPFD): Quantifies the number of PAR photons reaching a unit area per second ($\mu\text{mol}/\text{m}^2/\text{s}$) PPFD is crucial for assessing the efficacy of artificial lighting systems in indoor farms.
- Light-Use Efficiency (LUE): The ratio of biomass produced to light absorbed. High LUE indicates optimal growth conditions, whereas low LUE suggests inefficiencies in light delivery or plant stress.

2.9.2. Root Mean Square Error

To quantitatively evaluate the performance of predictive and simulation models, several statistical metrics are commonly employed, including the Root Mean Square Error (RMSE), the Coefficient of Determination (R^2), and the Pearson correlation coefficient (Hodson, 2022).

The Root Mean Square Error (RMSE), shown in Equation (19), quantifies the average deviation between simulated and observed values, where lower RMSE values indicate better predictive accuracy:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{\text{simulated},i} - X_{\text{observed},i})^2} \quad (19)$$

2.9.3. Coefficient of Determination

The Coefficient of Determination (R^2), presented in Equation (20), evaluates the goodness-of-fit by measuring the proportion of variance in the observed data explained by the model (Gao, 2023):

$$R^2 = 1 - \frac{\sum (X_{\text{simulated}} - X_{\text{observed}})^2}{\sum (X_{\text{observed}} - \bar{X})^2} \quad (20)$$

The Pearson correlation coefficient, expressed in Equation (21), measures the strength and direction of the linear relationship between two variables:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (21)$$

where X_i represents individual environmental inputs such as light intensity or moisture, Y_i denotes the corresponding plant-growth value, $\bar{X}$ and $\bar{Y}$ are the mean values of the respective variables, and n is the number of observations.

LITERATURE REVIEW

2.9.4. Biomass Conversion Metrics

In addition to statistical evaluation metrics such as Mean Squared Error (MSE) and R-Squared (R^2), there are other biological metrics that evaluate the efficiency of converting resources (light, energy, water) into harvestable biomass in indoor farming systems shown in Table 2.3:

Table 2.3. Biomass Conversion Efficiency Metrics in Controlled-Environment Agriculture

Metric	Full Name	Formula	Focus	Scope	Suitability for Indoor Farming
RUE	Radiation Use Efficiency	$RUE = \frac{\text{Biomass}}{\text{Intercepted PAR}}$	Conversion of intercepted PAR to biomass	Field and greenhouse systems	Limited indoors
NPP	Net Primary Productivity	$NP = GPP - R_a$ GPP: Total carbon assimilated R_a : Energy “cost” of plant metabolism	Net biomass production	All ecosystems	Indirect light focus
CYE	Crop Yield Efficiency	$CYE = \frac{\text{Crop Yield (kg)}}{\text{Intercepted PAR (MJ)}}$	Yield-based productivity metric	Crop yield studies	Dependent on multiple factors

2.9.5. Performance Metrics

The confusion matrix is a fundamental tool for evaluating classification model performance, providing a detailed representation of correct and incorrect predictions (Chifu et al., 2014). As illustrated in Fig. 2.18, the matrix compares predicted classes with observed classes, enabling the calculation of several important evaluation metrics.

This framework is particularly valuable in agricultural applications, where the consequences of false positives (FP), such as unnecessary pesticide application, and false negatives (FN), such as undetected disease outbreaks, must be carefully balanced.

Confusion Matrix

		Predicted Class	
		No	Yes
Observed Class	No	TN	FP
	Yes	FN	TP

Fig. 2.18. Confusion matrix for binary classification performance evaluation (Chifu et al., 2014).

LITERATURE REVIEW

The confusion matrix enables the calculation of four key evaluation metrics commonly used to assess classifier performance, where:

- a. TP (True Positives): Correctly predicted positive cases.
- b. TN (True Negatives): Correctly predicted negative cases.
- c. FP (False Positives): Incorrectly predicted positive cases.
- d. FN (False Negatives): Incorrectly predicted negative cases.

Accuracy: Measures the proportion of correctly predicted instances in the dataset.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: Measures the accuracy of positive predictions.

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall: Measures the ability to find all positive instances.

$$\text{Recall} = \frac{TP}{TP+FN}$$

F1 Score: The harmonic mean of precision and recall, providing a balance between the two.

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

2.10. Growth Metrics

2.10.1. Plant Height

Plant height is an important metric used to assess plant growth and development in controlled-environment agriculture. It is commonly measured as the vertical distance from the base of the plant to its highest point, typically expressed in centimeters (Jiang et al., 2016; Wang et al., 2018). Plant height provides insight into plant responses to environmental conditions and resource availability. Monitoring plant height over time supports the optimization of environmental inputs and contributes to efficient and productive crop management in controlled agricultural systems (Hu et al., 2018).

Plant height can be modeled as a function of environmental factors such as light intensity, water availability, nutrient supply, and temperature. A general empirical model is expressed in Equation (22):

$$H = b_0 + b_1L + b_2W + b_3N + b_4T + e \quad (22)$$

Where H represents plant height [cm], b_0 is the intercept representing baseline plant height when all inputs are zero [cm], b_1 is the coefficient representing the sensitivity of plant height to light

LITERATURE REVIEW

intensity $\left[\frac{\text{cm}}{\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}}\right]$, b_2 is the coefficient associated with water input $\left[\frac{\text{cm}}{\text{L} \cdot \text{day}^{-1}}\right]$, b_3 is the coefficient representing the effect of nutrient input $\left[\frac{\text{cm}}{\text{g} \cdot \text{day}^{-1}}\right]$, b_4 is the coefficient capturing the influence of temperature $\left[\frac{\text{cm}}{^\circ\text{C}}\right]$, L is light intensity $[\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}]$, W is water input $[\text{L} \cdot \text{day}^{-1}]$, N is nutrient input $[\text{g} \cdot \text{day}^{-1}]$, T is temperature $[^\circ\text{C}]$, e is error term accounting for unexplained variability $[\text{cm}]$.

2.10.2. Biomass Yield

Biomass yield is another important metric used in plant-growth analysis and controlled-environment agriculture. It represents the total fresh weight of plant material produced over a specific growth period and is typically expressed in grams (Szulczewski et al., 2018). Biomass yield provides a direct indicator of plant productivity and reflects the combined effects of environmental conditions, resource availability, and plant health status (Guillevic et al., 2024).

As a performance indicator, biomass yield is widely used to evaluate the effectiveness of cultivation strategies and the efficiency of resource utilization. Higher biomass yield values generally indicate favorable environmental conditions and efficient resource allocation, whereas lower values may suggest environmental stress or insufficient resource availability (Li et al., 2019).

Biomass accumulation is strongly influenced by factors including light intensity, water availability, nutrient supply, and temperature. Logarithmic or exponential growth models are commonly used to represent biomass accumulation dynamics and diminishing returns associated with increasing environmental inputs (Price & Munns, 2023). A general empirical formulation is expressed in Equation (23):

$$B = a_0 + a_1 \ln(L') + a_2 \ln(W') + a_3 \ln(N') + a_4 \ln(T') + e \quad (23)$$

Where:

B represents plant biomass $[\text{g}]$, a_0 is the intercept representing baseline biomass $[\text{g}]$, a_1 is the coefficient representing the sensitivity of biomass to light intensity (log-scale effect) $[\text{g}]$, a_2 is the coefficient associated with water input (log-scale effect) $[\text{g}]$, a_3 is the coefficient representing the effect of nutrient input (log-scale effect) $[\text{g}]$, a_4 is the coefficient capturing the influence of temperature (log-scale effect) $[\text{g}]$, L' is Normalized light intensity [dimensionless], W' is Normalized water input [dimensionless], N' is Normalized nutrient input [dimensionless], T' is Normalized temperature [dimensionless].

2.10.3. Leaf Area

In controlled-environment agriculture (CEA), leaf area is considered a critical indicator of plant health, growth rate, and resource-use efficiency because it determines the plant's capacity to intercept light for photosynthesis, directly influencing biomass production and overall development (Weraduwege et al., 2015). Larger and well-developed leaf areas generally contribute to improved light interception, more efficient utilization of water and nutrients, and enhanced air circulation within plant canopies (Addai & Alimiyawo, 2015).

LITERATURE REVIEW

Due to variability in environmental conditions, stochastic modeling approaches are increasingly applied to predict leaf-area dynamics under uncertain growth environments. These models account for both deterministic growth trends and random environmental fluctuations associated with factors such as light intensity, water availability, and temperature variability (Pérez et al., 2022).

Stochastic differential equation (SDE) models are commonly employed for this purpose, incorporating a drift term representing deterministic growth behavior and a diffusion term describing stochastic variability (Kim & van Iersel, 2022). Such approaches provide a robust framework for analyzing and predicting leaf-area development and plant-growth dynamics in controlled-environment agricultural systems.

Equation (24) presents the SDE model for leaf area, where the first term represents deterministic logistic growth and the second term accounts for random environmental fluctuations:

$$dA_t = rA_t \left(1 - \frac{A_t}{K}\right) dt + \sigma A_t dW_t \quad (24)$$

where:

A_t is leaf area [cm^2], r is growth rate [day^{-1}], K is carrying capacity [cm^2], dt is deterministic, time increment [day], σ is noise intensity [$\text{day}^{-1/2}$], dW_t is Wiener increment [$\sqrt{\text{day}}$].

2.11. Data Processing Methodologies

2.11.1. Data Transmission Protocols

While HTTP/REST APIs remain widely used for cloud integration due to their statelessness and compatibility with existing web infrastructures (Omar, 2021), their high overhead makes them less suitable for resource-constrained edge devices. Comprehensive reviews indicate that only 2% of surveyed smart farming systems relied on wired Ethernet (typically HTTP-based), with the majority adopting wireless protocols like LoRaWAN or cellular networks for long-range communication (Navarro et al., 2020; Terence & Purushothaman, 2020). Recent studies further demonstrate that HTTP/REST implementations in Azure IoT and AWS Greengrass exhibit 20% higher latency compared to MQTT in equivalent agricultural applications (Khanh et al., 2022).

Current research underscores MQTT's dominance in smart farming, particularly for real-time telemetry, citing its resilience in low-bandwidth conditions and flexible QoS levels (Khanh et al., 2022; Omar, 2021). The protocol has been identified as the preferred solution for greenhouse monitoring systems, where its broker architecture enables efficient multi-hop mesh topologies that effectively mitigate signal interference challenges (Navarro et al., 2020). Field studies report a 68% adoption rate of MQTT in precision irrigation systems due to its reliable message delivery mechanisms (Farooq et al., 2020).

Recent findings from multiple systematic reviews align with key protocol selection criteria:

- a. Message Queuing Needs: MQTT's QoS levels (particularly QoS-1 for guaranteed delivery) have proven critical for irrigation control systems and other time-sensitive agricultural applications (Khanh et al., 2022; Omar, 2021).

LITERATURE REVIEW

- b. Security Requirements: TLS 1.3 adoption remains limited in legacy agricultural IoT systems, with many deployments still relying on ZigBee's inherent encryption or older TLS versions (Farooq et al., 2020; Omar, 2021).
- c. Payload Optimization: CBOR encoding is emerging as a viable alternative to JSON in precision agriculture deployments, offering significant reductions in latency and bandwidth consumption (Khanh et al., 2022; Omar, 2021).

Recent advancements in edge computing architectures are gradually reducing dependency on HTTP/REST APIs, particularly in UAV-based monitoring systems where MQTT's lightweight protocol is preferred for real-time telemetry (Obaid et al., 2024). However, hybrid approaches combining HTTP for cloud integration with MQTT for field device communication are gaining traction in large-scale agricultural deployments (Navarro et al., 2020).

2.11.2. Data Acquisition Working Protocol

The data acquisition and control protocol in modern aeroponic and greenhouse systems represents a sophisticated integration of IoT architecture, cloud computing, and precision agriculture technologies. As shown in Fig. 2.19. , the system employs a hierarchical structure where sensor nodes continuously monitor critical growth parameters including ambient temperature, relative humidity, photosynthetically active radiation (PAR), CO₂ concentration, and hydroponic nutrient solution characteristics (EC, pH, and temperature). This sensor network sends raw data to an edge computing unit (typically Arduino or Raspberry Pi-based) that performs initial data validation and preprocessing before transmission via WiFi/LoRaWAN to the cloud infrastructure.

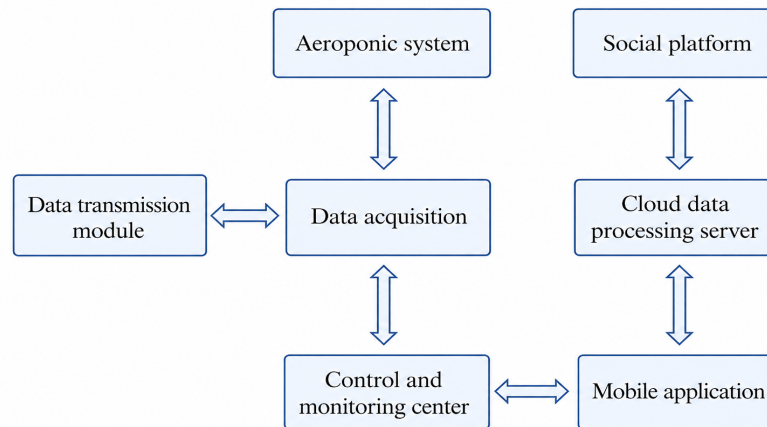


Fig. 2.19. Schematic protocol of wireless network in the aeroponic system intervention (Lakhia et al., 2018).

Fig. 2.20 provides crucial insight into the real-time control mechanisms, illustrating how the system implements a closed-loop feedback architecture. The process computer receives both sensor data and predefined growth recipes through a REST API interface, then calculates optimal adjustments for multiple actuator systems. These include precision control of: supplemental LED lighting (intensity and photoperiod), thermal screens, ventilation (opening percentage and airflow direction), CO₂ dosing systems (ppm concentration), and fertigation systems (nutrient composition and irrigation timing). The Fig.'s bidirectional arrows emphasize the continuous data exchange that enables dynamic system recalibration for instance, automatically increasing ventilation when humidity exceeds thresholds or modulating nutrient EC based on real-time plant uptake measurements. Advanced features like multispectral imaging (Ryu et al., 2023) integrate with this

LITERATURE REVIEW

framework to provide phenotypic stress indicators that trigger interventions before visible symptoms emerge.

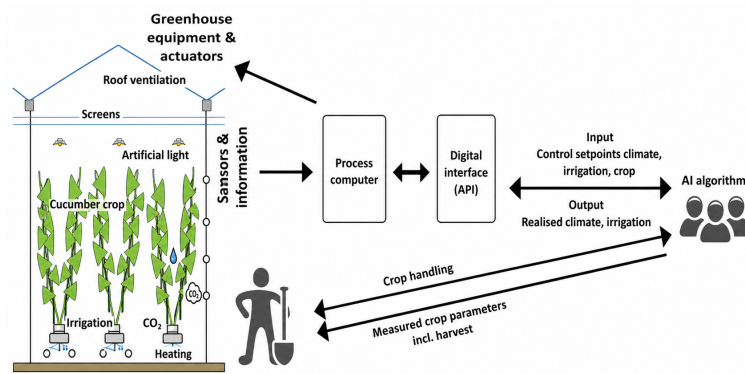


Fig. 2.20. Data exchange by a digital interface (REST API) towards the process computer and the greenhouse, actuators and data from sensors via the same way back (Hemming et al., 2019).

2.11.3. Data training

Recent advances in machine learning for agricultural applications have emphasized the importance of robust model training protocols to ensure reliable performance. A critical component of such protocols involves addressing class imbalance through hybrid sampling techniques, typically combining SMOTE oversampling of minority classes with strategic under sampling of majority classes (Zhao et al., 2025). As demonstrated in Fig. 2.21. , contemporary approaches employ k-fold cross-validation methodologies, where the training data is systematically partitioned to enable comprehensive model evaluation.

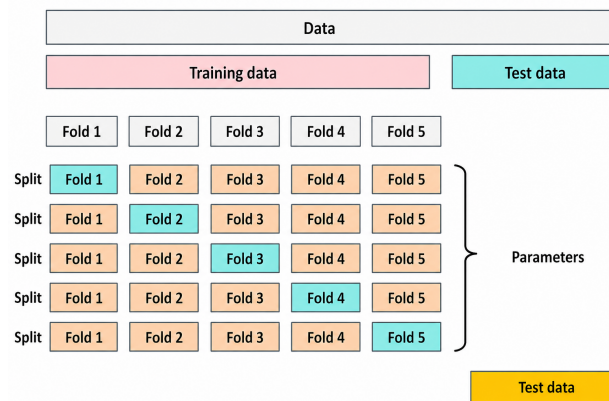


Fig. 2.21. Five-fold cross-validation (Zhao et al., 2025).

This validation framework, illustrated in Fig. 2.21, ensures each data segment serves as both training and validation set across successive iterations, thereby providing more reliable performance estimates while mitigating overfitting risks. Subsequent hyperparameter optimization further enhances model efficacy, with particular attention to architecture-specific parameters including SVM kernel functions, gradient boosting tree structures, DNN layer configurations, and time-series window sizing. The integration of these methodological components represents current

LITERATURE REVIEW

best practices in agricultural machine learning systems, balancing computational efficiency with predictive accuracy for crop performance applications.

2.12. Smart Farming Initiatives: EU and Hungary

2.12.1. EU smart farming initiatives (European Commission, 2023)

There is broad academic consensus that the European Union is a normative leader in institutionalizing smart farming through an interconnected policy-technology-funding paradigm (Bronson, 2022; European Commission, 2023). This approach aligns with what has been termed "mission-oriented innovation governance" (Klerkx et al., 2022), where the Farm to Fork Strategy (European Commission, 2020) and CAP 2023-2027 establish binding sustainability targets (% pesticide reduction) explicitly contingent on digital adoption—a policy design critiqued as technologically deterministic (Darnhofer, 2021) yet praised for its systemic ambition. The technological dimension reflects what has been conceptualized as "agriculture's fourth infrastructure" (Wolfert et al., 2023), where Horizon Europe's €8.9 billion R&D budget and the Digital Europe Programme's Testing Facilities operationalize innovations like the vertical farming system in Fig. 2.22 (European Commission, 2023). This case exemplifies contested claims about precision agriculture's scalability, achieving 100x yield/m³ gains (van der Burg et al., 2024) while replicating what are warned to be "islands of automation" (Carolan, 2024) inaccessible to 78% of sub-20ha EU farms (European Commission, 2023). Infrastructure development mirrors the "datafication of agri-governance" (Jakku et al., 2023), with 5G rollouts and agricultural data spaces (governed by the Data Governance Act) creating what some scholars label "digital lock-ins" (Rotz et al., 2023). Yet as Fig. 2.22 demonstrates, such systems enable real-time resource optimization; a capability studies associate with 20-35% input reduction (Zhang et al., 2024). Persistent adoption barriers substantiate "digital divide" thesis (Prause et al., 2023), particularly among smallholders facing:

- a. Data sovereignty concerns (Wiseman et al., 2023).
- b. Skill asymmetries (Labarthe et al., 2024).
- c. Path dependencies in conventional practices (Sutherland et al., 2024).

LITERATURE REVIEW

The EU's integrative model, while theoretically robust, faces empirical challenges in reconciling its twin goals of technological leadership in Fig. 2.22.'s contrast between high-tech efficiency and limited accessibility (Fielke et al., 2024).



Fig. 2.22. A smart indoor vertical farming system maximising yield efficiency and resource management (European Commission, 2023).

2.12.2. Hungary smart farming initiatives

Hungary has demonstrated significant progress in smart farming through innovative projects like Biopolus's LivingIsland vertical farm in Budapest (Biopolus, n.d.). As shown in Fig. 2.23., this facility (2016-2018) successfully implemented a hybrid aeroponic-hydroponic system that achieved remarkable resource efficiencies, using only 5% of the water required by conventional greenhouses while maintaining year-round production of leafy greens. The project's modular design and IoT-enabled climate control systems exemplify Hungary's technical capacity to implement precision agriculture solutions that align with EU sustainability targets under the Farm to Fork Strategy. However, the initiative also revealed substantial implementation challenges, particularly regarding economic viability - with €2.1 million in startup costs and 68% of operational expenses dedicated to LED lighting. These findings suggest that while Hungary possesses the technical expertise to develop advanced smart farming systems, wider adoption across Central Europe will require funding mechanisms, potentially through CAP digitalization measures, to address the high capital and energy requirements of such technologies. The LivingIsland case study ultimately presents both a model for sustainable urban food production and a cautionary tale about the financial barriers to smart farming implementation in the region.



Fig. 2.23. Hungary's Contribution to Smart Farming: The Biopolus Case (Biopolus, n.d.).

2.13. Critical research gaps

The literature review reveals a persistent lack of integration in the development of IoT-enabled agricultural systems. While significant progress has been made in individual domains, several critical research gaps remain, as outlined below:

a. Limitations of Unidimensional and Static Growth Models:

Classical logistic growth models and several recent plant-growth approaches primarily model single growth variables such as biomass or height independently (Verhulst, 1838; Lotka, 1925; Volterra, 1926; Gregorczyk, 1991; Gachoki et al., 2022). However, these models do not represent the multidimensional and interdependent evolution of physiological traits such as biomass, height, chlorophyll content, and leaf area. In addition, most existing models assume static carrying capacity and growth-rate parameters, failing to account for dynamic environmental variability including nutrient depletion and light stress (Evers et al., 2018; Niinemets, 2016).

b. Limitations of Isolated Environmental Factor Analysis:

Several studies investigated the influence of environmental parameters such as light intensity or temperature individually (Afzali et al., 2021; Jiang et al., 2021); however, the interactive and nonlinear effects among environmental variables and plant maturity remain insufficiently explored. Existing approaches therefore provide limited capability for predicting plant responses under combined environmental conditions (Khanna & Kaur, 2019).

c. Undefined Optimal Canopy Density for Light-Use Efficiency:

Although Light-Use Efficiency (LUE) has been introduced as a critical performance metric in controlled-environment agriculture, current studies provide limited quantitative analysis linking canopy structure and Leaf Area Index (LAI) to lighting efficiency (Khanna & Kaur, 2019). Furthermore, distinctions between lamp-level efficiency and canopy-level light absorption efficiency remain insufficiently addressed, limiting optimization of energy-efficient indoor farming systems.

d. Absence of Data-Driven Composite Plant Health Metrics:

Existing plant-health and performance indices commonly rely on fixed or arbitrarily weighted parameters such as plant height and leaf count. However, current approaches lack adaptive, data-driven frameworks capable of objectively quantifying the relative contribution of physiological traits including biomass, leaf area, and height to overall plant performance and resource-use efficiency.

e. Lack of Precision Nutrient Optimization Models:

Recent agricultural monitoring studies have focused on tracking nutrient concentrations and

LITERATURE REVIEW

environmental variables (Lavanya et al., 2020); however, limited attention has been given to predictive nutrient-dosing models capable of identifying optimal concentration thresholds and diminishing-return regions for plant growth (Jones et al., 2017). Consequently, nutrient application strategies often remain based on static schedules rather than adaptive optimization approaches.

f. Limitations in Dynamic Environmental Adaptation and Trait Coordination:

Although allometric relationships between plant height and biomass have been extensively studied (Niklas, 1994; Enquist et al., 2007), limited research has investigated whether these relationships remain stable under adaptive IoT-based environmental control systems. Existing studies also lack generalized empirical models capable of evaluating coordinated trait development under dynamic controlled-environment conditions.

3. MATERIALS AND METHODS

3.1. Experimental Setup for Plant Growth Evaluation

3.1.1. Environmental Control and Growth Conditions

This study implemented replicates of 8-week experimental setup to evaluate the effectiveness of an IoT-based system in optimizing indoor plant growth, with lettuce (*Lactuca sativa* L.) chosen as the test crop due to its fast growth cycle and sensitivity to environmental factors.

During the experiment, a new set of lettuce (*Lactuca sativa* L.) plants was grown each week, with environmental growing factors, including light intensity, moisture content, temperature, and nutrient levels, systematically controlled and varied weekly, as shown in the UML model of the system in Appendix A4. The growth metrics for each weekly set were recorded at the maturity level to ensure a comprehensive analysis of the relationship between environmental conditions and lettuce (*Lactuca sativa* L.) development. This approach allowed for a robust assessment of how each unique setup influenced plant growth and validated the system's ability to adapt to varying conditions. In this study, light intensity was measured using a standard lux meter (reporting lumens). While lux and lumens are convenient for general illumination monitoring, they do not accurately represent the photosynthetically active photon flux that plants utilize for growth, since plant photosynthesis is driven by photon quantity rather than brightness as perceived by the human eye. To address this, light measurements in lumens were converted to Photosynthetic Photon Flux Density (PPFD; $\mu\text{mol m}^{-2} \text{s}^{-1}$) using a conversion factor derived from published comparisons of lux to PPFD for broad-spectrum LED lighting (Sharakshane, 2018):

$$1000 \text{ lux} \approx 15 \mu\text{mol m}^{-2} \text{s}^{-1} \text{ under white LED spectral distributions}$$

Lumen values were multiplied by 0.015 to estimate corresponding PPFD values. IoT sensors continuously collected data measurements and transmitted them in real-time to Azure Blob Storage, as shown in the pipeline in Fig. 3.1.

MATERIALS AND METHODS

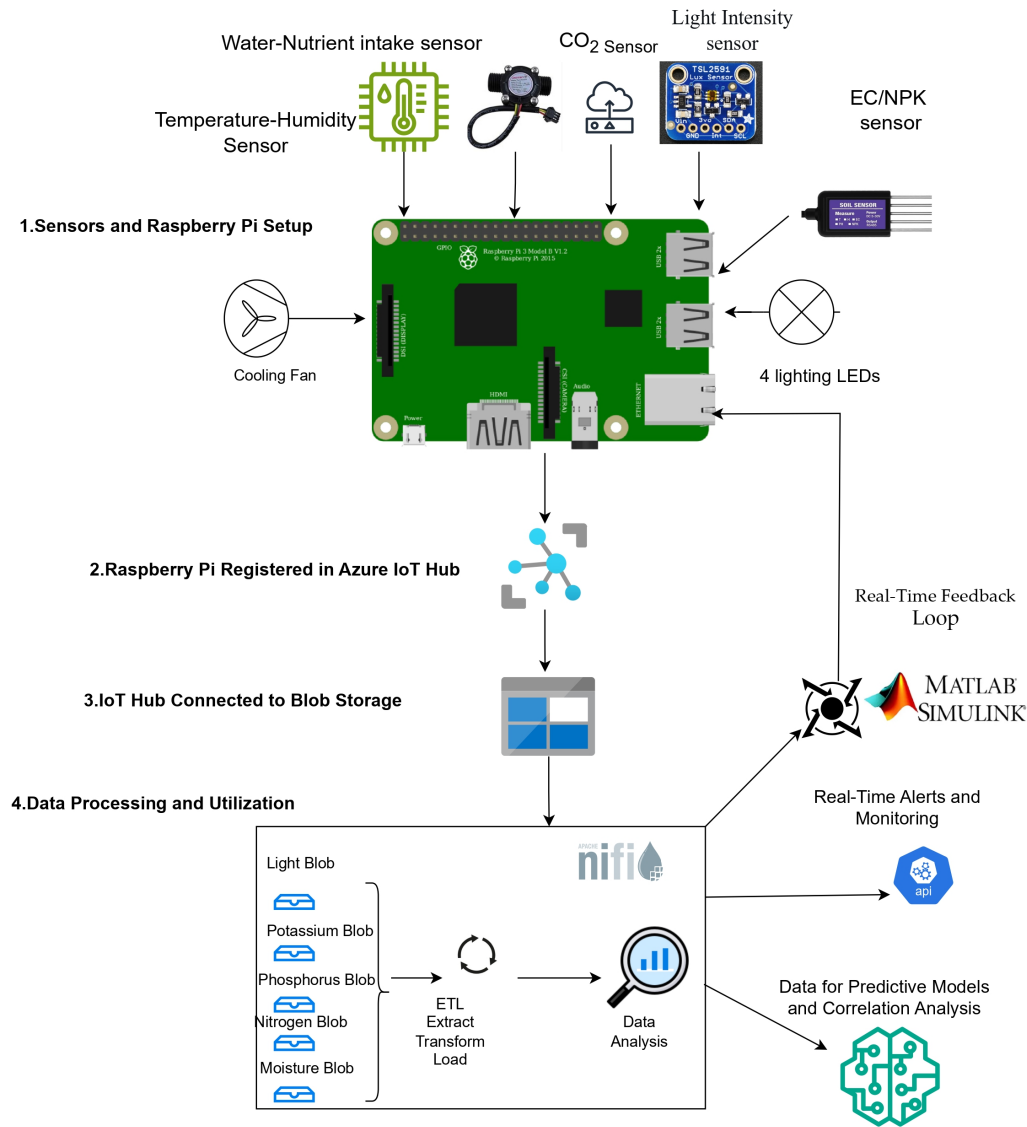


Fig. 3.1 IoT-Based pipeline for plant growth optimization

The Internet of Things (IoT) enables real-time environmental monitoring through connected sensors. In this study, IoT sensors were utilized to measure key environmental parameters, including light intensity, soil moisture, temperature, humidity, CO₂ levels, and nutrient concentrations. These sensors were connected to a Raspberry Pi 4 Model B (4 GB RAM, quad-core ARM Cortex-A72 processor, Raspberry Pi Foundation, Cambridge, United Kingdom) via General Purpose Input/Output (GPIO) pins as shown in Fig. 3.1. for data acquisition and transmission to a cloud-based storage system. The Raspberry Pi 4 Model B was chosen for its computational capabilities, low power consumption, and compatibility with IoT-based agricultural monitoring systems.

Each sensor was selected based on accuracy, sensitivity, and compatibility with IoT-driven farming as shown in Table 3.1 below:

MATERIALS AND METHODS

Table 3.1 Description of the sensors

Measured Parameter	Unit	Applied Sensor Type	Measurement Range	Accuracy	Notes
Temperature	°C	SCD41 (Sensirion)	−40 to 80 °C	±0.5 °C	Integrated with humidity and CO ₂ sensor.
CO ₂ concentration	ppm	SCD41 (Sensirion)	0–5000 ppm	±50 ppm or ±5 %	Same integrated sensor.
Electrical Conductivity (EC)	mS/cm	RS485 Modbus 7-in-1 soil sensor	0–23 mS/cm	±2 %	Also provides pH, temperature.
Nitrogen (N)	mg/kg	RS485 Modbus 7-in-1 soil sensor	0–1999 mg/kg	±2 %	Measured in soil solution.
Phosphorus (P)	mg/kg	RS485 Modbus 7-in-1 soil sensor	0–1999 mg/kg	±2 %	Measured in soil solution.
Potassium (K)	mg/kg	RS485 Modbus 7-in-1 soil sensor	0–1999 mg/kg	±2 %	Measured in soil solution.
Light intensity	Lux (converted to PPFD)	TSL2591 (Adafruit)	188 µlux to 88 klux	±10 %	Converted to PPFD (µmol·m ^{−2} ·s ^{−1}) using factor 0.015 for white LED spectrum.
Light spectrum ratio	–	Not directly measured	–	–	Set by LED controller
Soil moisture (VWC)	%	Capacitive soil moisture sensor (DFRobot)	0–100 %	±5 %	Used for irrigation control.
Water flow	L/min	Flow meter (DFRobot)	1–30 L/min	±3 %	For irrigation volume monitoring.
Visual Plant Traits (Leaf area, height, color)	N/A	Raspberry Pi High Quality Camera (12.3 MP, Sony IMX477) with 6 mm CS-mount wide-angle lens	4056 × 3040 pixels	N/A	Used for image-based phenotyping and growth analysis

Table 3.2 below shows the control system description defining the control logic, in A4, I provided the UML Model showing the interactions within the system.

MATERIALS AND METHODS

Table 3.2 Sensors Control Logic

Parameter	Sensor	Control Logic	Actuator
Temperature	SCD41	PID control (setpoint: 22 °C). If >setpoint, cool; if <setpoint, heat (if available).	Cooling fan (DC) / Heater (resistance)
CO ₂ concentration	SCD41	In Treatment Group C only: maintain 400 ppm ($\pm 3\%$). If CO ₂ <390 ppm, reduce fan speed; if >410 ppm, inject CO ₂ (from compressed bottle).	CO ₂ solenoid valve + Fan speed control
EC / NPK	RS485 Modbus 7-in-1	Maintain EC in range 1.5–2.0 mS/cm (≈ 750 –1000 ppm). If EC low, add nutrient solution; if high, add pure water.	Peristaltic pumps (one for nutrient stock, one for water)
Light intensity	TSL2591	Threshold-based feedback (Section 3.9.1). If PPFD < setpoint (e.g., 300 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), increase PWM duty cycle; if > setpoint, decrease.	COB LED grow lights (PWM-dimmed via MOSFET)
Irrigation (water)	Capacitive soil moisture sensor + Flow meter	If soil moisture < threshold (e.g., 60 % VWC), trigger irrigation until target volume (measured by flow meter) is delivered.	Submersible water pump + Solenoid valve

Table 3.3 defines the methods used to measure the main traits for this study:

Table 3.3 Measurement Methods for Biomass, Height, Chlorophyll, and Leaf Area

Quantity	Symbol (Revised)	Unit	Measurement Method
Biomass	B or B(t)	grams (g)	Fresh weight measured at harvest. Plants were gently blotted and weighed on a digital balance (0.01 g precision).
Plant height	H or H(t)	cm	Measured from substrate surface to the highest leaf tip using a rigid ruler (± 0.1 cm).
Chlorophyll	C or C(t)	SPAD units	Measured with a SPAD-502 Plus chlorophyll meter (Minolta). Three readings per leaf, averaged.
Leaf area	A	cm ²	Image-based measurement: Photographs were processed with PlantCV using binary thresholding and pixel-to-cm ² calibration.

MATERIALS AND METHODS

3.1.2. IoT-Based Data Collection and Storage of IOT Devices

Sensor data was collected hourly between April 2024 and September 2024, with each reading stored as a timestamped JSON file, ensuring chronological order.

The data was then transmitted from the IoT Hub to Azure Event Hub, where it underwent initial processing before being stored in Azure Blob Storage.

Each sensor's data was stored in a dedicated container named according to the corresponding sensor type. For instance, "light" contained light intensity measurements, "moisture content" held soil moisture data, "nitrogen" recorded nitrogen concentrations, "phosphorus" stored phosphorus levels, and "potassium" contained potassium concentration data. To maintain data accuracy and reproducibility, real-time sensor readings were saved in Azure Blob Storage using a timestamped naming format (yyyymmdd_hhmmss), ensuring a structured historical record for tracking as shown in Fig. 3.2. This approach enables data integrity, facilitates historical analysis, and enables scalable data retrieval for model training.

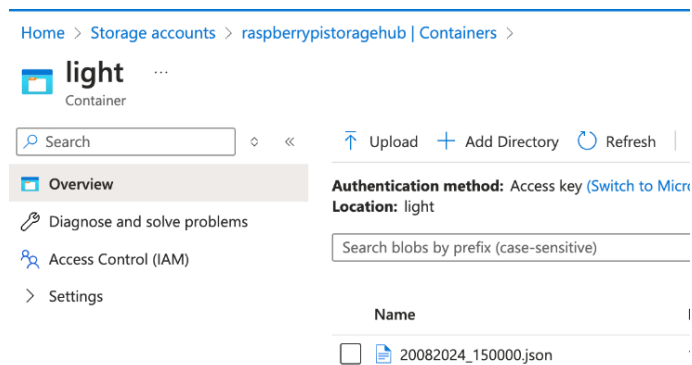


Fig. 3.2. Example of light sensor data in azure blob storage container.

3.1.3. Experimental Conditions

a. Resource-Use Efficiency for Growth Efficiency Optimization

Duration: 30 days.

Environmental Constants: Temperature maintained at 22°C ($\pm 1^\circ\text{C}$), relative humidity at 60% ($\pm 5\%$), CO₂ at 400 ppm.

Chamber: Single prototype chamber with 12 growing cubes, allowing 3 experimental repetitions to complete all treatment combinations over 30 days.

Plant Material: Lettuce (*Lactuca sativa* L.), 12 plants per treatment level, considered 8. Of these, 8 plants were retained for analysis, while 4 additional plants were included as backups to account for potential losses or sensor failures. The backup plants were maintained under identical conditions and continued to grow throughout the 8-week experimental period.

Sampling: Final harvest measurements only (non-destructive monitoring during growth)

Objective: To map individual resource-use efficiency (Growth Efficiency Ratio - GER) and identify optimal input levels for water, light, and nutrients in controlled environment lettuce (*Lactuca sativa* L.) production, Table 3.4 shows the detailed procedure.

MATERIALS AND METHODS

Table 3.4 Experimental design of Resource-Specific Response

Resource Varied	Treatment Levels	Constant Parameters	Plants per Level	Primary Output
Water	5, 7, 9 L/day	Light: 16h, Nutrients: Optimal, T: 22°C	12	Biomass, Yield, GER, Height
Light (Photoperiod)	6, 9, 12 h/day	Water: Optimal, Nutrients: Optimal, T: 22°C	12	Biomass, Yield, GER, Height
Nutrients	3, 4, 5 g/day	Water: Optimal, Light: 16h, T: 22°C	12	Biomass, Yield, GER, Height
Water × Fertilizer	5/7/10 L × 3-11 g	Light: Optimal, T: 22°C	12	Plant Height
Water × Light	5/7/10 L × 6-15 h	Nutrients: Optimal, T: 22°C	12	Plant Height

b. Leaf Area Expansion in response to graded levels of inputs

Duration: 30 days of continuous daily measurements.

Environmental Constants: Temperature: 22°C (±1°C), Relative Humidity: 60% (±5%), CO₂: 400 ppm.

Chamber: Single prototype chamber with parallel treatments.

Plant Material: Lettuce (*Lactuca sativa* L.) , 12 plants per treatment level. Of these, 8 plants were retained for analysis, while 4 additional plants were included as backups to account for potential losses or sensor failures. The backup plants were maintained under identical conditions and continued to grow throughout the 8-week experimental period.

Measurement Method: Daily image capture and analysis using PlantCV for non-destructive leaf area estimation.

Output: Complete 30-day growth trajectories for leaf area development under each resource gradient.

Three-Protocol Comparison for Resource-Efficient lettuce (*Lactuca sativa* L.) Production in Controlled Environments.

Objective: To characterize the daily progression of leaf area expansion in response to graded levels of water, nutrients, and light duration throughout a 30-day vegetative growth period, Table 3.5 shows the detailed procedure.

Table 3.5 Ressource Gradient Experimental Design

Resource Varied	Treatment Levels	Constant Parameters	Measurement Frequency	Data Type
Water	5, 7, 9 L/day	Light: 16h, Nutrients: Optimal, T: 22°C	Daily for 30 days	Leaf Area (cm ²)
Nutrients	3, 5, 7 g/day	Water: Optimal, Light: 16h, T: 22°C	Daily for 30 days	Leaf Area (cm ²)
Light (Photoperiod)	8, 10, 12 h/day	Water: Optimal, Nutrients: Optimal, T: 22°C	Daily for 30 days	Leaf Area (cm ²)

MATERIALS AND METHODS

c. Growth under three resource management strategies.

Duration: Each experiment lasted 8 weeks and was repeated five times for all Treatment Groups.
Environmental Constants: Temperature: 22°C ($\pm 1^\circ\text{C}$), Relative Humidity: 60% ($\pm 5\%$), Soil Moisture: 65-70% VWC. These parameters were kept consistent across all Groups to ensure comparability.

Chamber: A single prototype growth chamber was used sequentially for all trials to minimize environmental variability between experiments.

Plant Material: Lettuce (*Lactuca sativa* L.), 8 plants per trial considered. 12 plants were grown in the chamber per experiment but only 8 plants were retained for analysis, while 4 additional plants were included as backups to account for potential losses or sensor failures.

Sampling: Weekly destructive harvest (1 plant/week) for growth curve construction until final harvest at week 8, Table 3.6 shows the detailed procedure.

Objective: The objective of this experiment is to validate the interdependence between structural plant traits (height, biomass, and leaf area) and to support the four-dimensional logistic growth model by quantifying allometric scaling relationships under both static and adaptive resource conditions. In particular, the study aims to assess whether adaptive, IoT-based regulation preserves coordinated trait development rather than amplifying variability.

Table 3.6 Treatment Groups (A,B,C)

Parameter	Treatment Group A (Baseline Static Protocol)	Treatment Group B (Adaptive Dynamic Protocol)	Treatment Group C (Energy-Efficient Protocol)
System Type	Non-IoT (no sensor feedback)	IoT-based (sensor-driven control)	IoT-based (constrained sensor-driven control)
Control Strategy	Open-loop (no feedback, no adaptation)	Closed-loop (real-time feedback adaptation)	Closed-loop (resource-constrained adaptation)
Treatment Description	Static resource allocation	Dynamically adjusted based on feedback	Resource-constrained adaptive control
Light Protocol	Fixed setpoint: $775 \mu\text{mol m}^{-2} \text{s}^{-1}$ (no adaptation)	Weekly increments: $220 \rightarrow 1300 \mu\text{mol m}^{-2} \text{s}^{-1}$	Weekly increments: $180 \rightarrow 950 \mu\text{mol m}^{-2} \text{s}^{-1}$
Nutrient Protocol	Fixed concentration: 81 ppm (no adaptation)	Weekly increments: $22 \rightarrow 130 \text{ ppm}$	Weekly increments: $18 \rightarrow 95 \text{ ppm}$
CO₂ Protocol	Fixed nominal level: 950 ppm (no feedback control)	Weekly increments: $410 \rightarrow 550 \text{ ppm}$	Weekly increments: $405 \rightarrow 420 \text{ ppm}$
Experiment Replication	5 independent 8-week experiments	5 independent 8-week experiments	5 independent 8-week experiments
Plants per Experiment	8 plants measured weekly	8 plants measured weekly	8 plants measured weekly
Total Plants per Treatment Group	40 plants (8×5 experiments)	40 plants (8×5 experiments)	40 plants (8×5 experiments)

MATERIALS AND METHODS

d. Trait Response Across Light Configurations (A–D)

Duration: 30 days (vegetative stage)

Environmental Constants:

- Temperature: 22 °C (± 1 °C)
- Relative Humidity: 60% ($\pm 5\%$)
- CO₂: 400 ppm
- Chamber: Single prototype chamber

Plant Material: Lettuce (*Lactuca sativa* L.)

Plants per Configuration:

- 12 plants grown
- 8 plants retained for analysis
- 4 backup plants

Objective: To quantify how increasing lighting configurations (A–D) influence biomass accumulation, plant height, and to establish a quantitative link between lighting-system performance with biological responses, using light-use efficiency metrics (LUEP, LUEL) and canopy structure (LAI). Table 3.7 shows the detailed procedure.

Table 3.7 Lighting Configurations

Configuration	LEDs	Photoperiod (h/day)	Intensity ($\mu\text{mol}/\text{m}^2/\text{s}$)	Spectrum (B:R:FR)
A	2	12	150	0.3 : 0.7 : 0.0
B	3	16	250	0.2 : 0.7 : 0.1
C	4	18	350	0.2 : 0.6 : 0.2
D	5	20	400	0.1 : 0.8 : 0.1

Light intensity measurements were collected using a lux meter at two locations: directly at the light source (lamp level) and at canopy level above the plants. The measured illuminance values (lux) were converted into Photosynthetic Photon Flux Density (PPFD, $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) using the standard conversion factor for white LEDs described in Section 3.1.1 (Environmental Control and Growth Conditions). Based on these measurements, emitted Photosynthetically Active Radiation (PAR) was derived from lamp-level measurements, while received PAR was derived from canopy-level measurements. Absorbed PAR was subsequently estimated as a fraction of received PAR according to canopy light interception characteristics. The corresponding light-use efficiency metrics were then calculated, with LUEL defined as the ratio between received and emitted PAR, and LUEP defined as the ratio between absorbed and emitted PAR. Leaf Area Index (LAI) was determined from experimentally measured leaf area and planting density to characterize canopy structure and light interception efficiency. A detailed description of the complete computational procedure, including the parameter relationships, mathematical workflow, and step-by-step

derivation of emitted PAR, received PAR, absorbed PAR, LUEL, and LUEP across Configurations A–D.

For all remaining predictive analytics and machine learning validations presented in this thesis, the models were trained and tested using a consolidated dataset compiled from historical sensor measurements.

3.2. IoT System Architecture and Hardware

3.2.1. WiFi-Based MQTT Communication and Wi-Fi Device Communication with Ethernet-Backhauled MQTT

I implemented the IoT system architecture for lighting control using two primary communication setups. The first configuration (see Appendix A5 - Fig. 8.1.) illustrates a Wi-Fi-enabled integration where the light sensor measures intensity and transmits data to a Raspberry Pi, which then publishes it to Azure IoT Hub. A cloud-based application evaluates these light levels and sends MQTT commands back to the Raspberry Pi to dynamically adjust the LED system increasing intensity when too low, decreasing it when too high, or maintaining an optimal level. A second, more detailed architecture (see Appendix A5 - Fig. 8.2.) specifics the IoT-based control loop, where the light sensor relays data to a Raspberry Pi functioning as an MQTT broker. This broker publishes the measurements to Azure IoT Hub over a wired Ethernet connection for cloud processing and storage. Subsequently, MQTT commands are returned to the actuator, which executes the precise lighting adjustments required to maintain the target intensity for optimal plant growth.

In this configuration, the Wemos D1 R3 microcontroller acts as a Wi-Fi-based sensor and actuator node, digitizing the measured input data and transmitting it wirelessly to the Raspberry Pi gateway (see Appendix A6 – Fig. 8.3). The Raspberry Pi then relays the data to Azure IoT Hub using a wired Ethernet connection, which constitutes the Ethernet-based segment of the communication architecture. Control commands generated by the cloud application are returned via the same Ethernet backhaul to the Raspberry Pi and subsequently transmitted over Wi-Fi to the actuator node, which executes the required lighting adjustments to maintain the target intensity (see Appendix A6 – Fig. 8.4). No external Ethernet module was used on the Wemos D1 R3 device.

3.2.2. Time Responsiveness Analysis

To compute the end-to-end responsiveness of my MQTT-based IoT system, I model the overall delay, Ψ , as the sum of the delays in each segment of the communication path as shown in Equation (25):

$$\Psi = \beta_1 + \beta_2 + \beta_3 + \beta_4 \quad (25)$$

Where:

β_1 : represents the delay for the sensor node to measure and prepare the data.
 β_2 : is the communication delay from the sensor (or intermediate gateway) to the cloud, which I further define as in Equation (26):

$$\beta_2 = \frac{S}{B} + T_p \quad (26)$$

where S is the size of the message in bits, B is the effective bandwidth (differing between Wi-Fi and Ethernet), and T_p captures additional propagation or processing overhead.

β_3 : accounts for the processing delay within the cloud (including any function execution and data storage)

β_4 : is the delay for the command to travel back from the cloud to the actuator.

For instance, when comparing Wi-Fi versus Ethernet, I have in Equations (27) and (28) respectively:

$$\beta_2^{\text{WiFi}} = \frac{S}{B_{\text{WiFi}}} + T_p^{\text{WiFi}} \quad (27)$$

$$\beta_2^{\text{Ethernet}} = \frac{S}{B_{\text{Ethernet}}} + T_p^{\text{Ethernet}} \quad (28)$$

Fig.s 8.5 and 8.6 in Appendix A6 illustrate the four segments of the end-to-end delay Ψ from Equation (29), showing how each phase ($\beta_1, \beta_2, \beta_3, \beta_4$) contributes to overall responsiveness. The red section corresponds to sensor node preparation (β_1), blue indicates communication between the Raspberry Pi and the cloud (β_2), green represents cloud processing (β_3), and yellow covers the return path from the cloud to the actuator (β_4). In the Wi-Fi scenario (see Appendix A6 - Fig. 8.5), the communication delay (β_2) is typically larger due to lower bandwidth and higher overhead, whereas in the Ethernet scenario (see Appendix A6 - Fig. 8.6), this segment is reduced, reflecting the faster, more stable wired connection. Furthermore, if retransmissions are required due to packet loss, the effective delay increases by a factor of $1/P_{\text{success}}$, where P_{success} is the probability of a successful transmission as in Equation (29):

$$\Psi_{\text{eff}} = \Psi \times \frac{1}{P_{\text{success}}} \quad (29)$$

Time Responsiveness Ψ in point-to-point vs. cloud in HTTP-based IoT system.

Fig. 3.3. shows a setup where the Arduino-based microcontroller reads physical sensor data with a delay β_1 . The microcontroller is connected to the router via Wi-Fi, and from there the connection is wired to the cloud application with a communication delay β_2 . β_3 processing time is used by the cloud application to evaluate the sensor data, and β_4 communication delay is used to return the control data to the microcontroller. The microcontroller controls the actuator control with a delay β_5 .

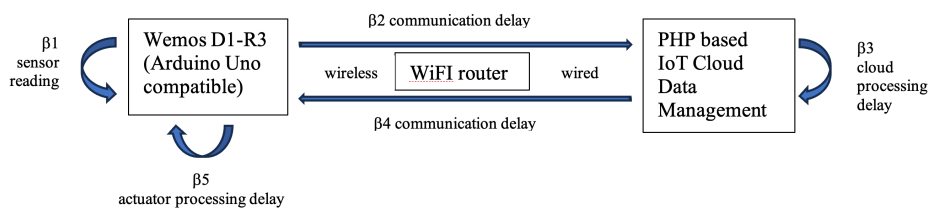


Fig. 3.3. Time distribution for Wi-Fi-based sensor to cloud communication using HTTP protocol.

MATERIALS AND METHODS

To compute the end-to-end responsiveness for Wi-Fi-based sensor to cloud communication using HTTP protocol, I model the overall delay, Ψ , as the sum of the delays in each segment of the communication path and shown in Equation (30):

$$\Psi = \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 \quad (30)$$

The measurement was performed within the microcontroller unit, where I measured the various latency times, while the execution time in the cloud was returned packaged in the response. The data arrived in a JSON structure in response to the HTTP request. Based on the received data, the PWM duty cycle of the output can be set, which, when connected to the input of the LED light source, determines the intensity of the grow light. The total communication delay consists of 5 elements.

Fig. 3.4. shows another setup where one Arduino-based microcontroller reads physical sensor data with a delay β_1 and then connects to the other microcontroller via Wi-Fi with a delay β_2 . The actuator control takes β_3 time. The communication is acknowledged with a delay β_4 .

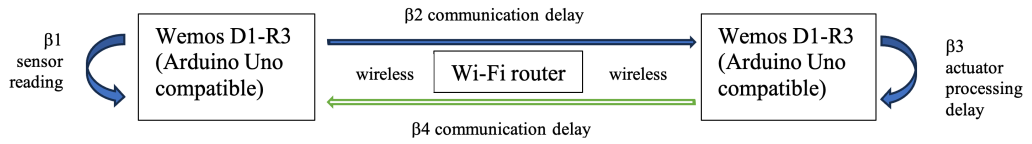


Fig. 3.4. Time distribution for Wifi-based sensor and actuator with point-to-point communication using HTTP protocol.

The series of measurements consisted of 500 queries. To compute the end-to-end responsiveness for Wi-Fi-based sensor point-to-point communication using HTTP protocol, I model the overall delay, Ψ , as the sum of the delays in each segment of the communication path shown in Equation (31):

$$\Psi = \beta_1 + (\beta_2 + \beta_4) / 2 + \beta_3 \quad (31)$$

The measurement was performed within the sensor microcontroller unit, where I were able to measure the different latency times. The latency time indicated by the green arrow is not included in the total processing time; it was only included for accurate measurement purposes. The actuator unit's latency time was received in the response. The response to the HTTP request returned the data in JSON format. The received data is evaluated by the actuator unit, which sets the PWM duty cycle for the output, determining the light intensity of the grow light connected to the LED light source input. The total communication latency consists of 3 elements.

In evaluating the time responsiveness of my MQTT-based IoT system, it's critical to contextualize β_2 (communication delay) within broader network characteristics. Ethernet offers low latency and high bandwidth but limited range and scalability, making it ideal for static, high-throughput scenarios. Wi-Fi provides moderate latency and bandwidth, with better mobility support but higher energy consumption and interference susceptibility.

3.3. Mathematical Modeling of Plant Growth

3.3.1. Multidimensional Logistic Growth state vector $X(t)$

To extend the traditional logistic growth model to accommodate multidimensional traits, I define a multidimensional logistic growth state vector $X(t)$ that incorporates the key plant traits as in Equation (32):

$$X(t) = \begin{bmatrix} H(t) \\ B(t) \\ C(t) \\ A(t) \end{bmatrix} \quad (32)$$

I redefine the carrying capacity K and growth rate r as vectors to reflect the maximum achievable values for each trait under given environmental conditions as represented in Equation (33) and Equation (34) respectively:

$$K = \begin{bmatrix} K_H \\ K_B \\ K_C \\ K_A \end{bmatrix} \quad (33)$$

$$r = \begin{bmatrix} r_H \\ r_B \\ r_C \\ r_A \end{bmatrix} \quad (34)$$

Therefore, the logistic growth equation becomes as in Equation (35):

$$\frac{dX(t)}{dt} = r \odot \left(1 - \frac{X}{K}\right) \odot X \quad (35)$$

where $\odot$ denotes element-wise multiplication, r is the vector of growth rates, K = the vector of carrying capacities, and division of vectors is performed element-wise.

where:

- $r = [r_H, r_B, r_C, r_A]^T$ is the vector of growth rates,
- $K = [K_H, K_B, K_C, K_A]^T$ is the vector of carrying capacities,
- $\odot$ denotes element-wise multiplication,
- The division $\frac{X}{K}$ is also performed element-wise.

Because measurements were collected at weekly intervals, Equation (35) was approximated using the forward Euler method with step size $\Delta t = 1$ week as illustrated in Equation (36):

$$X_{i+1} = X_i + r \odot X_i \odot \left(1 - \frac{X_i}{K}\right), i = 0, 1, \dots, 6, 7 \quad (36)$$

Here:

$$X_i = X(t_i)$$

denotes the predicted trait vector at week $i + 1$, with $i = 0$ corresponding to week 1.

For each Treatment Group, the experimental observations are organized in an 8×4 matrix as Equation (37).

$$D = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_7 \end{bmatrix} \quad (37)$$

where each row d_i contains the four measured traits at week $i + 1$.

The parameters are the growth rates and carrying capacities to estimate p in Equation (38):

$$p = \begin{bmatrix} r \\ K \end{bmatrix} = [r_H, r_B, r_C, r_A, K_H, K_B, K_C, K_A]^T \quad (38)$$

For a given parameter vector p , the model prediction $\hat{X}_i(p)$ is obtained by iterating Equation (36) starting from $X_0 = d_0$.

The residuals between model predictions and experimental observations are defined as Equation (39):

$$e_i(p) = \hat{X}_i(p) - d_i, i = 1, \dots, 7. \quad (39)$$

The optimal parameter vector is obtained by minimizing the sum of squared residuals in Equation (40):

$$F(p) = \sum_{i=1}^7 \sum_{j=1}^4 e_{i,j}^2(p) = \|e(p)\|_2^2. \quad (40)$$

The minimization problem is solved using the Levenberg–Marquardt nonlinear least-squares.

3.3.2. Proposed Interdependence Modeling in the 4D Model

In the extended four-dimensional logistic growth model, the plant traits biomass (B), leaf area (A), chlorophyll content (C), and plant height (H) were treated as interdependent variables that dynamically influence one another.

MATERIALS AND METHODS

Based on the general relationship introduced in Section 2.7, I hypothesized that $g(B(t))$ follows an allometric relationship of the form shown in Equation (41):

$$H(t) = \alpha_H \cdot B(t)^{\beta_H} \quad (41)$$

where α_H and β_H are species-specific parameters. I tested this hypothesis empirically using the collected dataset.

The parameter α_H represents a scaling constant, while β_H defines the degree of nonlinearity in the height–biomass relationship. This formulation enables the model to represent structural growth dynamics as a function of biomass within the proposed four-dimensional framework.

3.4. Machine Learning and Data Processing Analytics

3.4.1. Hybrid Model: Plant Growth and Resource Efficiency

The process begins as shown in Fig. 3.5 with the stochastic model, which takes the initial environmental inputs derived from the prototype such as light, water, temperature, humidity, and growing medium as its baseline. These inputs represent the planned or controlled operating conditions of the system. Using stochastic differential equations (SDEs), the model simulates the dynamic and uncertain behavior of these environmental factors by capturing both deterministic trends and random fluctuations.

The outputs of this step ($dL_t, dGM_t, dH_t, dF_t, dT_t, dWA_t$) represent these variations and provide realistic environmental scenarios. These outputs, together with the original inputs, are then used in the empirical modeling stage. The empirical model applies regression-based relationships to predict key plant growth metrics, including biomass, plant height, and growth rate, thereby quantifying the influence of environmental conditions on plant development.

These predicted outputs are subsequently used as inputs to the optimization model, which aims to determine the optimal combination of environmental parameters (e.g., light, water, temperature, and nutrients) that maximize or minimize a defined objective function related to plant growth performance. In this framework, the input variables include light, moisture, temperature, and nutrients (N, P, K), while the output variables correspond to biomass (g) and plant height (cm). Prior to modeling, the data are standardized using Z-score normalization to ensure comparability across variables.

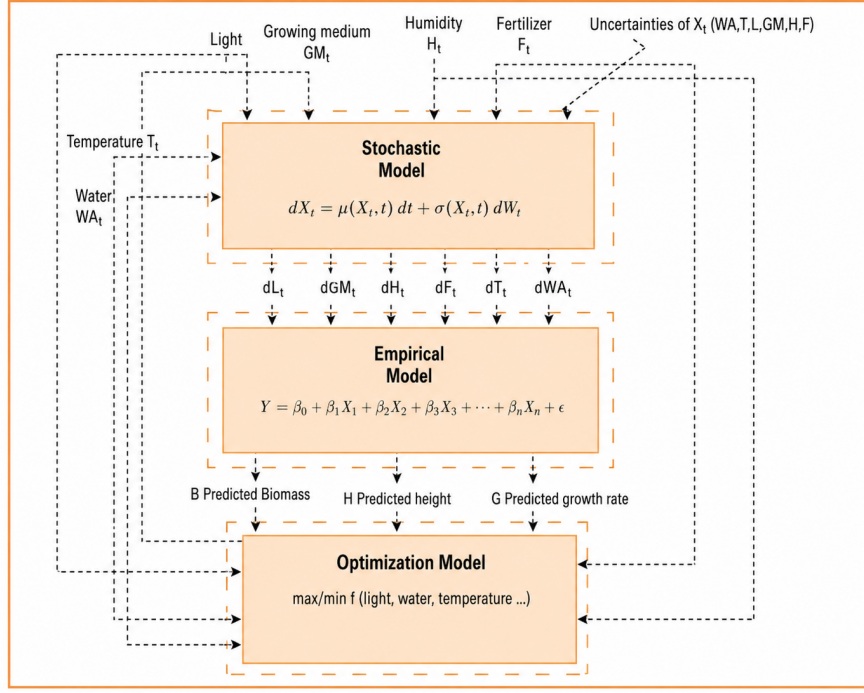


Fig. 3.5. Plant Growth Prediction and Resource Efficiency

3.4.2. Dataset Structure and Model Inputs

All machine learning models were trained on a 12-week dataset collected from IoT sensors, capturing key environmental parameters, including light intensity, moisture, temperature, nitrogen, phosphorus, and potassium, alongside corresponding plant growth metrics such as biomass (g) and height (cm). Each data point represents a weekly observation, where environmental variables serve as inputs, and growth metrics function as outputs. To ensure scale uniformity, the dataset underwent z-score normalization as a preprocessing step.

The input elements consisted of light (lumens), moisture (%), temperature ($^{\circ}\text{C}$), nitrogen (mg/kg), phosphorus (mg/kg), and potassium (mg/kg), while the output variables included crop yield (g) and plant height (cm).

3.4.3. Cloud-Based Data Processing

Apache NiFi was employed as the ingestion and data processing tool. The data flow, illustrated in Fig. 3.6, begins with the ListBlob and FetchBlob processors, which systematically receive sensor data from the storage system. Once the data is obtained, the InvokeHTTP processor transmits the entire JSON payload received from the blob directly to an API for further processing and evaluation.

MATERIALS AND METHODS

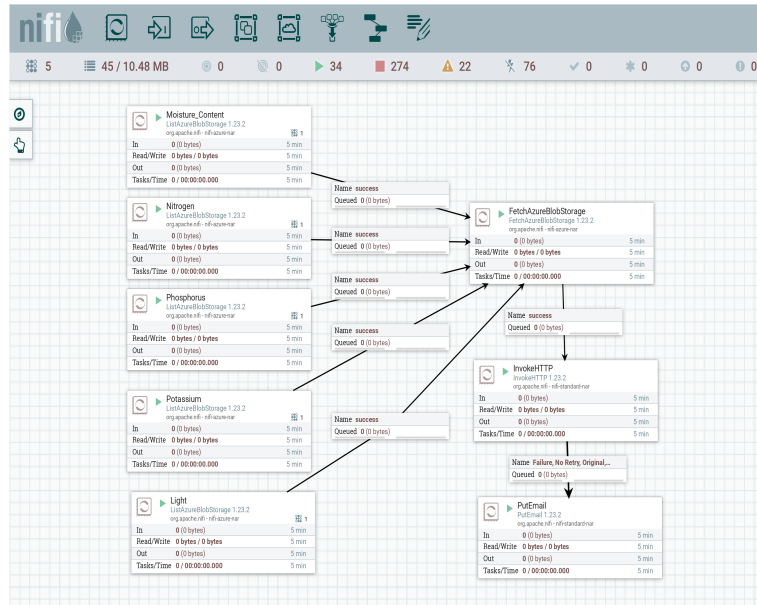


Fig. 3.6. Apache NiFi data flow for IoT sensor processing and API integration.

3.4.4. Data Collection

The prototype growing chamber in Fig. 3.7 is a controlled environment system designed to optimize and monitor plant growth conditions. It features a transparent enclosure from plexiglass to allow visual inspection. A cooling fan is installed to ensure air circulation, preventing stagnation and regulating temperature and humidity. The system includes adjustable LED lights mounted inside the chamber enabling precise control of light conditions to suit the needs of different growth stages.

Fig. 3.8 shows the electronic elements of the prototype. A Raspberry Pi serves as the central processing unit for my IoT data collection and control, while a relay module connected to an Arduino manages devices such as the fan, and LED lights. A breadboard facilitates prototyping and the integration of the combined sensor for temperature, humidity, and CO₂.

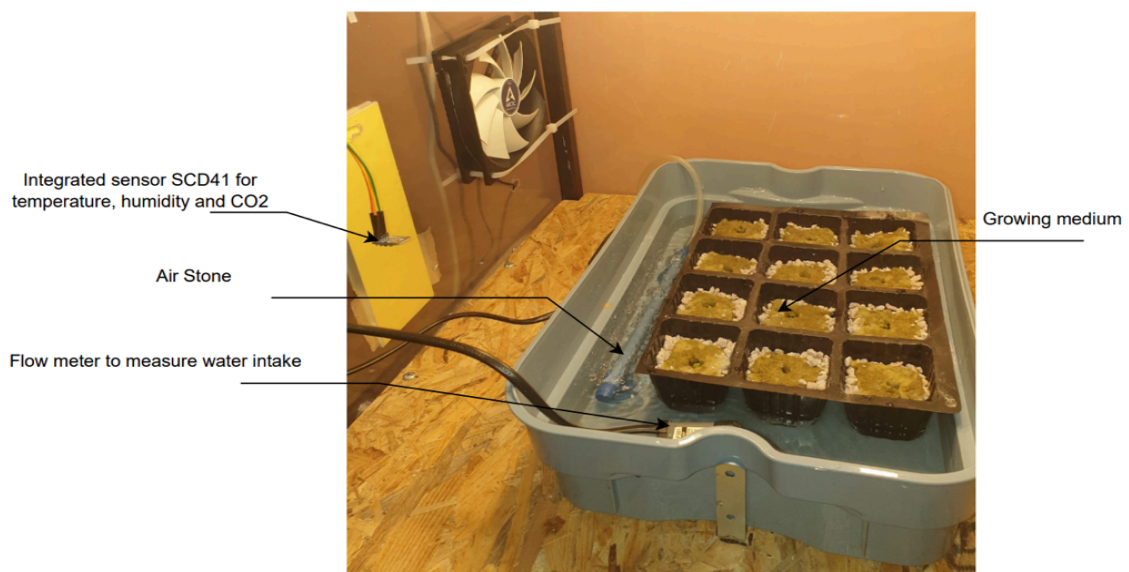


Fig. 3.7. IoT-Integrated Growing Chamber from inside

MATERIALS AND METHODS

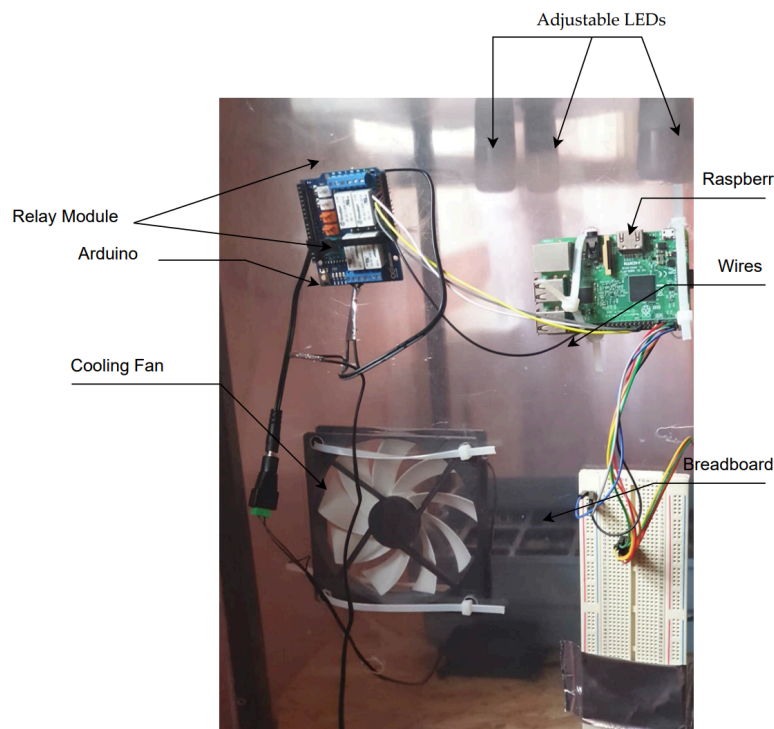


Fig. 3.8. Side view of the prototype showing the electronic part for monitoring.

Cooling Fan: Regulates the internal microclimate by providing air circulation and temperature stabilization within the chamber.

Raspberry Pi: Serves as the central processing unit responsible for data acquisition, system control, and communication with connected sensors.

Relay Module: Enables low-power control signals to switch higher-power devices such as the ventilation fan.

Arduino Microcontroller: Interfaces with environmental sensors and transmits measured data to the main processing unit.

Breadboard: Provides a temporary prototyping platform for assembling and testing electronic circuits without soldering.

Electrical Wiring: Ensures power distribution and signal transmission between system components.

Transparent Enclosure: Creates a controlled cultivation environment while allowing visual monitoring of plant development.

SCD41 Environmental Sensor: Measures carbon dioxide concentration, temperature, and relative humidity in real time.

Air Stone: Enhances dissolved oxygen levels in the nutrient solution to support root respiration.

Flow Meter: Monitors the volume of water supplied to the plants within the irrigation system.

MATERIALS AND METHODS

Growing Medium (Rockwool): Provides structural support for plant roots while facilitating water and nutrient retention. Initial data on environmental inputs, including light intensity, water intake, nutrient levels were collected directly from the prototype using IoT sensors and image analysis.

To validate early growth stages in Fig. 3.9 , lettuce (*Lactuca sativa* L.) seedlings were monitored inside the controlled chamber.

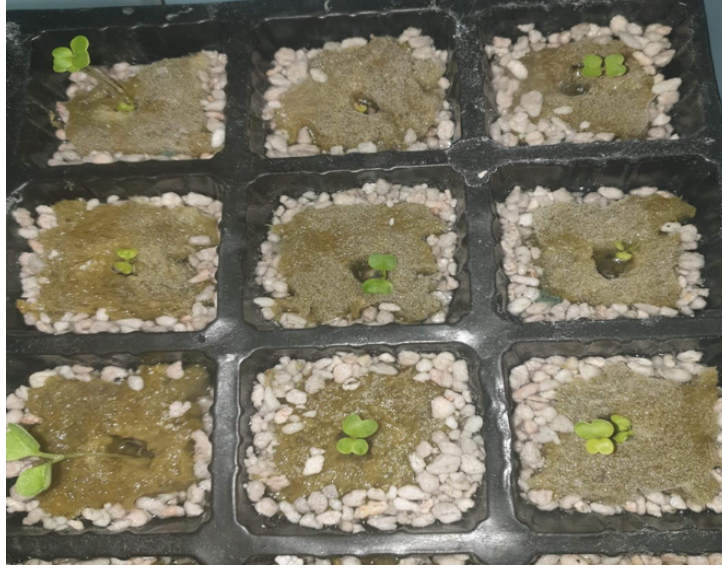


Fig. 3.9. Early growth stage of lettuce (*Lactuca sativa* L.) taken from the inside of the growth chamber

3.5. Novel Growth Performance Metrics

3.5.1. Integrated Agricultural Efficiency Metric (IAEM)

The Integrated Agricultural Efficiency Metric (IAEM), presented in Equation (42), introduces a novel metric for evaluating agricultural system performance by integrating real-time data within a unified model tailored for technology-driven farming.

IAEM extends beyond spatial insights by incorporating real-time resource utilization, alert accuracy, and system latency, ensuring a comprehensive assessment of system responsiveness and operational efficiency. IAEM serves as a practical operational performance metric, bridging the gap between technical system optimization and agricultural productivity.

By integrating key sub-metrics such as Resource Efficiency (RE), Real-Time Alert Accuracy (RA), Data Latency (DL), and Cloud Data Availability (CDA), IAEM effectively addresses gaps in existing agricultural performance metrics. This makes it a performing tool for real-time evaluation of data-driven agricultural systems, particularly those leveraging ETL solutions such as Apache NiFi to enable seamless data ingestion and analysis.

$$IAEM = \frac{\omega_1.RE + \omega_2.RA - \omega_3.DL}{CDA} \quad (42)$$

The Integrated Agricultural Efficiency Metric (IAEM) is computed using normalized sub-metrics, including Resource Efficiency (RE), Real-Time Alert Accuracy (RA), Data Latency (DL), and Cloud Data Availability (CDA). The lowest IAEM values occur when RE and RA are minimal,

MATERIALS AND METHODS

DL is at its highest, and CDA is low, indicating suboptimal system performance. Conversely, the highest IAEM values are achieved when RE and RA are maximized, DL is minimized, and CDA remains high, reflecting an efficient and well-optimized agricultural system. Table 3.8. presents the proposed scale for interpreting IAEM values.

Table 3.8. IAEM performance scale

IAEM Range	Performance Level
IAEM < 0.1	Poor Performance
$0.1 \leq \text{IAEM} < 0.2$	Moderate Performance
$0.2 \leq \text{IAEM} < 0.3$	Good Performance
IAEM ≥ 0.3	Excellent Performance

The IAEM framework is composed of several sub-metrics that collectively evaluate agricultural system performance:

a. Data Latency Metric (DL)

The metric part in Equation (43) represents the time delay between data generation and actionable insights or system response.

$$DL = \frac{\sum_{i=1}^n (T_{\text{response},i} - T_{\text{generation},i})}{n} \quad (43)$$

Where:

$T_{\text{response},i}$ is the time when the system generates a response for the i-th data point, $T_{\text{generation},i}$ is the time when the i-th data point is generated, n is the total number of data points.

b. Resource Efficiency (RE)

The metric in Equation (44) quantifies the efficiency of the system in processing data relative to the resources used:

$$RE = \frac{\sum \text{Resource Optimization}}{\text{Number of Resources}} \quad (44)$$

c. Real-Time Alert Accuracy (RA)

The metric in Equation (45) evaluates the accuracy of real-time alerts generated by the system.

$$RA = \frac{\text{Correct Alerts}}{\text{Total Alerts Generated}} \times 100 \quad (45)$$

d. Cloud Data Availability (CDA)

The metric in Equation (46) represents Cloud Data Availability, it shows the ability of a cloud storage system; Azure Blob Storage in my case, to ensure full access to data for read and write

MATERIALS AND METHODS

operations over a given period. This metric measures the reliability and accessibility of data in the cloud, expressed as a percentage of uptime or successful operations versus total operational time.

$$\text{CDA (\%)} = \text{SR(\%)} \times \text{Uptime (\%)} \quad (46)$$

where:

$$\text{SR} = \frac{\text{Successful Requests}}{\text{Total Requests}} \times 100$$

$$\text{Uptime (\%)} = \frac{\text{System Uptime}}{\text{Total Operational Time}} \times 100$$

e. Pairwise Comparison algorithm using AHP for weights

In this study, a Python-based computational framework (Python 3.11) was developed to determine the weights of the newly proposed Integrated Agricultural Efficiency Metric (IAEM). The script was specifically adapted for this research and employs the Analytic Hierarchy Process (AHP), a well-established multi-criteria decision-making methodology. The flowchart in Appendix A7 illustrates the step-by-step process, which includes the construction of a pairwise comparison matrix, normalization of the matrix, and calculation of the average normalized values to derive the relative importance of each sub-metric. To ensure the accuracy and validity of the computed weights, the algorithm incorporates consistency checks, mitigating potential biases in decision-making. For a detailed breakdown of the mathematical formulations and computational steps, refer to Appendix A7.

3.5.2. The Growth Efficiency ratio GER

The Growth Efficiency Ratio (GER), shown in Equation (47), is a novel metric developed during this research to evaluate the efficiency of resource utilization in plant growth, particularly in biomass production.

GER represents the relationship between biomass generation and the total normalized environmental resources supplied during cultivation. It therefore quantifies how effectively resources such as energy, water, and nutrients are converted into plant biomass. In this study, the proposed hybrid model aims to maximize GER by optimizing resource allocation to achieve high biomass production with reduced resource consumption.

The objective is to improve resource-use efficiency in controlled-environment agriculture by minimizing energy, water, and nutrient inputs while maintaining high biomass productivity. Such optimization is particularly important in indoor farming systems, where energy consumption associated with lighting and cooling represents a major operational cost.

$$\text{GER} = \frac{B^*(t)}{\text{CRU}(t)} = \frac{B^*(t)}{E^*(t) + W^*(t) + N^*(t)} \quad (47)$$

MATERIALS AND METHODS

The Combined Resource Use (CRU) represents the aggregated normalized resource input in dimensionless form and serves as the denominator of the GER metric, as defined in Equation (48).

$$CRU(t) = E^*(t) + W^*(t) + N^*(t) \quad (48)$$

Where:

$$B^*(t) = \frac{B(t)}{B_{\max}}, \quad W^*(t) = \frac{W(t)}{W_{\max}}, \quad N^*(t) = \frac{N(t)}{N_{\max}}, \quad E^*(t) = \frac{E(t)}{E_{\max}}$$

where $B(t)$ represents the total biomass yield corresponding to the final fresh weight of the plant at time t (g), $E(t)$ denotes energy consumption corresponding to electricity usage at time t (kWh), $W(t)$ represents the amount of water consumed during the plant growth cycle at time t (L), and $N(t)$ denotes the nutrient input supplied at time t (g). $B_{\max}$, $E_{\max}$, $W_{\max}$, and $N_{\max}$ represent the maximum biomass, energy consumption, water usage, and nutrient input values, respectively. The normalized variables B^* , E^* , W^* , and N^* correspond to the dimensionless normalized forms of biomass, energy, water, and nutrient inputs.

3.5.3. Plant Growth Index (PGI)

The Plant Growth Index (PGI) is a metric I developed specifically for this case study, integrates key growth factors such as height, biomass, and leaf area to provide a composite measure of plant health and resource efficiency. By employing a data-driven approach to determine the relative weights of each factor, this index represents a unique contribution to controlled environment agriculture, offering a scalable and adaptable tool for assessing plant performance under varying growth conditions.

The PGI in Equation (49) allows for tracking plant health over time and across multiple growth factors. The maximization of the PGI is ensured by the hybrid model to get optimal plant health and growth efficiency within the given environmental conditions. This composite metric makes the comparison of the different plant growth conditions easier.

$$PGI = w_1 \cdot \frac{H}{H_{\max}} + w_2 \cdot \frac{B}{B_{\max}} + w_3 \cdot \frac{A}{A_{\max}} \quad (49)$$

$$w_1 = \frac{H_{\text{norm}}}{H_{\text{norm}} + B_{\text{norm}} + A_{\text{norm}}}$$

$$w_2 = \frac{B_{\text{norm}}}{H_{\text{norm}} + B_{\text{norm}} + A_{\text{norm}}}$$

$$w_3 = \frac{A_{\text{norm}}}{H_{\text{norm}} + B_{\text{norm}} + A_{\text{norm}}}$$

Where

$$X_{\text{norm}} = \frac{\sum_{i=1}^n X_i}{X_{\max}}$$

where $X \in \{H, B, A\}$, i represents the day of measurement, and n denotes the total number of

MATERIALS AND METHODS

measured days. H represents plant height normalized by the maximum observed height, B represents plant biomass normalized by the maximum observed biomass, and A represents leaf area normalized by the maximum observed leaf area.

To compute the Plant Growth Index (PGI), I employed a machine learning approach, specifically linear regression, to determine the relative importance of key growth factors: plant height, biomass, and leaf area.

The logistic regression process, developed specifically for this case study, was implemented using a Python script shown in Appendix A8 and followed the process below:

a. Data Preparation:

The dataset was divided into training and testing subsets, with 80% of the data allocated for training the model and 20% reserved for testing. Input data (height, biomass, and leaf area) was structured into a feature matrix (X) and paired with the target variable (P_t), representing overall growth efficiency.

b. Model Training:

A linear regression model was applied to compute coefficients (w_1, w_2, w_3) for each growth factor. These coefficients quantified the contribution of each factor to the target metric (P_t).

c. Normalization:

The coefficients are normalized to sum to 1, ensuring interpretability as weights in the PGI formula. By training the model using Python, the coefficients were calculated to represent the relative influence of each factor on growth efficiency.

The logit transformation ensures that the probabilities are correctly modeled while maintaining a linear relationship between the factors and the target metric.

3.5.4. Relative Growth Rate (RGR) Metric

The Relative Growth Rate (RGR) was selected to represent the response variable because it normalizes biomass accumulation relative to plant size and allows comparisons across different developmental stages.

The Relative Growth Rate is defined in Equation (50):

$$RGR = \frac{B_{t+1} - B_t}{B_t} \quad (50)$$

where:

B_t is plant biomass at week t. [g], B_{t+1} is plant biomass at week t + 1. [g]

MATERIALS AND METHODS

Because both numerator and denominator are expressed in grams, the resulting quantity is dimensionless and represents a relative growth rate per week.

Using the centered variables, the growth response was approximated by the following second-order regression model in Equation (51):

$$\text{RGR} = \beta_0 + \beta_1 I_c + \beta_2 P_c + \beta_3 W_c + \beta_{11} I_c^2 + \beta_{22} P_c^2 + \beta_{33} W_c^2 + \beta_{12} I_c P_c + \beta_{13} I_c W_c + \beta_{23} P_c W_c \quad (51)$$

To improve numerical stability during regression estimation, the predictors were centered around their mean values:

$$\begin{aligned} I_c &= I - \bar{I} \\ P_c &= P - \bar{P} \\ W_c &= W - \bar{W} \end{aligned}$$

Where:

I represents light intensity ($\mu\text{mol m}^{-2} \text{s}^{-1}$), P represents photoperiod (hours day^{-1}), W represents cultivation week, and $\bar{I}$, $\bar{P}$, and $\bar{W}$ denote the mean values of light intensity, photoperiod, and cultivation week, respectively.

Therefore:

The regression coefficients therefore have the following units: β_0 is dimensionless, $\beta_1 (\mu\text{mol m}^{-2} \text{s}^{-1})^{-1}$, $\beta_2 (\text{h day}^{-1})^{-1}$, β_3 dimensionless, $\beta_{12} (\mu\text{mol m}^{-2} \text{s}^{-1} \cdot \text{h day}^{-1})^{-1}$, $\beta_{23} (\text{h day}^{-1})^{-1}$, $\beta_{22} (\text{h day}^{-1})^{-2}$, and β_{33} dimensionless.

To obtain a parsimonious model, a stepwise regression procedure was applied using MATLAB's `stepwiselm` function with the Akaike Information Criterion (AIC) as the selection criterion, the procedure.

The algorithm iteratively removed statistically insignificant predictors while monitoring changes in model information criteria. Terms whose removal improved the AIC score were excluded from the model.

As a result, the final model retained only the statistically relevant predictors shown in Equation (52):

$$\text{RGR} = \beta_0 + \beta_1 I_c + \beta_2 P_c + \beta_3 W_c + \beta_{12} I_c P_c + \beta_{23} P_c W_c + \beta_{22} P_c^2 + \beta_{33} W_c^2 \quad (52)$$

The terms I_c^2 and $I_c W_c$ were removed during stepwise selection because they did not significantly improve model performance.

The regression coefficients were estimated using ordinary least squares (OLS) implemented in MATLAB.

MATERIALS AND METHODS

The model can be written in matrix form as in Equation (53):

$$y = X\beta + \varepsilon \quad (53)$$

where:

y represents the vector of observed RGR values, X is the regression design matrix, β is the vector of regression coefficients, ε is the residual error vector.

The design matrix was constructed as Equation (54):

$$X = [1 \ I_C \ P_C \ W_C \ I_C \cdot P_C \ P_C \cdot W_C \ P_C.^2 \ W_C.^2] \quad (54)$$

The OLS estimator is shown in Equation (55):

$$\beta = (X^T X)^{-1} X^T y \quad (55)$$

Model performance was evaluated using the coefficient of determination in 2.9.3.

3.5.5. Light Use Efficiency (LUE) Metrics

Light Use Efficiency (Plant Canopy, LUEP):

- a. LUEP = PAR absorbed by leaves / PAR emitted by lamps
- b. LUEP = PAR emitted by lamps / PAR absorbed by leaves

Light Use Efficiency (Lamp Level, LUEL):

- a. LUEL = PAR received at canopy / PAR emitted by lamps
- b. LUEL = PAR emitted by lamps / PAR received at canopy

Indoor farming systems rely on efficient light utilization to maximize lettuce (*Lactuca sativa* L.) productivity while minimizing energy consumption. Metrics such as Light Use Efficiency at the Plant Canopy Level (LUEP) and Light Use Efficiency at the Lamp Level (LUEL) are essential to evaluate how well artificial lighting supports photosynthesis and biomass production. These metrics highlight the importance of balancing energy output, light absorption, and lettuce (*Lactuca sativa* L.) health.

This metric represents the efficiency with which plants convert photosynthetically active radiation (PAR) into biomass. It is influenced by canopy light interception, spectral distribution, and plant spacing. It can be computed according to Equation (56):

$$LUEP = \frac{\text{PAR Absorbed by Leaves}}{\text{PAR emitted by Lamps}} \quad (56)$$

LUEL: Light energy use efficiency with respect to PAR emitted from lamps.

This metric represents the efficiency of converting light emitted from lamps into usable energy for plants. It can be computed according to the Equation (57):

$$LUEL = \frac{\text{PAR Received at the Canopy}}{\text{PAR emitted by Lamps}} \quad (57)$$

3.5.6. System-Level Energy Efficiency

The system-level energy efficiency evaluates the electrical-to-optical conversion efficacy of the IoT-controlled LED system. It is computed as shown in Equation (58):

$$\text{Energy Efficiency (\%)} = \left(\frac{\text{PAR}_{\text{canopy}}}{\text{Electrical Energy Input}} \right) \cdot K \cdot 100 \quad (58)$$

where K is a conversion factor ($\mu\text{mol/J}$) derived from LED spectral data. Key implementation details include:

- Electrical Input: Measured in real-time via smart meters integrated with the Raspberry Pi control system.
- PAR Output: Quantified using calibrated quantum sensors at canopy height.
- Adaptive Optimization: The IoT framework dynamically adjusted LED intensity to maintain optimal efficiency while reducing power waste.
- This metric complements LUEP/LUEL by directly linking energy consumption to usable light for plant growth.

The calculation of Light Use Efficiency of Photosynthesis (LUEP) and Light Use Efficiency of Lighting (LUEL) for LED configurations is detailed in Appendix A9. The computational algorithm processes inputs such as the number of LEDs, light intensity, the percentage of photosynthetically active radiation (PAR) lost to the canopy, and PAR absorbed by the leaves. As previously stated, Apache NiFi is responsible for collecting environmental data and transmitting it to a dedicated API, which evaluates these values against predefined thresholds to assess plant environmental conditions.

3.5.7. Ressource Use Efficiency

To evaluate the effectiveness of the environmental control strategies, four key resource-use efficiencies (RUE) were calculated: Light Use Efficiency (LUE), CO₂ Use Efficiency (CUE), Nutrient Use Efficiency (NUE), and Photoperiod Use Efficiency (PUE).

The principle underlying these indicators is the relationship between plant productivity and the total environmental resources supplied during the cultivation period. In this study, resource-use efficiency was defined as the ratio of the final plant biomass to the cumulative environmental resource input over the entire experimental cycle.

The general expression used to calculate resource-use efficiency is in Equation (59):

$$RUE = \frac{B_{\text{final}}}{\sum_{t=1}^n R(t)} \quad (59)$$

where:

B_{final} represents the final plant biomass (g) measured at the end of the 8-week.

MATERIALS AND METHODS

$R(t)$ denotes the value of the environmental resource supplied during week t . Depending on the efficiency being calculated, this variable corresponds to:

Light intensity ($\mu\text{mol m}^{-2} \text{s}^{-1}$) for LUE, CO_2 concentration (ppm) for CUE, Nutrient concentration (ppm or EC equivalent) for NUE, Photoperiod duration (hours per day) for PUE.

n is the total number of weeks in the experiment, where $n = 8$.

The denominator $\sum_{t=1}^8 R(t)$ represents the cumulative amount of the given resource supplied during the entire experimental period. This cumulative value was used to normalize the final biomass production and quantify how efficiently each environmental resource contributed to plant growth.

I applied the calculation procedure independently for each environmental factor using the weekly measurements recorded during the experiment.. The resulting efficiencies were then compared between the control treatment (Treatment Group A) and the IoT-enabled dynamic control treatment (Treatment Group B) to assess the impact of the intelligent environmental management system on resource utilization.

3.6. Image Analysis

The workflow for image analysis to evaluate the impact of different lighting conditions on plant features is shown in Appendix A10. Images of plants and illustrated in Fig. 3.10 (week 5) and Fig. 3.11 (week 8) were captured under controlled lighting setups, followed by preprocessing to standardize dimensions and reduce noise. Using segmentation algorithms, key metrics such as leaf count, plant area, and HSV (hue, saturation, value) characteristics were extracted. Binary masks were generated to isolate plant features and enhance visualization of edges and structural details under varying light intensities. These analyses provided a quantitative basis to assess the effects of lighting on plant growth and health, forming the foundation for the results presented.

All processing was performed using PlantCV and OpenCV. Appendix A11 Fig. 8.7, Fig. 8.8, and Fig. 8.9 present additional segmentation results for other growth stages (week 2, week 3, and week 7). Wet pixel analysis was used to estimate water absorption and nutrient uptake. A summary of parameters is provided in Appendix A12.

This method builds upon the approach discussed in previous work on plant growth monitoring through phenotyping and image analysis, where similar techniques were employed to study photosynthesis efficiency and resource utilization (Poorter et al., 2011). Appendix A12 summarizes the key parameters used for image processing.

MATERIALS AND METHODS

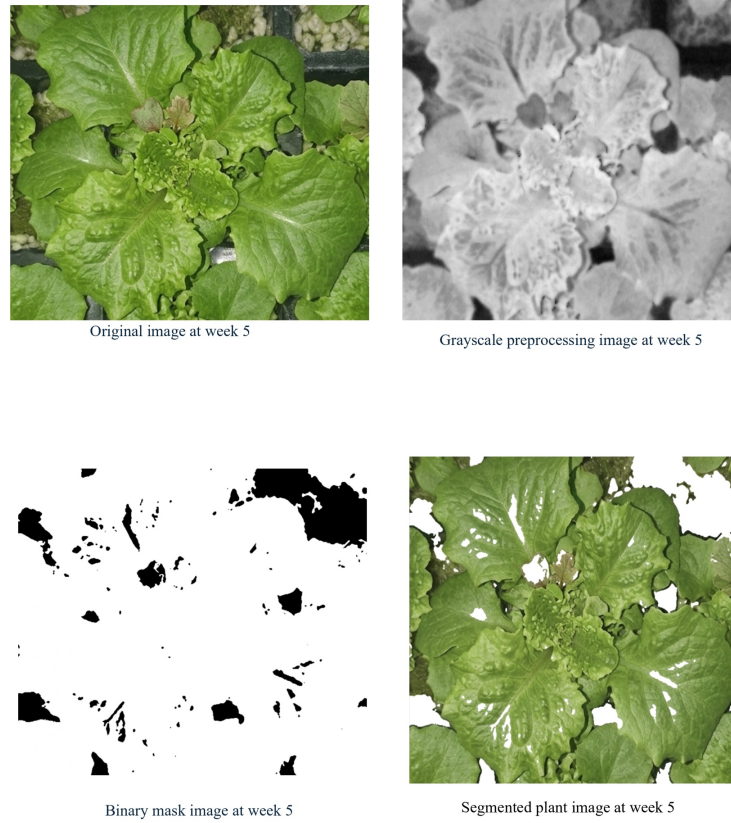


Fig. 3.10. Plant image segmentation workflow at week 5.

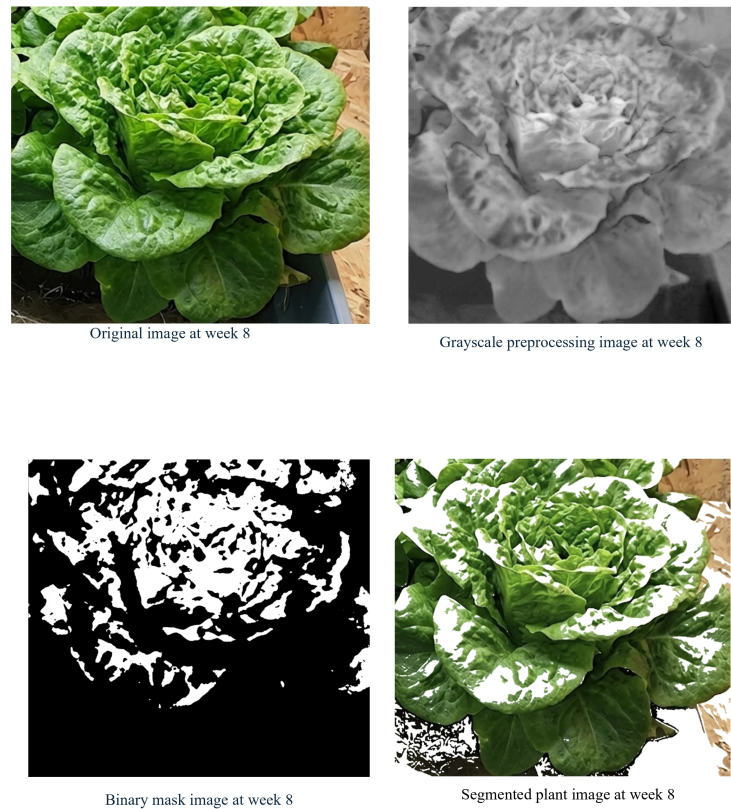


Fig. 3.11. Plant image segmentation workflow at week 8.

3.7. Feedback Control and Simulation

To further enhance the IoT-based environmental monitoring system, MATLAB Simulink (MATLAB R2024a) was utilized to model the dynamic growth of plants based on real-time sensor data. The Simulink model (shown in Appendix A13) solves the model in 2.5.5 and implements a feedback control system designed for the IoT-based plant growth monitoring framework with the following components.

The model incorporates several key components:

The Gain Block ($-k$) applies a constant adjustment factor to modulate the growth rate based on incoming sensor data, utilizing a negative feedback mechanism.

The Integrator Block ($1/s$) computes the total plant size by integrating the growth rate over time.

The Multiplication Block ($\times$) updates plant size dynamically by multiplying the growth rate by the current size.

The Feedback Loop continuously adjusts plant size predictions based on real-time environmental inputs, ensuring adaptive growth modeling.

4. RESULTS AND DISCUSSION

4.1. Growth Efficiency Ratio GER

4.1.1. Combined Resource Use vs GER

The plot in Fig. 4.1 shows X-axis representing energy consumption (kWh used per plant or system) and Y-axis representing the Growth Efficiency Ratio (GER), which was calculated as the biomass yield per unit of energy consumed, the data used to plot the graph and to do the computations of GER values are included in Appendix A14.

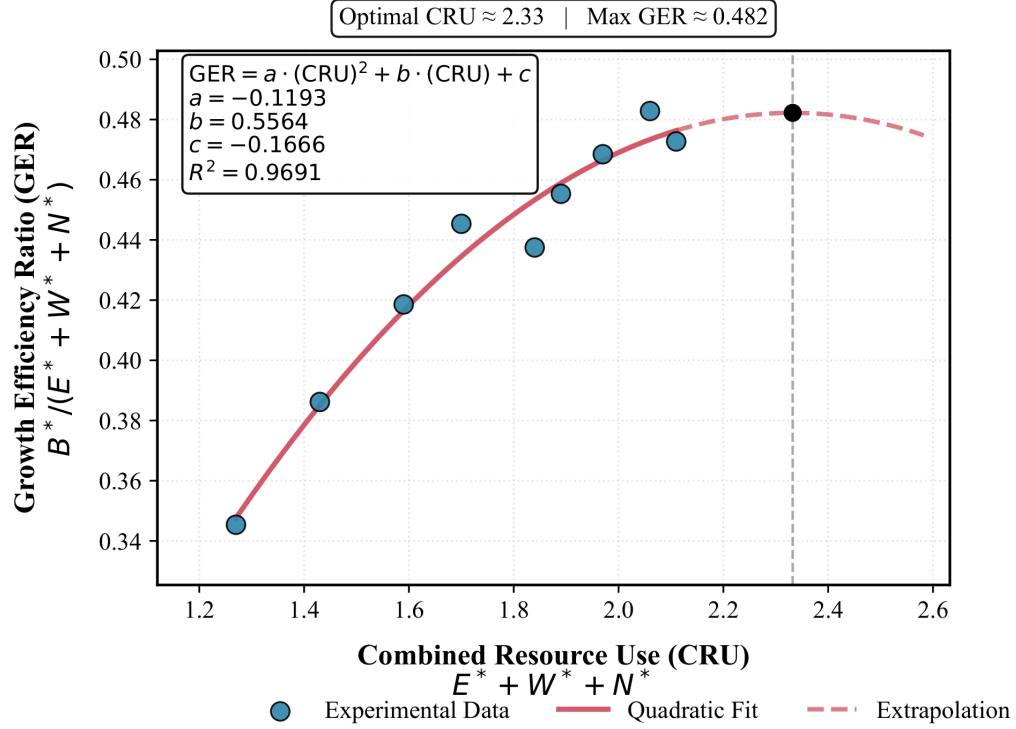


Fig. 4.1. Growth Efficiency Ratio vs Combined Resource Use Quadratic Regression Analysis

As illustrated in the plot in Fig.4.1, GER increases with increasing CRU in the lower range (approximately CRU = 1.2–1.6), indicating that initial increases in resource input lead to improved efficiency in biomass production.

Beyond this range, the rate of increase in GER progressively decreases, and the curve begins to flatten as CRU approaches values above approximately 1.8–2.0. This behavior reflects a diminishing-returns effect commonly observed in controlled-environment agriculture systems, where physiological and resource-use limitations reduce the efficiency of biomass production at higher input levels.

The plot demonstrates a non-linear relationship, which was mathematically modeled using a polynomial equation, as shown in Equation (60). This equation was derived based on the observed quadratic behavior of GER as a function of combined resource input:

$$\text{GER} = a \cdot (\text{Combined Resource Use})^2 + b \cdot (\text{Combined Resource Use}) + c \quad (60)$$

Combined Resource Use is the sum of normalised energy consumption (kWh), water usage (liters),

RESULTS AND DISCUSSION

and nutrient input (grams) as defined previously in 3.5.2. a , b , and c are coefficients determined by fitting the equation to the data.

In my case study, the regression analysis yielded the following coefficients:

$$a=-0.1193, b=0.5564, c=-0.1666$$

The negative quadratic coefficient ($a < 0$) indicates that the fitted curve is concave, confirming the presence of diminishing efficiency gains at higher resource-input levels. Furthermore, the model achieved a high coefficient of determination ($R^2 = 0.9691$), demonstrating excellent agreement between the regression model and the experimental observations.

To identify the operational point corresponding to maximum GER, the vertex of the quadratic function was determined. Since the parabola opens downward, the vertex corresponds to the optimal CRU value and is calculated using Equation (61):

$$CRU_{\max} = -\frac{b}{2a} \quad (61)$$

This gives the level of combined resource use that maximizes GER. I substitute this value back into the equation to get the maximum GER.

I can find the level of resource input where adding more inputs doesn't lead to significant efficiency gains, I analyze the rate of change of GER, by finding the second derivative of the GER equation shown in Equation (62):

$$\frac{d^2(GER)}{d(CRU)^2} = 2a \quad (62)$$

Because $2a < 0$, the model confirms a continuously decreasing marginal efficiency response as CRU increases. Overall, the developed quadratic model provides a robust quantitative framework for characterizing growth-efficiency behavior and identifying efficient operational zones in indoor farming systems. In this study, CRU values above approximately 2.0 correspond to a near-optimal region where further increases in resource consumption generate progressively smaller efficiency gains.

4.1.2. GER Heatmap

To further examine how individual resources influence GER, I analyzed the interaction between light intensity and water usage using additional visualizations. Specifically, I investigated how different combinations of light intensity and water input affect plant growth efficiency. To represent these interactions, a heatmap was generated illustrating GER variations across multiple combinations of light and water levels.

The heatmap displays GER values corresponding to different combinations of normalized light intensity and water usage. The data used to generate this visualization are provided in Appendix

RESULTS AND DISCUSSION

A14. Each cell in Fig. 4.2 represents a specific experimental condition, while the color gradient reflects the corresponding GER value. The heatmap highlights regions where growth efficiency is maximized and demonstrates the influence of resource balance on biomass production efficiency.

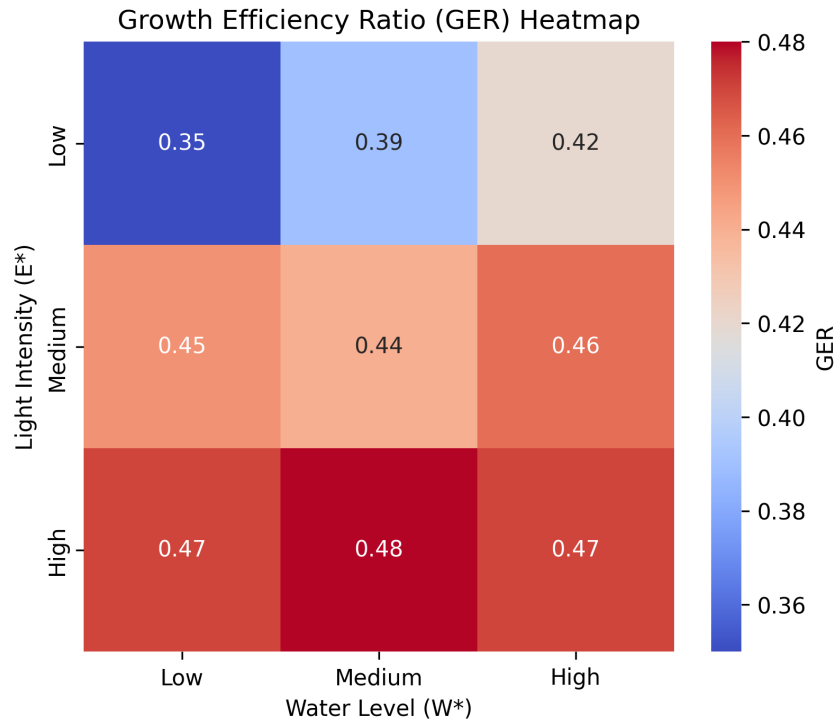


Fig. 4.2. GER Heatmap: Light vs Water.

Under low light conditions, GER values range from 0.35 to 0.42 and generally increase with higher water availability. This suggests that increasing water input partially improves growth efficiency under limited energy conditions, although the overall GER values remain relatively moderate.

Under medium light conditions, GER values vary between 0.44 and 0.46. The highest value within this range is observed under high water conditions, indicating that moderate increases in both light intensity and water availability positively influence growth efficiency.

The highest GER values are observed under high light conditions, ranging from 0.47 to 0.48. The maximum GER value (0.48) occurs under high light intensity combined with medium water input. A slight decrease is observed at the highest water level, suggesting diminishing returns when water input exceeds the optimal balance relative to light availability.

Although the maximum GER value is obtained under high light conditions, this operating point also requires substantially greater energy input. From a sustainability and resource-management perspective, lower-light conditions with moderate GER values may represent a more balanced operating strategy by reducing energy consumption while maintaining acceptable biomass productivity. For example, low light intensity combined with high water usage achieves a GER value of approximately 0.42 with lower energy demand compared to the maximum GER configuration.

RESULTS AND DISCUSSION

Overall, the heatmap demonstrates that both insufficient and excessive resource inputs may reduce growth efficiency, emphasizing the importance of balanced environmental management in controlled-environment agriculture. The results further indicate that maximizing biomass production alone does not necessarily correspond to maximum resource-use efficiency, highlighting the need for optimized resource allocation to support sustainable indoor crop production.

4.2. Plant Growth Traits Characteristics

4.2.1. Biomass yield

Biomass yield reflects the actual growth and productivity of the plants, which is crucial for maximizing output while growing plants indoor. However, producing biomass efficiently measured by the Growth Efficiency Ratio (GER) ensures that resources like water, light, and nutrients are not wasted. Both biomass yield and GER together can give the optimal conditions that provide the highest yield while minimizing resource consumption.

There three important factors that I considered in my study that decide about the growth rate are: Light, Water and Nutrient intake. Therefore, I plotted graphs as shown in Fig. 4.3 combining two parameters at the same time, a bar chart for Biomass yield and a line plot for GER to compare two different parameters from Light, Water and Nutrients on the same graph. The data used for this plot is shown in Appendix A15.

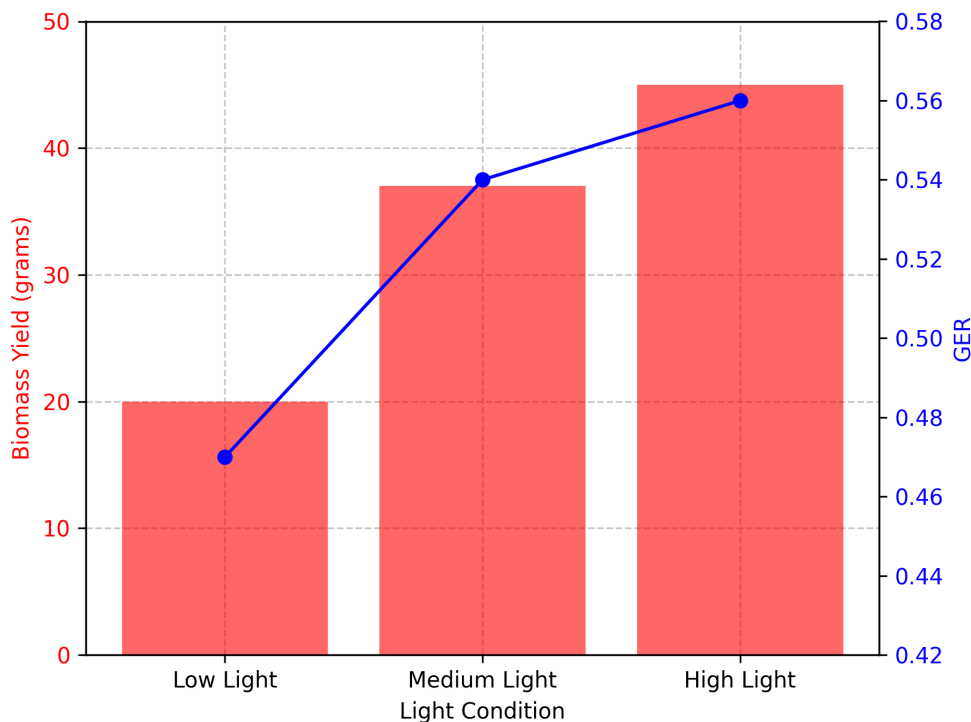


Fig. 4.3. Biomass Yield and GER Comparison Across Conditions.

The bars represent biomass yield in grams under different light conditions: low, medium, and high light conditions. The line plot (GER) indicates the efficiency of biomass production relative to energy consumption or resource use. High light conditions yield the most biomass (around 42 grams) and the highest GER (approximately 0.56). However, achieving higher GER with low light conditions is beneficial in my case study as resource efficiency is prioritized over maximizing biomass yield.

RESULTS AND DISCUSSION

The graph in Fig. 4.4 shows biomass yield and GER at different daily water intake rates (5L, 7L, and 9L per head). The data used for this plot is shown in Appendix A16. Increased water intake correlates with higher biomass yield, reaching about 210 grams at 9L. GER also improves with increased water usage, reaching about 0.52 at 9L. This suggests that, within these conditions, higher water intake optimizes both growth and resource efficiency.

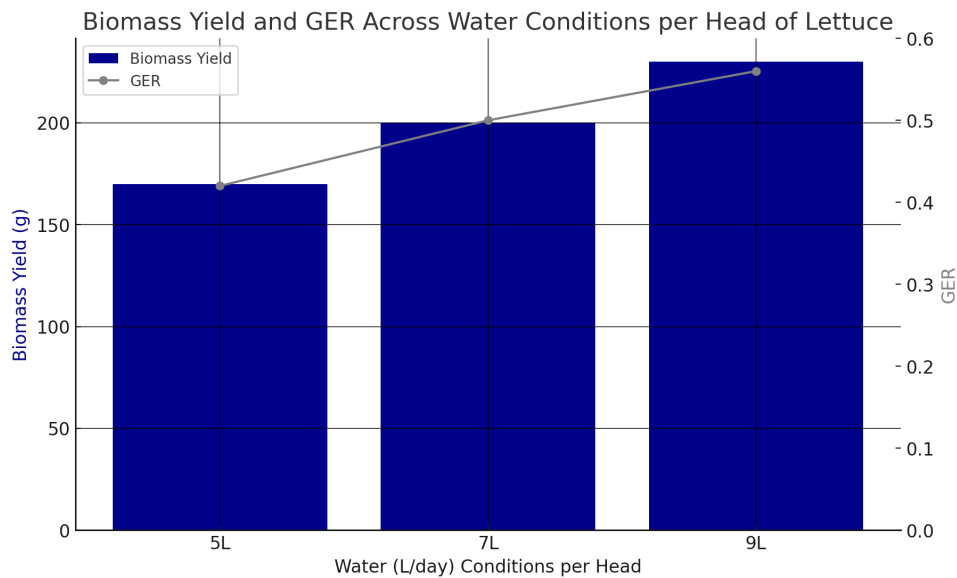


Fig. 4.4. Biomass Yield and GER Comparison Across Water Conditions per Head of Lettuce (*Lactuca sativa* L.).

In Fig. 4.5 biomass yield and GER are compared across different daily nutrient intake levels (3g, 4g, and 5g per head). The data used for this plot is shown in Appendix A17.

Biomass yield peaks at 5g (close to 200 grams), while GER also improves, reaching about 0.51. This indicates that increased nutrient input improves both growth and efficiency, but it's essential to weigh the costs of nutrients against the efficiency gains.

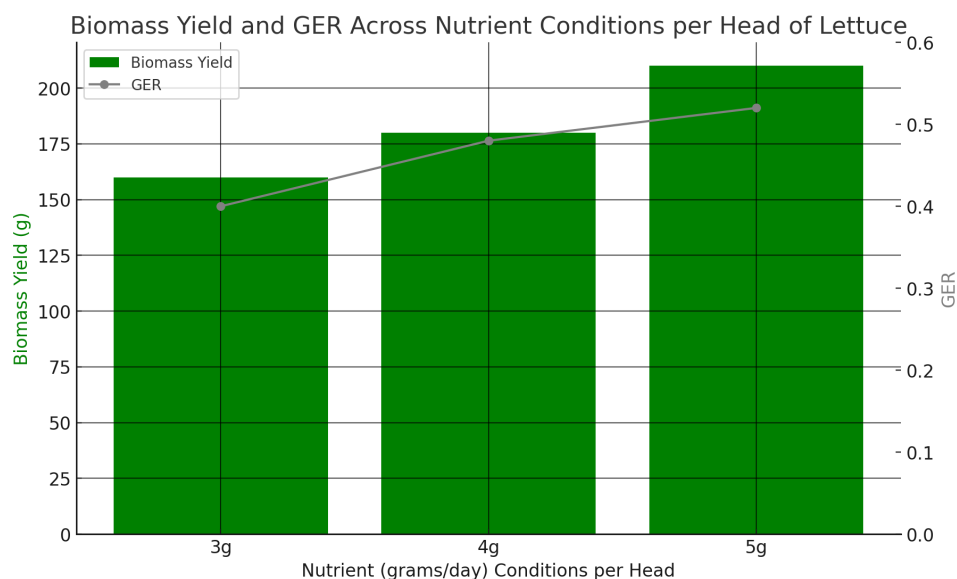


Fig. 4.5. Biomass Yield and GER Across Nutrient Conditions per Head of Lettuce (*Lactuca sativa* L.).

RESULTS AND DISCUSSION

To sum up, from the Biomass Yield and GER across Water Conditions plot, 9L per day appears to yield the highest biomass while still maintaining a high GER. Therefore, 9L/day seems to be an optimal water level. The Biomass Yield and GER across Nutrient Conditions plot suggests that 5 grams per day of nutrients provides the highest biomass yield and the best GER. Thus, 5g/day is optimal for nutrient input. Based on the Biomass Yield and GER Comparison Across Light Conditions and the GER Heatmap, high light levels (around 12–14 hours of light) seem to support the best growth efficiency and biomass yield. However, medium light conditions (10–12 hours) also show efficient GER levels and might be more sustainable for energy use.

4.2.2. Plant height

Plant height is closely related to biomass yield, as both metrics reflect the plant's overall growth and health.

The graph in Fig. 4.6 shows specific quantitative relationships between water and nutrient intake and plant height growth.

The data used are shown in Appendix A18. The slope of each regression line represents the response of plant height to increasing nutrient levels at each watering rate.

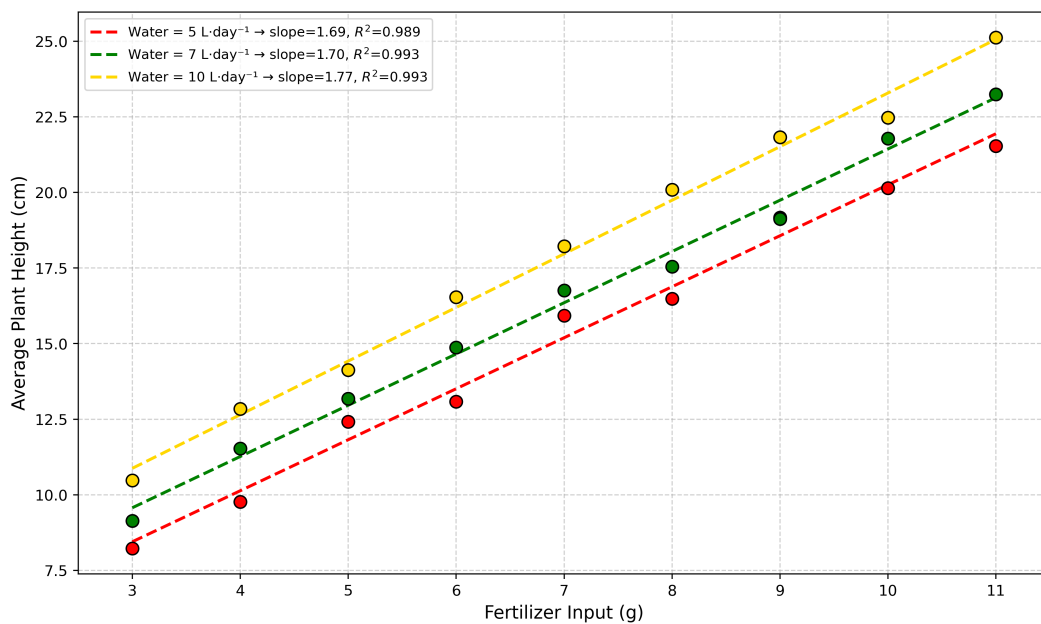


Fig. 4.6. Fertilizer Input vs Plant Growth with Watering Rates

At 10 L/day watering, the slope is higher (1.77 cm per gram of fertilizer), meaning that at this watering rate, plants gain more height per additional gram of nutrient input compared to the two lower watering rates.

At 5 L/day watering, the slope is lower (1.69 cm per gram), indicating a smaller increase in height for each additional gram of nutrients.

Within the tested experimental range, a combination of high watering (10 L/day) and fertilizer levels between 9 and 11 g produced the highest plant heights, reaching approximately 25 cm at 11

RESULTS AND DISCUSSION

g of fertilizer.

The data used are presented in Appendix A19. A positive correlation between light exposure and plant height is observed across all watering conditions. As shown in Fig. 4.7, plant height increases consistently with increasing lighting duration.

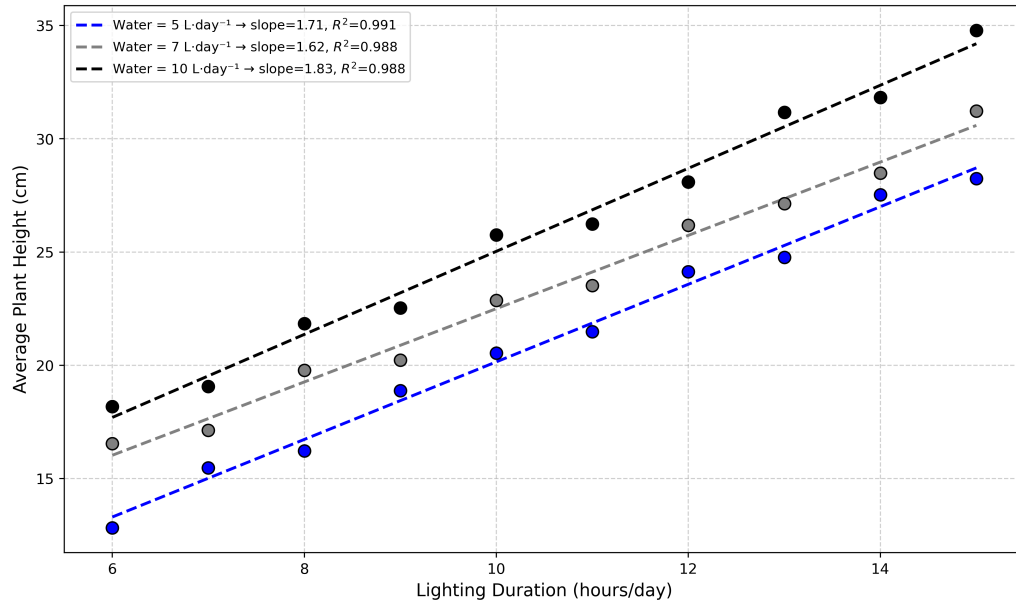


Fig. 4.7. Light Consumption vs Plant Growth with Watering Rates

Three watering levels were analyzed (5, 7, and 10 L/day), each showing a linear relationship between light exposure and plant height. The slope of this relationship increases with water availability.

At 10 L/day watering, the slope is the highest (1.83 cm per additional hour of light), indicating that plant height responds most strongly to increased light under optimal water supply. At 7 L/day, the slope is slightly lower (1.62 cm per hour), while at 5 L/day, the slope is 1.71 cm per hour, reflecting a reduced growth response under limited water conditions.

For optimal height growth, the combination of 14–15 hours of light per day with 10 L/day watering provides the most efficient conditions.

4.2.3. Leaf Area

As a part of analyzing the factors impacting the PGI index, I focused on how different environmental factors including water levels, nutrient intake, and light intensity affect the leaf area of plants over a 30-day growth period. Leaf area is a crucial indicator of a plant's photosynthetic capacity and overall health. Fig. 4.8 shows the change in leaf Area across different water levels, I can see that as water levels increase from 5L to 9L per day, there is a clear upward trend in leaf area growth, with 9L/day yielding the highest leaf area by day 30. The data used to plot this graph are in Appendix A20.

RESULTS AND DISCUSSION

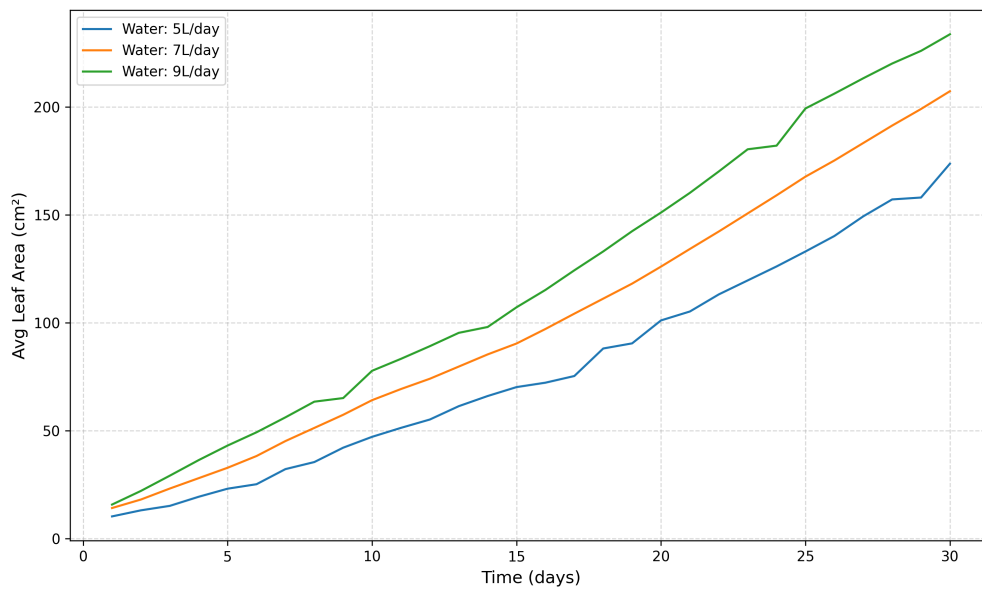


Fig. 4.8 Avg Leaf Area vs Water Levels.

Similar to water levels, increased nutrient intake in Fig. 4.9 shows that a change from 3g to 7g per day leads to a larger leaf area. Higher nutrient intake correlates with more significant leaf growth, reaching the maximum around 7g/day. The data used to plot this graph are in Appendix A21.

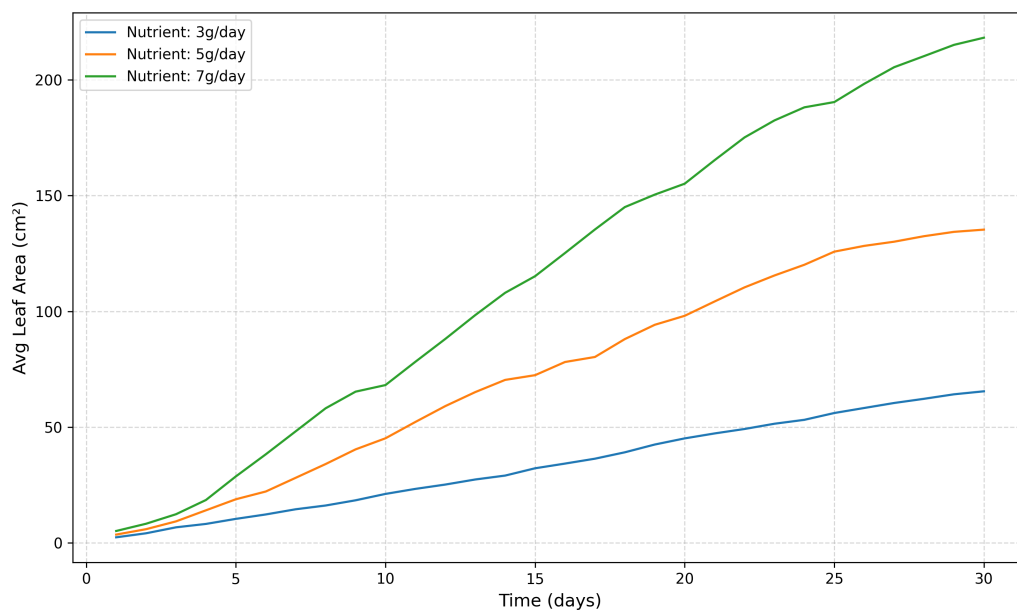


Fig. 4.9 Leaf Area vs Nutrient Intake

The impact of light exposure shown in Fig. 4.10 is also evident, with higher light intensity (12 hours/day) resulting in the most substantial leaf area growth, followed by moderate and low light conditions. The data used to plot this graph are in Appendix A22.

RESULTS AND DISCUSSION

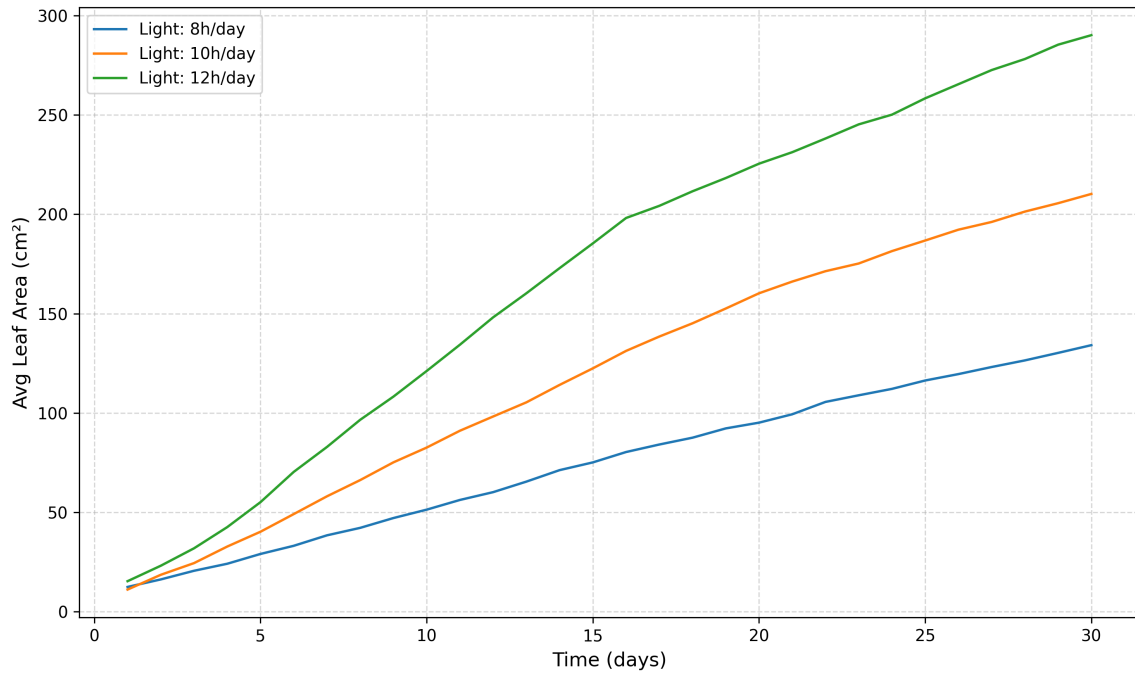


Fig. 4.10. Leaf Area vs Light Intensity.

4.3. Machine Learning for Plant Growth Index (PGI)

To determine the contributions of plant height, biomass, and leaf area to the Plant Growth Index (PGI), I implemented a linear regression model as described in the methodology.

The maximum values used for normalization correspond to theoretical upper bounds representing the potential growth limits under ideal conditions. Historical sensor measurements were then used to infer plant states at earlier growth stages, describing the progression toward these predefined maxima. This approach represents a bounded, data-driven reconstruction of growth trajectories, where empirical sensor data anchor the temporal evolution while the theoretical maxima define the asymptotic limits.

The input data is collected iteratively for each data point until the user indicates completion by typing "done." The python script is detailed in Appendix A8. The results, shown in Appendix A23, illustrate the input process and the normalized weights.

The normalized weights computation is shown in Appendix A24, the results derived from the linear regression model are as follows:

The highest weight ($W_3 = 0.3401$) indicates that Leaf Area has the most significant impact on the PGI. This suggests that Leaf Area accumulation is the primary driver of growth efficiency in the studied Lettuce (*Lactuca sativa* L.).

Plant height ($W_1 = 0.3335$) contributes moderately.

Biomass ($W_2 = 0.3264$) has a lower contribution.

RESULTS AND DISCUSSION

These results emphasize the need to prioritize leaf area accumulation strategies, , while monitoring height and biomass to avoid resource inefficiencies. Fig. 4.11 illustrates the progression of the Plant Growth Index (PGI) over a 30-day period, using data presented in Appendix A25, and the computed PGI values are shown in Appendix A26.

The PGI exhibits a gradual initial increase, starting at approximately 0.07 on Day 0 and reaching around 0.30–0.35 by Day 10, indicating steady early-stage growth as resources are primarily allocated toward initial plant establishment and structural development.

Between Days 10 and 20, the PGI increases more rapidly, rising from approximately 0.35 to 0.80. This phase corresponds to the main growth period, during which the plant efficiently accumulates height, biomass, and leaf area under favorable environmental conditions. After Day 20, the rate of increase slows, and the PGI asymptotically approaches the target value of 1.0 by Day 30, without fully reaching it. This saturation behavior indicates that the plant is approaching its theoretical growth potential, as defined by the PGI normalization framework and experimental constraints.

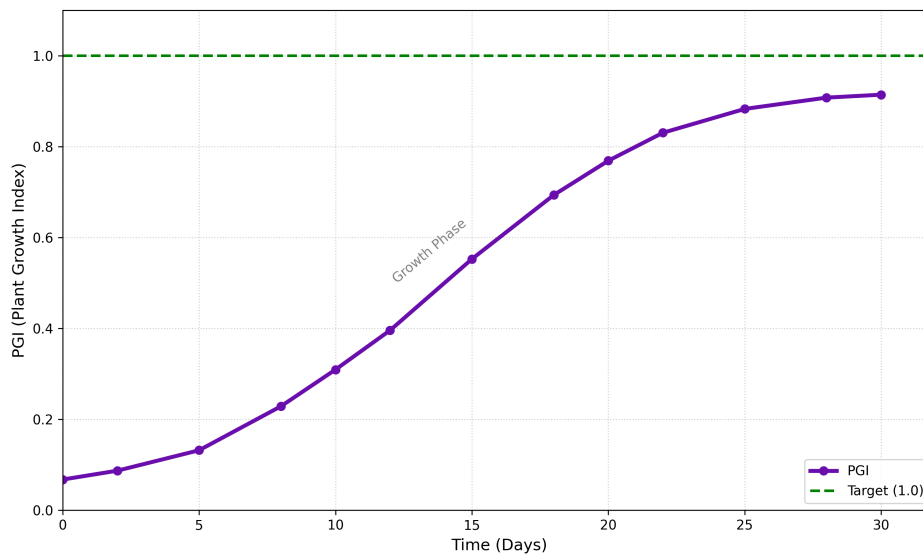


Fig. 4.11. PGI (Plant Growth Index) over 30 Days.

4.4. IoT and Real-Time Monitoring

4.4.1. Multivariate Regression Analysis of Growth Drivers

To complement the correlation analysis and account for simultaneous interactions among environmental variables, a multivariate linear regression model was applied using observed growth as the dependent variable. This analysis was performed on the dataset provided in Appendix A27 and aims to quantify the marginal contribution of each factor while controlling for the influence of others.

The estimated regression coefficients are reported in Appendix A28, which summarizes the standardized β -values for each environmental parameter. The results identify phosphorus as the strongest positive contributor to plant growth ($\beta = 0.54$), reinforcing its dominant role in biomass accumulation. Potassium exhibits a pronounced negative coefficient ($\beta = -0.43$), indicating potential overuse or inhibitory effects beyond optimal levels. Temperature contributes moderately ($\beta = 0.23$), while moisture ($\beta = 0.06$), nitrogen ($\beta = 0.04$), and light intensity ($\beta = 0.03$) show comparatively weak marginal effects once other variables are controlled.

4.4.2. Predictive Models Performance

The integration of machine learning models with the IoT-driven pipeline demonstrated substantial potential for optimizing indoor farming. Fig. 4.12 presents a detailed performance breakdown of each individual model, showcasing how each predictive algorithm forecasts plant growth. The graphs illustrate how closely each model aligns with the experimental growth trends. The raw data are shown in Appendix A29.

RESULTS AND DISCUSSION

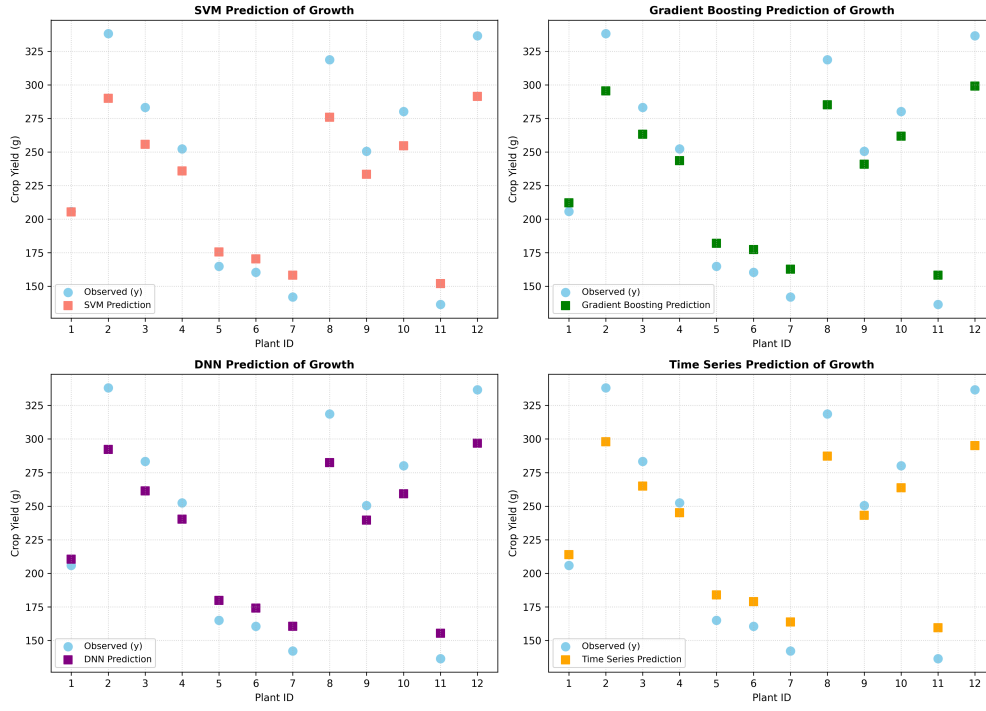


Fig. 4.12. Comparison of observed and predicted plant growth for 12 studied plants.

4.4.3. Resource Optimization with IoT-Based Monitoring

In this experiment, two cultivation strategies were evaluated: a static cultivation protocol (Treatment Group A, Non-IoT) and an adaptive IoT-based control protocol (Treatment Group B). The baseline (Treatment Group A) was defined as a non-IoT static protocol in which environmental inputs were maintained at fixed setpoints throughout the experiment, without any sensor-driven feedback or adaptive control. Light intensity, irrigation rates, and fertilization schedules were predefined and kept constant, with no adjustment based on real-time plant or environmental responses, reflecting a conventional cultivation approach. These setpoints were selected based on literature-reported optimal conditions for *Lactuca sativa* L. growth, ensuring a representative controlled-environment baseline.

A detailed description of the experimental setup and treatment protocols is provided in Section Growth under three resource management strategies (Table 3.6).

In contrast, the IoT-based protocol allowed these environmental parameters to be adjusted dynamically during plant development in order to better match the physiological needs of the crop at different growth stages.

Fig. 4.13 compares the evolution of three key plant growth indicators under the two Groups' strategies over the eight-week experimental period: plant height (cm), fresh biomass (g), and leaf area (cm²).

Across all three growth traits, the adaptive control strategy consistently resulted in improved plant development. In the case of plant height, Treatment Group B plants showed faster vertical growth, reaching approximately 30 cm at week 8, compared with about 20 cm under the static protocol. Similar trends were observed for biomass accumulation, where the IoT-controlled environment produced approximately 221 g of fresh biomass, compared with 129 g under the non-IoT condition. Leaf area development followed the same pattern, with the IoT-based treatment

RESULTS AND DISCUSSION

reaching nearly 248 cm², while the static protocol reached approximately 179 cm².

The dashed segments shown between the curves illustrate the growth improvement achieved under adaptive environmental control. These results demonstrate that dynamically adjusting environmental conditions using IoT during the growth cycle can significantly enhance plant growth performance compared with a fixed environmental schedule. The complete experimental dataset used to generate these growth curves is provided in Appendix A30.

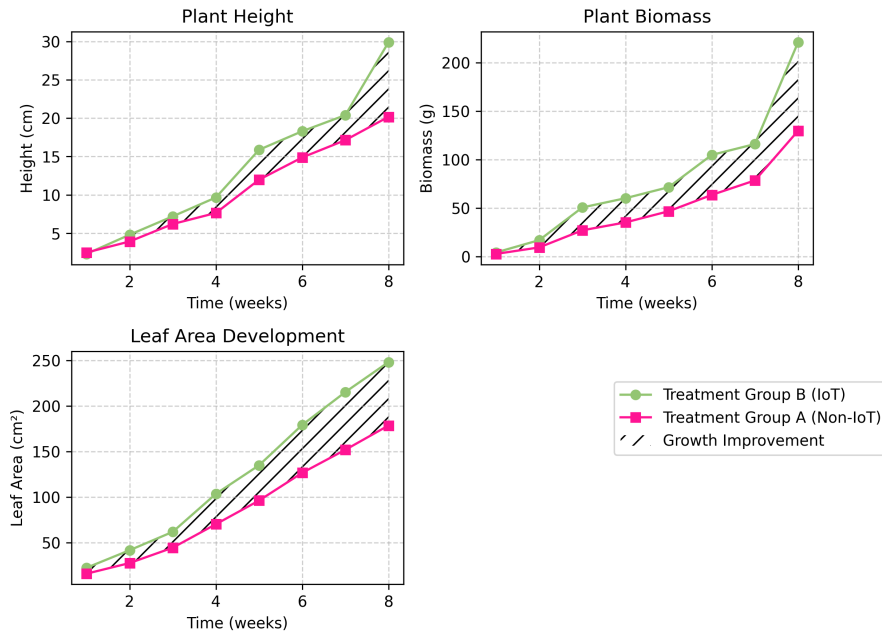


Fig. 4.13. Comparison of Plant Growth: IoT and Non-IoT Monitoring (in weeks).

The graph in Fig. 4.14 compares the resource-use efficiency of the conventional static cultivation protocol (Non-IoT, Treatment Group A) and the adaptive IoT-based control strategy (Treatment Group B).

Resource-use efficiency was calculated as the ratio between the final biomass measured at week 8 and the total supplied resource over the eight-week experiment. The results indicate that the adaptive IoT-based control strategy improved the conversion of environmental inputs into plant biomass. In particular, the biomass produced per unit of supplied light increased from 0.0209 to 0.0353 g per $\mu\text{mol m}^{-2} \text{s}^{-1}$, while CO₂-use efficiency increased from 0.0171 to 0.0580 g per ppm. Similarly, nutrient-use efficiency increased from 0.200 to 0.352 g per ppm, and photoperiod efficiency increased from 1.014 to 1.590 g per hour.

These results demonstrate that the IoT-controlled cultivation system improves the efficiency with which environmental resources are converted into plant biomass. The complete dataset used to compute these efficiency indicators is provided in Appendix A31.

RESULTS AND DISCUSSION

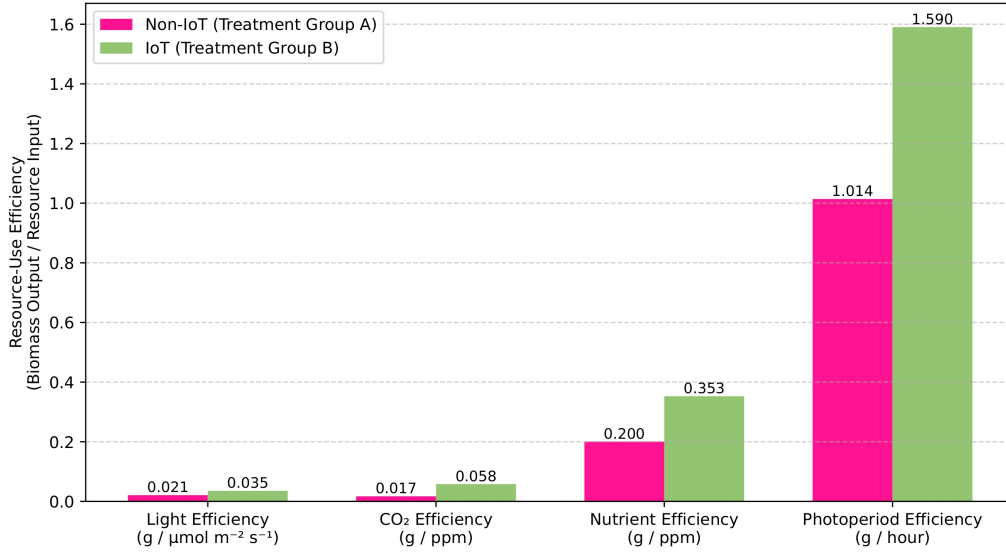


Fig. 4.14 Resource optimization before and after IoT implementation.

4.5. Integrated Agricultural Efficiency Metric (IAEM)

4.5.1. Data Latency Metric (DL)

The average data latency was calculated across 10,000 data points:

Baseline Configuration: DL = 120 ms (without real-time IoT-based monitoring)

Optimized Configuration: DL = 85 ms (with the full IoT + cloud integration)

The optimized configuration reduced latency by ~29%.

4.5.2. Resource Efficiency (RE)

The achieved resource optimization rates were 22.08% for light, 27.55% for nitrogen, 32.22% for phosphorus, 36.58% for potassium, and 28.25% for water.

$$RE = \frac{22.08+27.55+32.22+36.58+28.25}{5} = 29.34 \%$$

4.5.3. Real-Time Alert Accuracy (RA)

The graph in Fig. 4.15 demonstrates a visual improvement for the optimized system, represented by the green line. The baseline system is shown in red.

RESULTS AND DISCUSSION

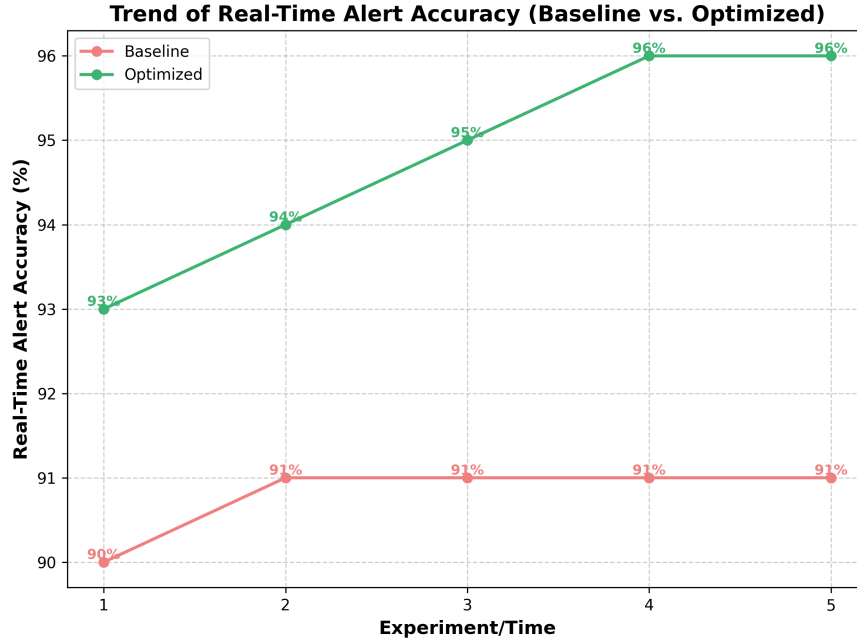


Fig. 4.15. Real-time alert accuracy over experiments.

The baseline system consistently performs at 91% accuracy across the experiments which represents a system without optimization as shown in Appendix A32. The optimized system starts at 93% and improves over experiments as more alerts were forced and thus were implemented in the API. The total improvement in accuracy from the baseline to the optimized system:

$$\Delta RA = RA(\text{optimized}) - RA(\text{baseline}) = 96 - 91\% = 5\%.$$

4.5.4. Cloud Data Availability (CDA)

Azure Blob Storage logs track the number of successful and failed operations.

If:

Total Requests = 10,000

Successful Requests = 9900

Failed Requests = 100

System Uptime = 23.8 h/day

Total Operational Time = 24 h/day

$$SR = \frac{9900}{10000} \times 100 = 99\%$$

$$\text{Uptime (\%)} = \frac{23.8}{24} \times 100 = 99.17\%$$

$$CDA (\%) = 99 \times 99.17 / 100 = 98.17\%$$

4.5.5. Pairwise Comparison algorithm using AHP for weights

The algorithm, specifically tailored for this case study, calculated the weights for the selected metrics:

Resource Efficiency (RE), Real-Time Alert Accuracy (RA), and Data Latency (DL).

The derived weights are as follows:

Resource Efficiency (RE): $\omega_1 = 0.6333$.

Real-Time Alert Accuracy (RA): $\omega_2 = 0.2605$.

Data Latency (DL): $\omega_3 = 0.1062$.

RESULTS AND DISCUSSION

The consistency of the comparisons was validated using the Consistency Ratio (CR), which was calculated as 0.0477 below the accepted threshold of 0.1. With an IAEM value of 0.2105, the system falls into the “Good Performance” category.

Table 4.1. presents the analysis of latency components and success probability for sensor-to-cloud communication using WiFi and Ethernet. The total latency (Ψ) and effective delay (Ψ_{eff}) are calculated for both network types. Additionally, the probability of successful transmission without retries (P_{success}) is provided, showing Ethernet’s higher reliability.

The major difference in performance is observed in β_2 (Communication Delay), where Ethernet has almost half the communication delay of WiFi, leading to an overall improvement in response time. This suggests that network stability is a key factor in optimizing sensor-to-cloud communication. Additionally, Ethernet ensures a 99% success rate, making data transmission more reliable and reducing the need for retransmissions, whereas WiFi’s 95% success rate lead to retries, further increasing its actual response time in real-world scenarios.

Table 4.1. WiFi vs. Ethernet in Sensor-to-Cloud Communication.

Parameter	WiFi	Ethernet
β_1 (Sensor Delay)	5 ms	5 ms
β_2 (Comm Delay)	2.001 ms	1.0001 ms
β_3 (Cloud Processing)	10 ms	10 ms
β_4 (Command Back)	5 ms	5 ms
Sum of β_s (Ψ)	22.001 ms	21.0001 ms
P_{success} (No Retries)	0.95	0.99
Effective Delay Ψ_{eff}	23.158 ms	21.212 ms

4.6. IoT Framework Performance

4.6.1. Analysis of Light intensity

I generated a three-dimensional response surface methodology (RSM) plot to analyze the relationship between Relative Growth Rate (RGR), light intensity, and time, as shown in Fig. 4.16 below. The figure combines the experimental measurements obtained from the three experimental groups: Treatment Group A, Treatment Group B and Treatment Group C with a fitted multivariate regression surface derived from the experimental dataset in Appendices A33,A34 and A35.

RESULTS AND DISCUSSION

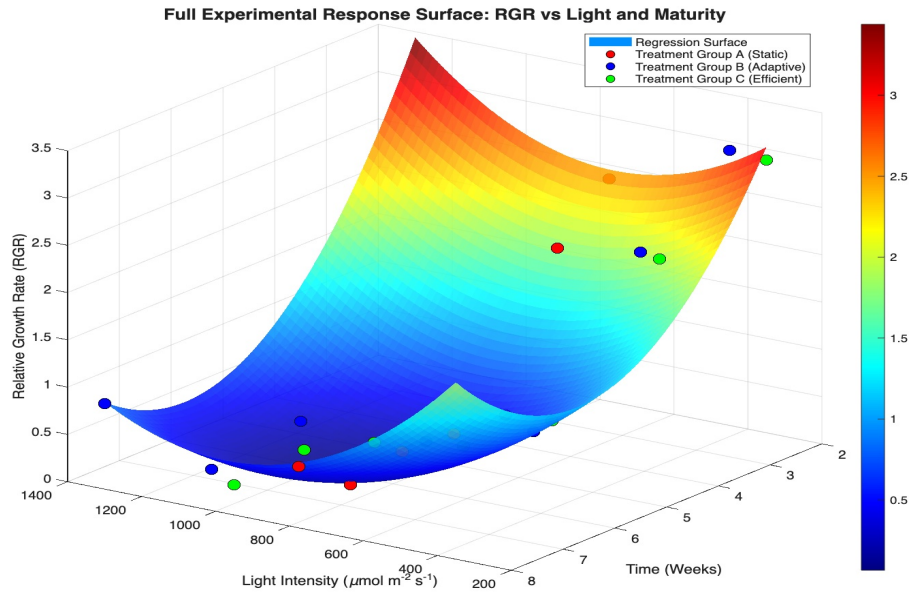


Fig. 4.16. Growth rate vs Light Intensity and Time.

The horizontal axes represent light intensity ($\mu\text{mol m}^{-2} \text{s}^{-1}$) and time (weeks), while the vertical axis represents the Relative Growth Rate (RGR) derived from biomass measurements. The color gradient indicates the magnitude of RGR, ranging from low growth rates (blue) to high growth rates (red).

The experimental points distributed across the surface reveal a non-linear relationship between plant growth and environmental conditions.

This U-shaped curvature along the maturity axis highlights the combined influence of plant development and environmental conditions on biomass accumulation. The smooth regression surface further indicates that the growth response can be approximated by a continuous mathematical relationship between light intensity, photoperiod conditions, and time (weeks).

These observations motivated the development of a multivariate growth response model capable of quantifying the combined influence of environmental factors on plant productivity.

4.6.2. Development of the Growth Response Model

The response surface shown in Fig. 4.16 indicates that plant growth is influenced simultaneously by light intensity, photoperiod, and time (weeks).

Response Surface Methodology (RSM) is widely used in controlled-environment agriculture to describe the combined influence of environmental variables on biological responses. In this approach, the response variable is approximated by a second-order polynomial expansion that captures linear effects, curvature effects, and interactions between factors.

Based on the experimental observations, the Relative Growth Rate (RGR) was selected as the response variable because it normalizes biomass accumulation relative to plant size and allows comparisons across different developmental stages.

RESULTS AND DISCUSSION

The regression coefficients were estimated using ordinary least squares (OLS) implemented in MATLAB in Appendix A36.

Applying the response surface regression model described in Section 3.5.4 to the experimental dataset produced the coefficient estimates shown in Table 4.2.

Table 4.2 Estimated regression coefficients of the RGR model

Term	Estimate	Standard Error	t-Statistic	p-Value
Intercept (β_0)	0.42529	0.17713	2.401	0.032
I_c	-0.00040864	0.00097077	-0.421	0.681
P_c	-0.070762	0.097384	-0.727	0.480
W_c	-0.30738	0.062485	-4.919	0.00028
$I_c P_c$	0.0033401	0.0021076	1.585	0.137
$P_c W_c$	-0.33512	0.28836	-1.162	0.266
P_c^2	-0.13305	0.06398	-2.080	0.0579
W_c^2	0.12503	0.033105	3.777	0.00231

Substituting the estimated coefficients into the response surface equation yields the following predictive model for the relative growth rate in Equation (63):

$$\text{RGR} = 0.4253 - 4.09 \times 10^{-4} I_c - 0.0708 P_c - 0.3074 W_c + 0.00334 (I_c P_c) - 0.3351 (P_c W_c) - 0.1331 P_c^2 + 0.1250 W_c^2 \quad (63)$$

The coefficients were rounded to four significant digits to preserve the predictive accuracy of the regression model.

4.6.3. RGR Model Validation

To evaluate the predictive performance of the stepwise regression model, the experimentally measured relative growth rates (RGR) for Treatment Groups A, B, and C were compared with the model-predicted values over the experimental weeks (weeks 2 to 8). Fig. 4.17 presents these comparisons as time-series plots.

RESULTS AND DISCUSSION

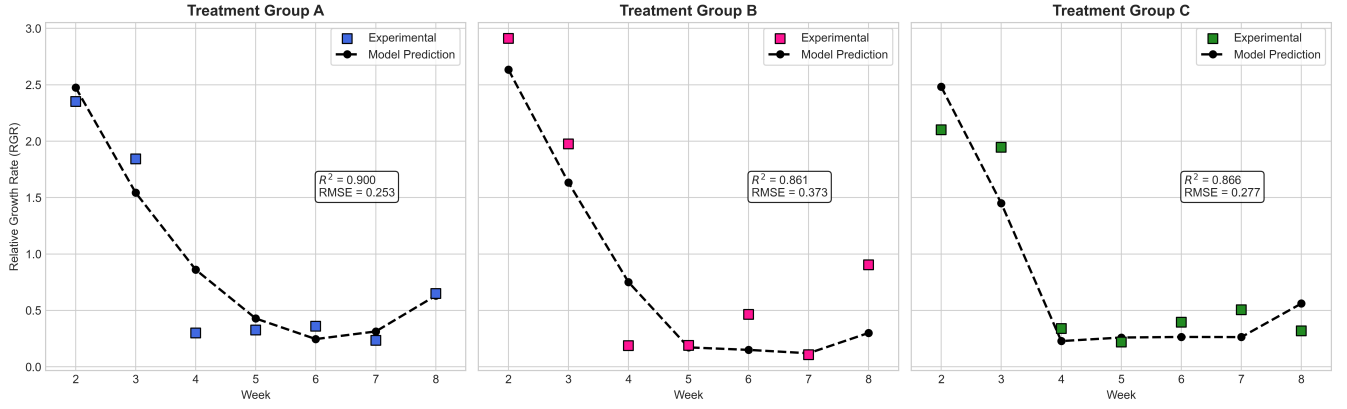


Fig. 4.17. Predictive Accuracy of the RGR Model Across Treatment Groups A, B and C

The predictive accuracy of the model was quantified using the coefficient of determination (R^2) and the root mean square error (RMSE). For each group, these metrics were computed from the experimental and predicted RGR values and are summarised below:

Treatment Group A: $R^2 = [0.900]$, RMSE = [0.253]

Treatment Group B: $R^2 = [0.920]$, RMSE = [0.283]

Treatment Group C: $R^2 = [0.866]$, RMSE = [0.277]

The high R^2 values (all above [0.8]) indicate that the model explains a large proportion of the variance in RGR, while the low RMSE values (all below [0.5]) confirm the model's precision.

4.7. Optimal Configuration

4.7.1. Light Use Efficiency (LUE) Metrics

Detailed quantitative metrics for Configurations A–D, describing the relationships between photosynthetically active radiation (PAR) emitted by the lighting system, PAR received at canopy level, and PAR absorbed by the crop, together with the resulting Light Use Efficiency of Photosynthesis (LUEP) and Light Use Efficiency of Lighting (LUEL), are provided in Appendix A37.

For example, in Configuration A, $300 \mu\text{mol m}^{-2} \text{s}^{-1}$ of PAR is emitted, of which $260.4 \mu\text{mol m}^{-2} \text{s}^{-1}$ is received at the canopy and $228 \mu\text{mol m}^{-2} \text{s}^{-1}$ is absorbed, yielding a Light Use Efficiency of Photosynthesis (LUEP) of 0.76 and a Light Use Efficiency of Lighting (LUEL) of 0.87. As light intensity increases, Configuration D emits $2000 \mu\text{mol m}^{-2} \text{s}^{-1}$, with $1688 \mu\text{mol m}^{-2} \text{s}^{-1}$ received and $1540 \mu\text{mol m}^{-2} \text{s}^{-1}$ absorbed. This results in a slightly lower canopy-level efficiency (LUEP = 0.77) compared to the peak observed in Configuration C (LUEP = 0.83), while the lamp-level efficiency decreases to 0.84 from a maximum of 0.89 in Configuration B.

Fig. 4.18 presents a comparative bar chart of LUEP and LUEL across Configurations A–D. The results show that canopy-level efficiency (LUEP) improves from low to moderate light intensities and then declines at higher irradiance, indicating the onset of saturation effects. Similarly, lamp-level efficiency (LUEL) reaches its maximum at intermediate configurations before decreasing under high light conditions.

RESULTS AND DISCUSSION

These results highlight that optimal efficiency is achieved at intermediate light intensities, where both light delivery and canopy absorption are balanced, rather than at the highest irradiance levels.

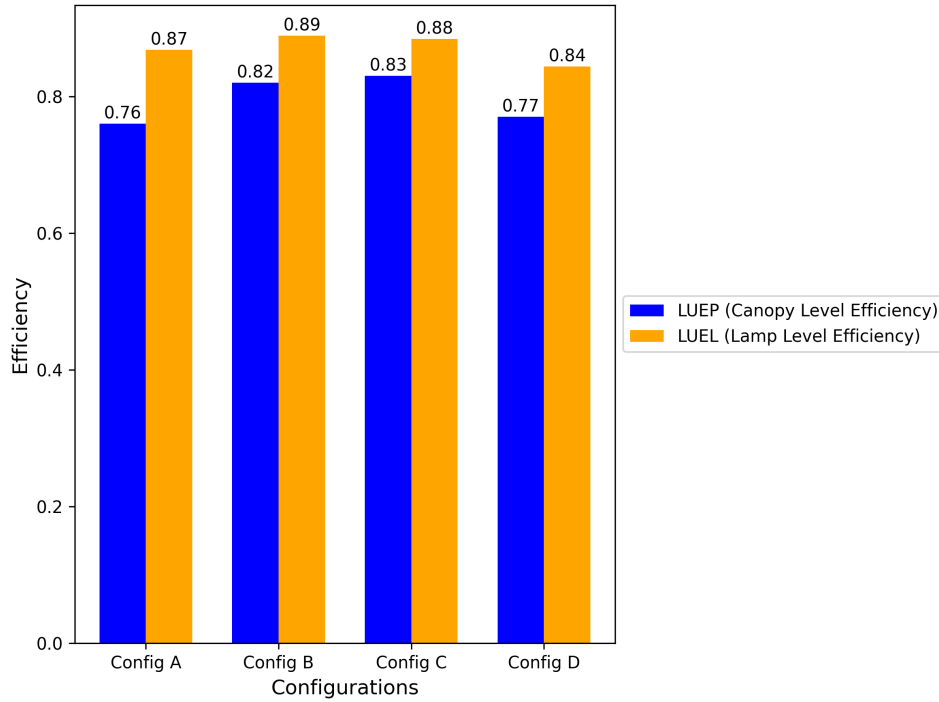


Fig. 4.18 LUEP and LUEL Comparisons Across Configurations.

4.7.2. Energy Efficiency Analysis

Fig. 4.19 illustrates the relationship between Leaf Area Index (LAI) and light use efficiency metrics (LUEP and LUEL). LUEP and LUEL were computed from the four configurations shown in Appendices A9 and A37. These equations quantify the efficiency of light utilization at both the plant canopy and lamp levels, as a function of the Leaf Area Index (LAI). A growth model was developed to assess the interaction between light and the plant canopy. The model incorporated parameters such as light intensity, spectral distribution, and canopy structure (measured by LAI). LAI was varied across a range of values to get different plant densities and foliage coverage, where LUEP and LUEL were computed as following:

The fraction of PAR absorbed by the canopy for LUEP.

The fraction of PAR transmitted to the canopy for LUEL.

RESULTS AND DISCUSSION

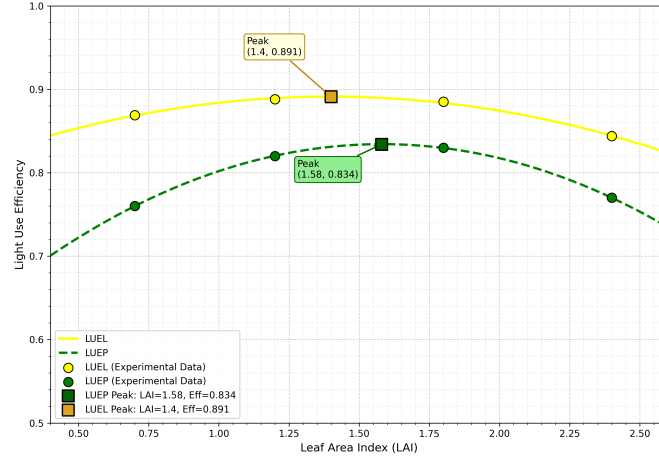


Fig. 4.19. Light Use Efficiency vs. LAI.

The data shown in Fig. 4.19 was analyzed using regression techniques to determine the mathematical relationship between LAI and the efficiency metrics (LUEP and LUEL).

Results were recorded for each configuration run. The equations were fitted to the data using a polynomial regression approach, the model in Appendix A37 is leading to the following results in Equation (64) and Equation (65):

LUEP (Canopy Level Efficiency):

$$\text{LUEP} = 0.5948 + 0.3029x - 0.0958x^2 \quad (64)$$

LUEL (Lamp Level Efficiency):

$$\text{LUEL} = 0.7999 + 0.1306x - 0.0467x^2 \quad (65)$$

Both equations in (64) and (65) show a non-linear relationship, where the efficiency metrics increase with LAI up to a certain point and then decline as LAI becomes too high (due to shading effects and light saturation).

4.8. Comparative Growth Performance of Treatment Groups

4.8.1. Chlorophyll Evolution

Fig. 4.20 shows the evolution of chlorophyll content (SPAD) over eight weeks for the three treatment groups. Data are presented as mean $\pm$ SEM based on five experimental replicates. Treatment Group A increased steadily under static conditions. Treatment Group B exhibited a stronger adaptive response, with relatively low variability across weeks. Treatment Group C reached the highest final chlorophyll level with slightly greater variability, indicating higher sensitivity to environmental fluctuations and resource constraints.

The regression analysis demonstrated strong linear temporal progression in chlorophyll accumulation across all treatments, with high coefficients of determination ($R^2_{\text{GroupA}} = 0.973$, $R^2_{\text{GroupB}} = 0.964$, $R^2_{\text{GroupC}} = 0.972$). These high R^2 values indicate that chlorophyll development followed a highly consistent growth pattern over time under all experimental conditions.

RESULTS AND DISCUSSION

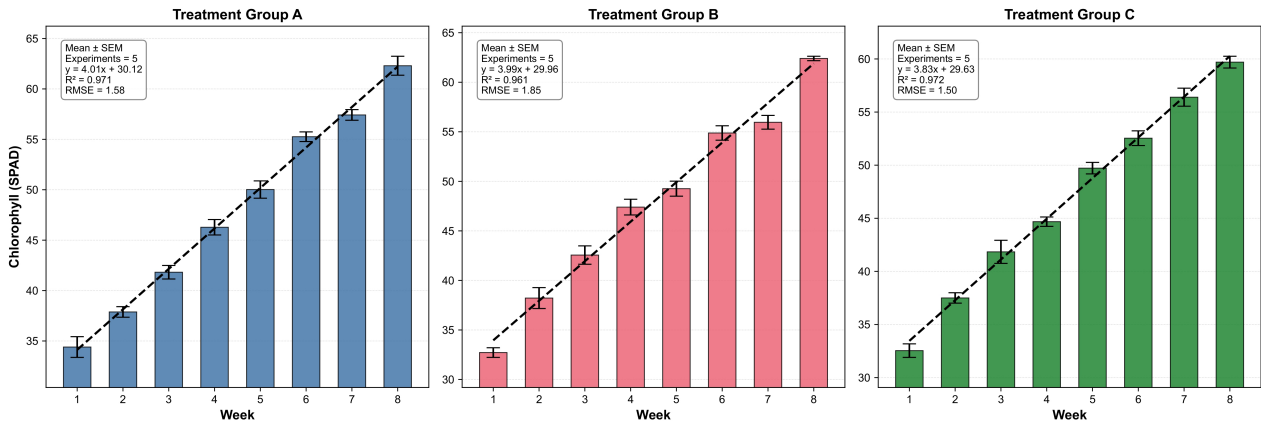


Fig. 4.20. Chlorophyll content (SPAD) under different treatment conditions

4.8.2. Biomass Accumulation

Fig. 4.21 presents biomass accumulation over eight weeks for the three Treatment Groups. Data are presented as mean $\pm$ SEM based on five experimental replicates. Treatment Group A increased gradually under static conditions. Treatment Group B showed the strongest biomass production, reflecting the effectiveness of adaptive resource regulation and optimized environmental control. Treatment Group C showed substantial biomass accumulation but lower final productivity compared with Treatment Group B.

The regression analysis revealed strong temporal biomass progression across all treatments, with high coefficients of determination ($R^2_{\text{GroupA}} = 0.910$, $R^2_{\text{GroupB}} = 0.964$, $R^2_{\text{GroupC}} = 0.889$). These results indicate that biomass accumulation followed a highly structured growth pattern over time, particularly under adaptive environmental regulation in Treatment Group B.

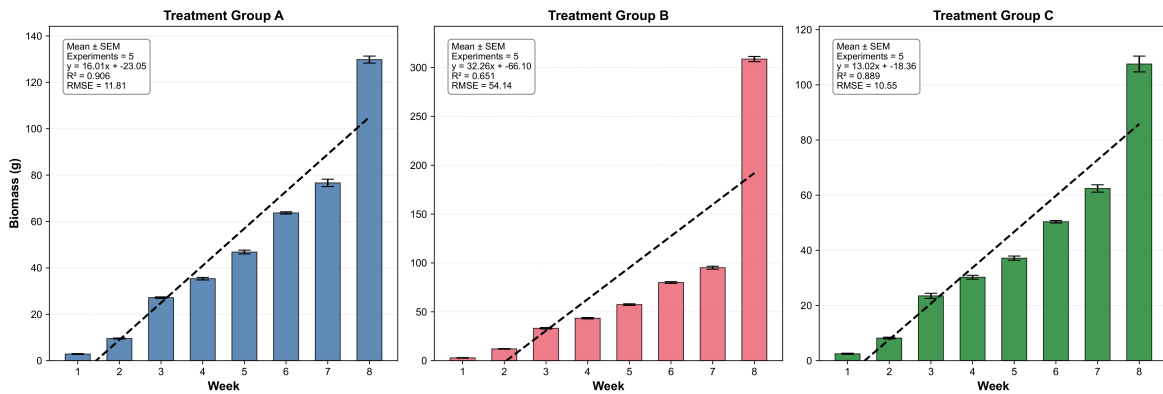


Fig. 4.21. Biomass Accumulation under different treatment conditions

4.8.3. Height Progression

Fig. 4.22 shows plant height progression over eight weeks for the three Treatment Groups. Data are presented as mean $\pm$ SEM based on five experimental replicates. Treatment Group A increased steadily under static conditions. Treatment Group B exhibited the greatest height gain, reflecting effective adaptive regulation of environmental growth drivers. Treatment Group C indicated consistent but less pronounced vertical growth compared with Treatment Group B. The regression analysis demonstrated highly linear height progression across all treatments, with strong coefficients of determination ($R^2_{\text{GroupA}} = 0.983$, $R^2_{\text{GroupB}} = 0.977$, $R^2_{\text{GroupC}} = 0.985$). These

RESULTS AND DISCUSSION

results indicate that plant height development followed a highly consistent temporal growth pattern throughout the experimental period.

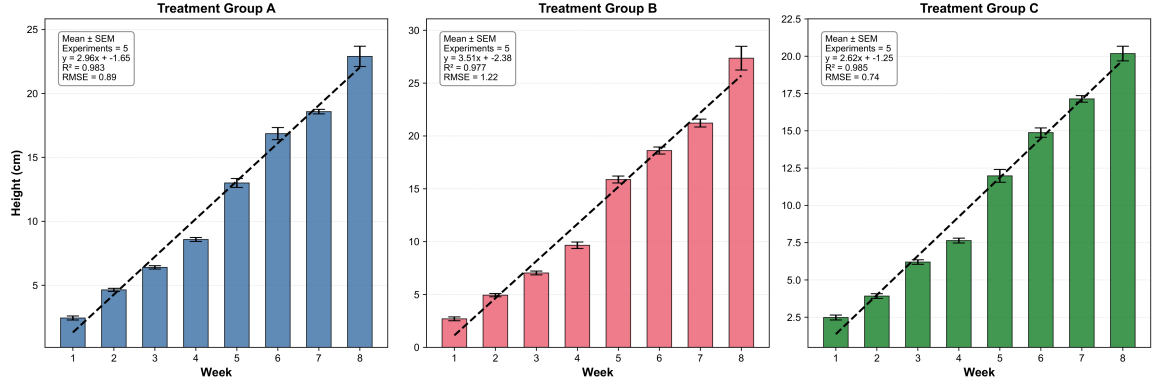


Fig. 4.22. Height Progression under different treatment conditions

4.8.4. Leaf Area Development

Fig. 4.23 presents leaf area development over eight weeks for the three Treatment Groups. Data are presented as mean $\pm$ SEM based on five experimental replicates. Treatment Group A increased progressively under static conditions. Treatment Group B showed the most pronounced leaf-area expansion, indicating strong adaptive allocation toward canopy development and light interception. Treatment Group C reflected efficient leaf expansion but lower maximal surface area compared with Treatment Group B.

The regression analysis demonstrated highly linear temporal development of leaf area across all treatments, with strong coefficients of determination ($R^2_{\text{GroupA}} = 0.988$, $R^2_{\text{GroupB}} = 0.985$, $R^2_{\text{GroupC}} = 0.987$). These results indicate that canopy expansion followed a highly structured and consistent growth pattern throughout the experimental period.

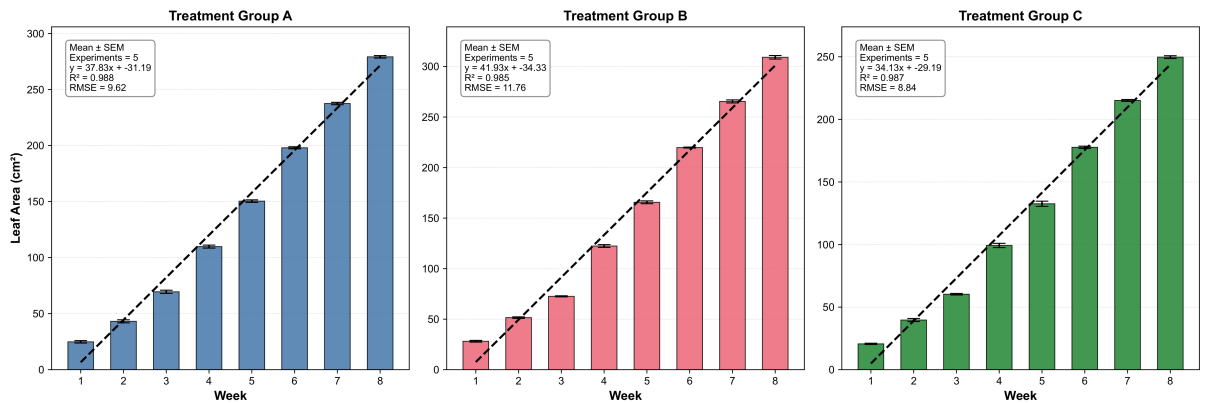


Fig. 4.23. Leaf Area under different treatment conditions

4.8.5. Allometric Relationship Between Height (H) and Biomass (B)

Fig. 4.24 quantifies the allometric relationship between plant height and biomass using the power-law formulation $H = \alpha B^\beta$, revealing distinct trait-coupling behaviors across Treatment Groups. Treatment Group A exhibits a scaling exponent ($\beta = 0.5885$), indicating sublinear growth where

RESULTS AND DISCUSSION

height increases at a diminishing rate relative to biomass accumulation, consistent with resource saturation under static environmental control. Treatment Group C shows a slightly higher exponent ($\beta = 0.6086$), reflecting stronger biomass-driven height expansion but with increased dispersion, suggesting greater sensitivity to fluctuating resource availability.

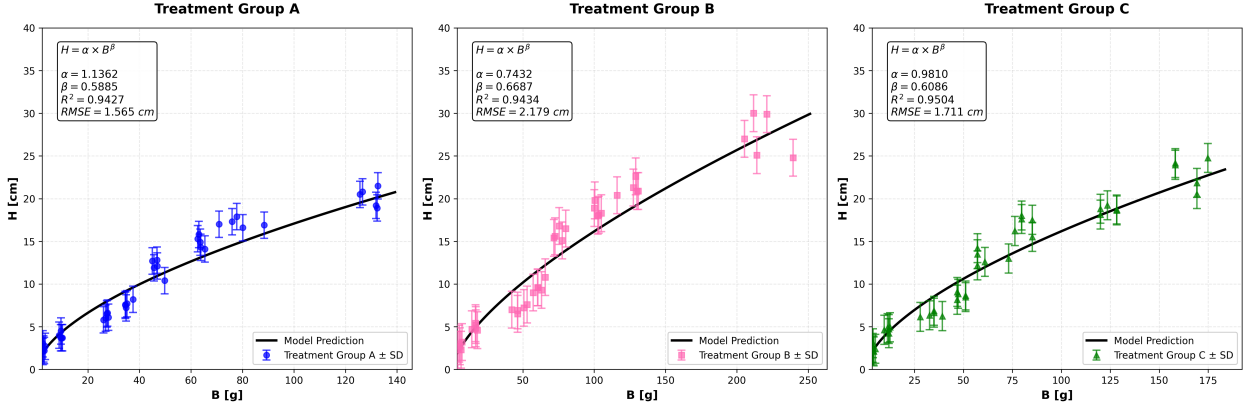


Fig. 4.24. Allometric Relationship Between Height (H) and Biomass (B) Across Treatment Groups

Treatment Group B demonstrates the most coherent allometric structure, with the highest scaling exponent ($\beta = 0.6687$) and the strongest coefficient of determination ($R^2 = 0.9434$), indicating tightly coordinated biomass partitioning and vertical growth. Despite operating across a wider biomass range, Treatment Group B maintains relatively low prediction error ($RMSE = 2.179 \text{ cm}$), confirming that adaptive environmental regulation preserves proportional trait development under variable conditions.

The consistently high coefficients of determination across all treatments ($R^2_{\text{GroupA}} = 0.9427$, $R^2_{\text{GroupB}} = 0.9434$, $R^2_{\text{GroupC}} = 0.9504$) further validate the robustness of the allometric formulation and support the biological realism of the proposed four-dimensional growth model. The reduced structural dispersion observed under adaptive regulation suggests that intelligent environmental control enhances trait coordination rather than amplifying growth variability. The full dataset of measured traits is provided in Appendix A38.

To further evaluate the generalizability of the allometric relationship, the data from the three Treatment Groups were pooled and fitted using a single global scaling function, as illustrated in Fig. 4.25.

RESULTS AND DISCUSSION

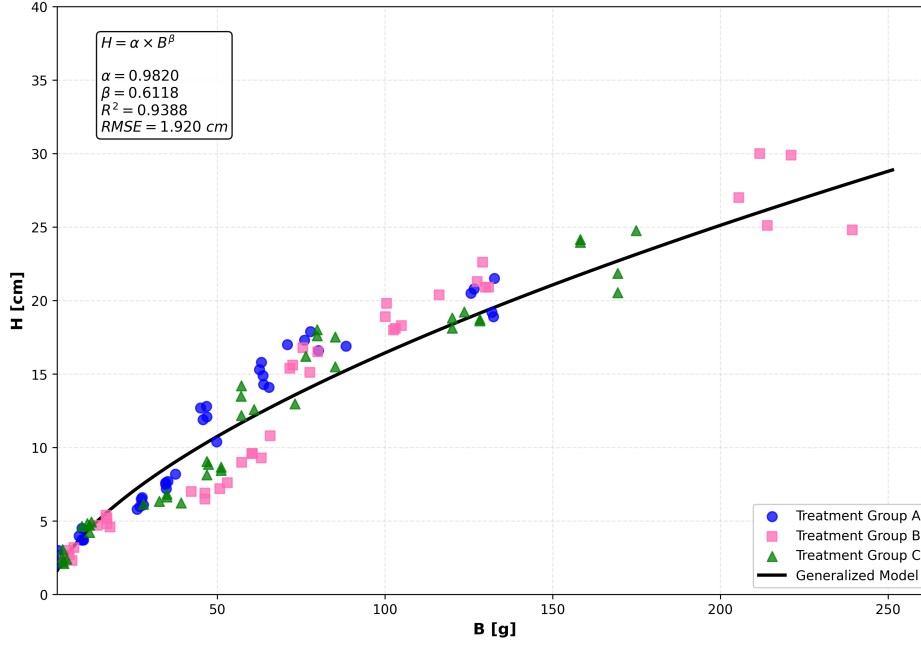


Fig. 4.25. Generalized Allometric Model $H = \alpha B^\beta$

The resulting generalized model maintained strong statistical performance across the complete dataset, with a coefficient of determination ($R^2 = 0.9388$) and a root mean square error ($RMSE = 1.920 \text{ cm}$). The estimated scaling exponent ($\beta = 0.6118$) indicates a consistent sublinear relationship between biomass accumulation and plant height across all treatment conditions.

4.8.6. Sensitivity of Growth Rate to Nutrient Concentration

Fig. 4.26 illustrates the relationship between nutrient concentration and growth rate across the three Treatment Groups. Treatment Group A shows a flat response due to its static nutrient regime, confirming that growth rate remains largely insensitive to nutrient variation under fixed environmental conditions. In contrast, Treatment Groups B and C exhibit positive growth responses as nutrient concentration increases. Treatment Group B demonstrates a moderate and controlled increase in growth rate, reflecting efficient nutrient utilization under adaptive regulation. Treatment Group C shows a steeper response, indicating stronger sensitivity to nutrient input but also greater variability in growth behavior. The relatively low coefficients of determination observed for the adaptive treatments ($R^2_{\text{GroupB}} = 0.261$, $R^2_{\text{GroupC}} = 0.245$) suggest that the relationship between nutrient concentration and growth rate is not fully explained by a simple linear model, indicating the presence of nonlinear physiological responses and biological variability.

RESULTS AND DISCUSSION

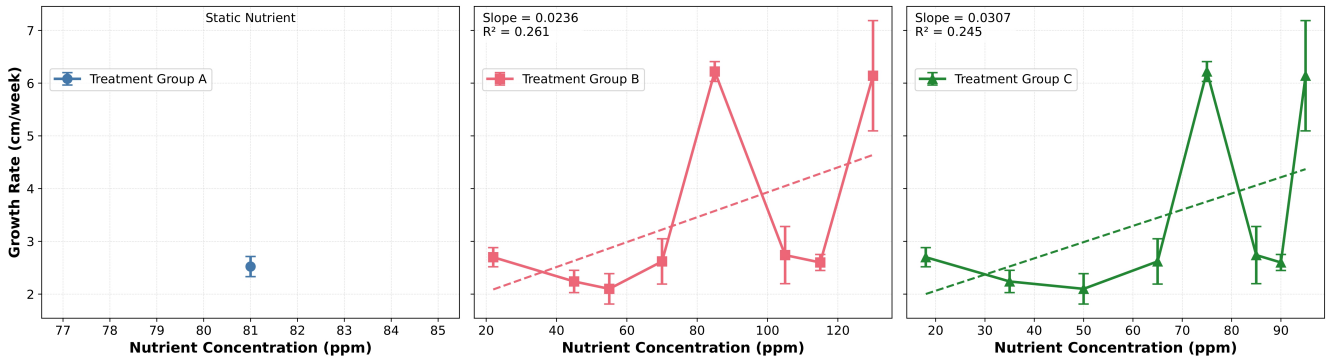


Fig. 4.26. Effect of Nutrient Concentration (ppm) on Growth Rate

Fig. 4.27 presents the raw sensitivity of growth rate to nutrient concentration, expressed as the change in growth rate per unit variation in nutrient input. Treatment Group A exhibits negligible sensitivity due to static nutrient conditions. Treatment Group B shows a raw sensitivity value of 0.0236, indicating a controlled growth response under adaptive nutrient regulation. Treatment Group C displays the highest raw sensitivity (0.0307), reflecting stronger responsiveness to nutrient variation but also greater variability compared with Treatment Group B. The relatively low coefficients of determination for the adaptive treatments ($R^2_{\text{Group B}} = 0.261$, $R^2_{\text{Group C}} = 0.245$) indicate that nutrient-driven growth responses are not fully explained by a simple linear relationship, suggesting nonlinear physiological dynamics and biological variability.

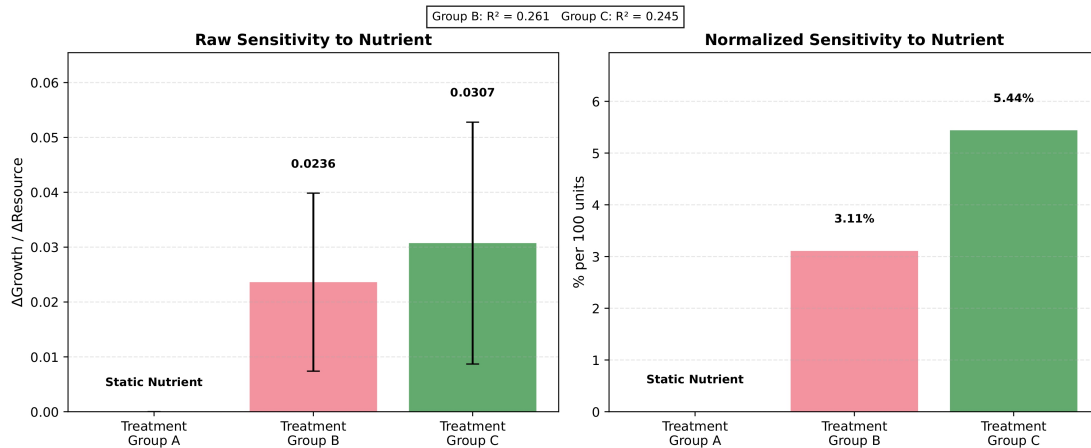


Fig. 4.27. Raw Sensitivity of Growth Rate to Nutrient Concentration

The normalized sensitivity analysis further confirms these observations. Treatment Group B achieves a normalized sensitivity of 3.11%, reflecting efficient nutrient utilization under adaptive control, whereas Treatment Group C reaches 5.44%, indicating stronger dependence on nutrient enrichment. Overall, Treatment Group B maintains a more balanced trade-off between responsiveness and stability under controlled environmental conditions.

4.8.7. Sensitivity of Growth Rate to Light

Treatment Group A exhibits a flat response due to its static light regime, confirming the absence of measurable growth sensitivity under constant illumination, as shown in Fig. 4.28. In contrast,

RESULTS AND DISCUSSION

Treatment Groups B and C demonstrate positive relationships between light intensity and growth rate, indicating that increased irradiance enhances growth under adaptive or efficiency-based regulation strategies. Treatment Group B shows a controlled and gradual increase in growth rate, reflecting stable adaptive light regulation, whereas Treatment Group C presents a steeper response, indicating stronger sensitivity to light input but with increased variability at higher irradiance levels.

The relatively low coefficients of determination for the adaptive treatments ($R^2_{\text{GroupB}} = 0.261$, $R^2_{\text{GroupC}} = 0.245$) suggest that the relationship between light intensity and growth rate is not fully explained by a linear model, indicating nonlinear physiological responses and biological variability under changing illumination conditions.

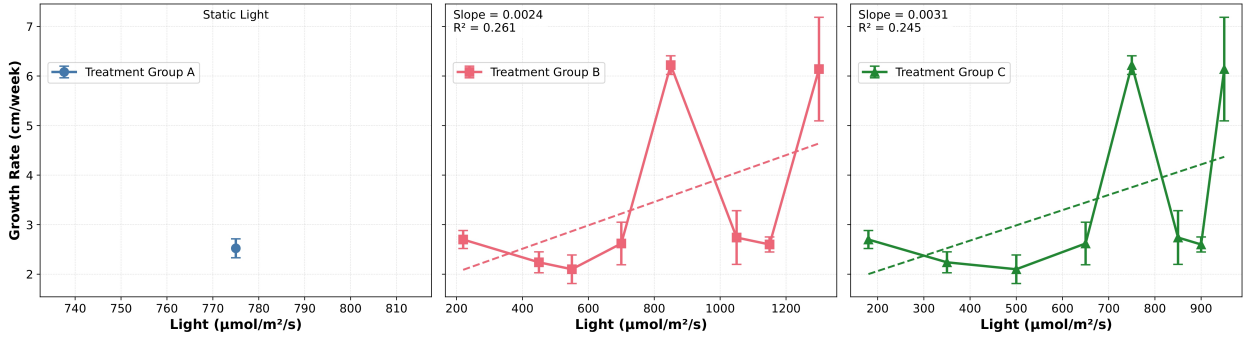


Fig. 4.28. Effect of Light Intensity ($\mu\text{mol/m}^2/\text{s}$) on Growth Rate

Fig. 4.29 presents the raw sensitivity of growth rate to light intensity, expressed as the change in growth rate per unit variation in light input. Treatment Group A shows negligible sensitivity due to static lighting conditions. Treatment Group B exhibits a raw sensitivity value of 0.0024, indicating a moderate and stable growth response under adaptive light regulation. Treatment Group C displays the highest raw sensitivity (0.0031), reflecting stronger responsiveness to increasing light intensity but also greater variability compared with Treatment Group B.

The relatively low coefficients of determination for the adaptive treatments ($R^2_{\text{GroupB}} = 0.261$, $R^2_{\text{GroupC}} = 0.245$) suggest that the relationship between light intensity and growth rate is not fully explained by a simple linear model, indicating nonlinear physiological responses and biological variability.

RESULTS AND DISCUSSION

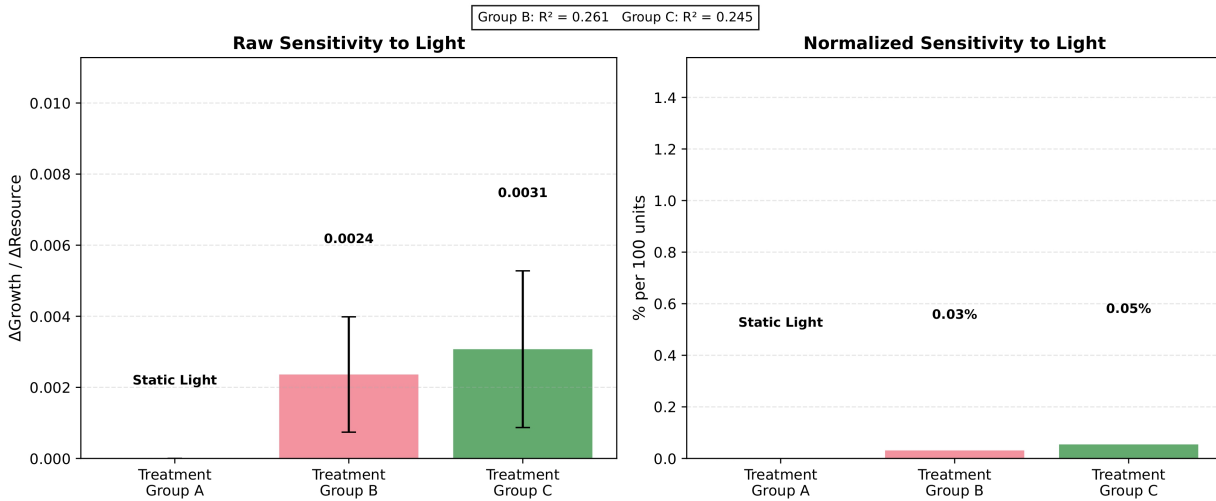


Fig. 4.29. Raw and Normalized Sensitivity of Growth Rate to Light Intensity

The normalized sensitivity analysis further supports these observations. Treatment Group B achieves a normalized sensitivity of 0.03%, indicating efficient utilization of additional light under adaptive control, whereas Treatment Group C reaches 0.05%, confirming stronger dependence on light input. While Treatment Group C maximizes responsiveness, Treatment Group B maintains a more balanced trade-off between light-use efficiency and system stability.

4.8.8. Sensitivity of Growth Rate to CO₂

Fig. 4.30 illustrates the effect of CO₂ concentration on growth rate across the three treatment groups. Treatment Group A shows no measurable response, as CO₂ levels were held constant, resulting in a flat growth profile. In contrast, Treatment Group B exhibits a modest positive relationship, indicating a limited but controlled growth response to increasing CO₂ concentration. Treatment Group C displays a steeper positive relationship, demonstrating a stronger dependence of growth rate on CO₂ availability. However, the increased dispersion observed in Treatment Group C suggests heightened sensitivity and reduced stability under fluctuating CO₂ conditions. The relatively low coefficients of determination observed for the adaptive treatments ($R^2_{\text{Group B}} = 0.260$, $R^2_{\text{Group C}} = 0.249$) indicate that the relationship between CO₂ concentration and growth rate is not fully explained by a simple linear model, suggesting the presence of nonlinear growth dynamics and biological variability. These results further highlight the role of CO₂ as a significant growth driver under dynamic environmental regulation.

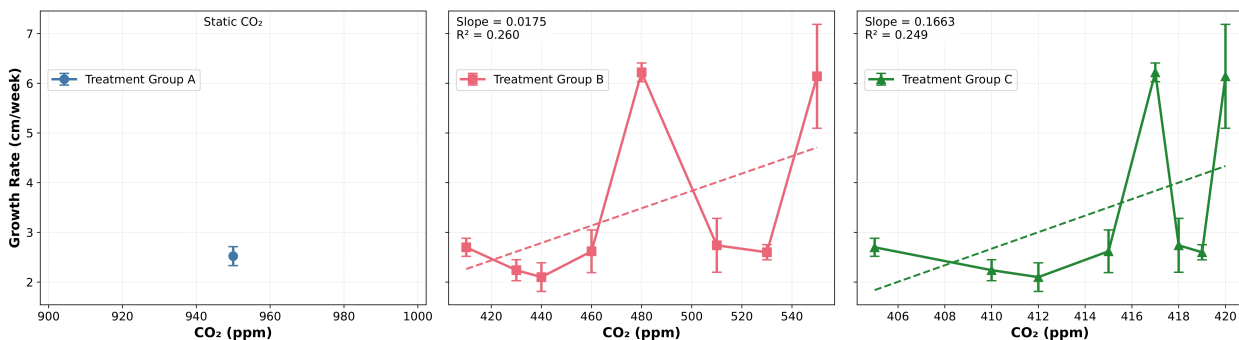


Fig. 4.30. Effect of CO₂ (ppm) on Growth Rate

Fig. 4.31 presents the raw sensitivity of growth rate to CO₂ concentration, expressed as the change

RESULTS AND DISCUSSION

in growth rate per unit variation in CO₂ input. Treatment Group A exhibits negligible sensitivity due to static CO₂ conditions. Treatment Group B shows a low raw sensitivity value of 0.0175, indicating a conservative and controlled growth response under adaptive regulation. In contrast, Treatment Group C demonstrates a substantially higher raw sensitivity (0.1663), nearly an order of magnitude greater than Treatment Group B. Although this indicates a strong CO₂-driven growth response, the larger associated variability suggests reduced robustness compared with the adaptive strategy implemented in Treatment Group B. The regression fit values remain relatively similar between the adaptive treatments ($R^2_{\text{GroupB}} = 0.260$, $R^2_{\text{GroupC}} = 0.249$), indicating comparable explanatory consistency despite differences in response magnitude.

The normalized sensitivity analysis further confirms these observations. Treatment Group A again shows negligible sensitivity under fixed CO₂ conditions. Treatment Group B achieves a normalized sensitivity of 0.36%, reflecting limited but efficient CO₂ utilization under adaptive control. Treatment Group C reaches a markedly higher normalized sensitivity of 4.03%, indicating strong dependence on CO₂ enrichment. While Treatment Group C maximizes responsiveness, Treatment Group B maintains a more balanced trade-off between responsiveness and stability, reinforcing its suitability for controlled and sustainable growth conditions.

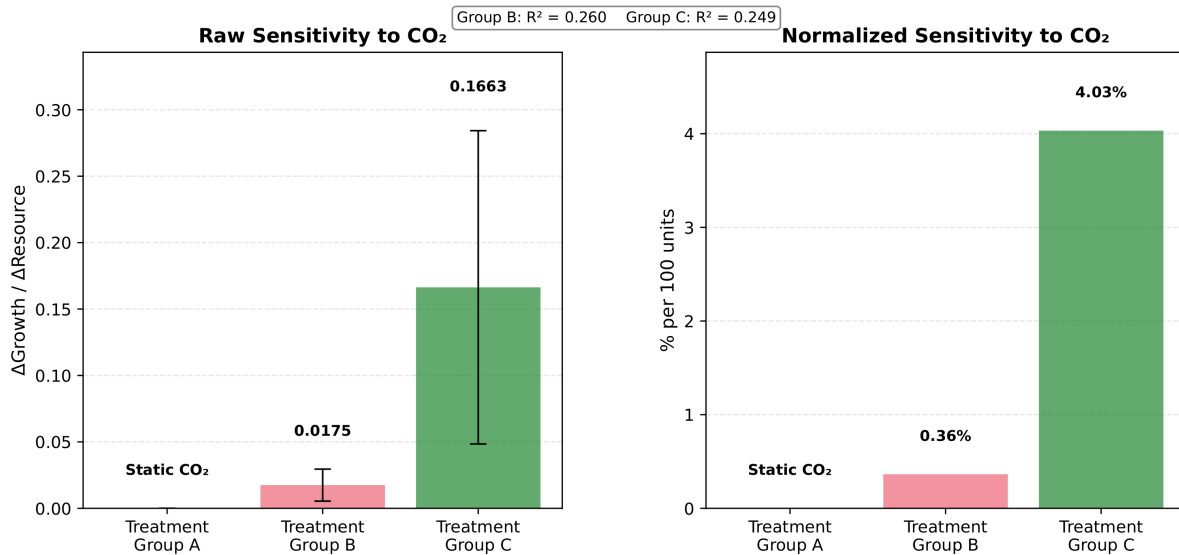


Fig. 4.31. Raw and Normalized Sensitivity of Growth Rate to CO₂ Concentration

Overall, growth sensitivity followed the biologically coherent hierarchy CO₂ > nutrient > light, confirming the validity of the experimental design and model structure. Treatment Group B consistently achieved the best trade-off between responsiveness and stability, whereas Treatment Group C maximized sensitivity at the cost of robustness and Treatment Group A remained insensitive due to static control.

4.9. Advanced Growth Four-Dimensional Model Validation

The multi-dimensional logistic model was parameterized using the euler approximation model for each Treatment Group (A, B, C) in A39. I estimated the parameters K (carrying capacity) and r (intrinsic growth rate) directly from the observed time-series data:

RESULTS AND DISCUSSION

For each Treatment group, the experimental measurements are stored in an 8×4 matrix whose columns correspond to:

1. Height (cm)
2. Biomass (g)
3. Chlorophyll (SPAD)
4. Leaf area (cm²)

Each row represents a weekly observation from week 1 to week 8 in Equation (66), Equation (67) and Equation (68):

$$D_A = \begin{bmatrix} 2.48 & 2.852 & 34.4 & 15.91 \\ 3.92 & 9.56 & 37.88 & 27.62 \\ \vdots & \vdots & \vdots & \vdots \\ 20.18 & 129.78 & 62.30 & 178.66 \end{bmatrix} \quad (66)$$

$$D_B = \begin{bmatrix} 2.3 & 4.36 & 31.7 & 22.4 \\ 4.8 & 17.05 & 37.4 & 41.7 \\ \vdots & \vdots & \vdots & \vdots \\ 29.9 & 221.07 & 60.5 & 248.3 \end{bmatrix} \quad (67)$$

$$D_C = \begin{bmatrix} 3.06 & 3.83 & 31.8 & 13.0 \\ 4.23 & 11.88 & 38.5 & 29.6 \\ \vdots & \vdots & \vdots & \vdots \\ 23.96 & 158.25 & 61.7 & 212.8 \end{bmatrix} \quad (68)$$

The same procedure was applied to all traits (height, biomass, chlorophyll, and leaf area) and for each experimental Treatment Group (A, B, and C).

The final calibrated parameters for Treatment Group A are summarized below in Equation (69), Equation (70), Equation (71) and Equation (72). The corresponding fitted logistic curves and their agreement with the measured growth traits are illustrated in Fig. 4.32 (Measured Growth Traits vs Logistic Model Treatment Group A):

$$S_A^{\text{chlorophyll}} = \{K_A^{\text{chlorophyll}} = 98.38 \text{ SPAD}, r_A^{\text{chlorophyll}} = 0.166 / \text{week}\} \quad (69)$$

$$S_A^{\text{Biomass}} = \{K_A^{\text{Biomass}} = 168.85 \text{ g}, r_A^{\text{Biomass}} = 0.716 / \text{week}\} \quad (70)$$

$$S_A^{\text{Height}} = \{K_A^{\text{Height}} = 25.2 \text{ cm}, r_A^{\text{Height}} = 0.509 / \text{week}\} \quad (71)$$

$$S_A^{\text{Leaf Area}} = \{K_A^{\text{Leaf Area}} = 214.05 \text{ cm}^2, r_A^{\text{Leaf Area}} = 0.581 / \text{week}\} \quad (72)$$

RESULTS AND DISCUSSION

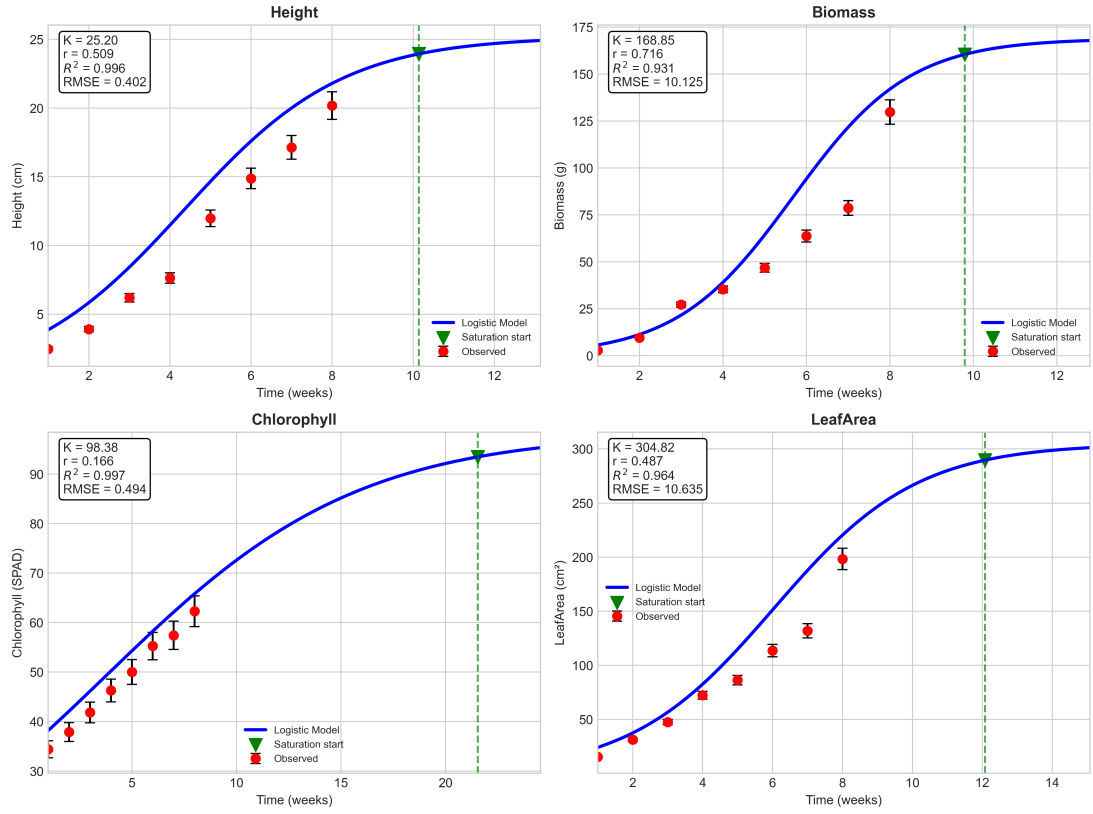


Fig. 4.32. Measured Growth Traits vs Logistic Model (Treatment Group A)

The final calibrated parameters for Treatment Group B are summarized below in Equation (73), Equation (74), Equation (75) and Equation (76). The corresponding fitted logistic curves and their agreement with the measured growth traits are illustrated in Fig. 4.33 (Measured Growth Traits vs Logistic Model Treatment Group B):

$$S_B^{\text{chlorophyll}} = \{K_B^{\text{chlorophyll}} = 70.68 \text{ SPAD}, r_B^{\text{chlorophyll}} = 0.274 / \text{week}\} \quad (73)$$

$$S_B^{\text{Biomass}} = \{K_B^{\text{Biomass}} = 332.96 \text{ g}, r_B^{\text{Biomass}} = 0.688 / \text{week}\} \quad (74)$$

$$S_B^{\text{Height}} = \{K_B^{\text{Height}} = 37.20 \text{ cm}, r_B^{\text{Height}} = 0.546 / \text{week}\} \quad (75)$$

$$S_B^{\text{Leaf Area}} = \{K_B^{\text{Leaf Area}} = 292.09 \text{ cm}^2, r_B^{\text{Leaf Area}} = 0.594 / \text{week}\} \quad (76)$$

RESULTS AND DISCUSSION

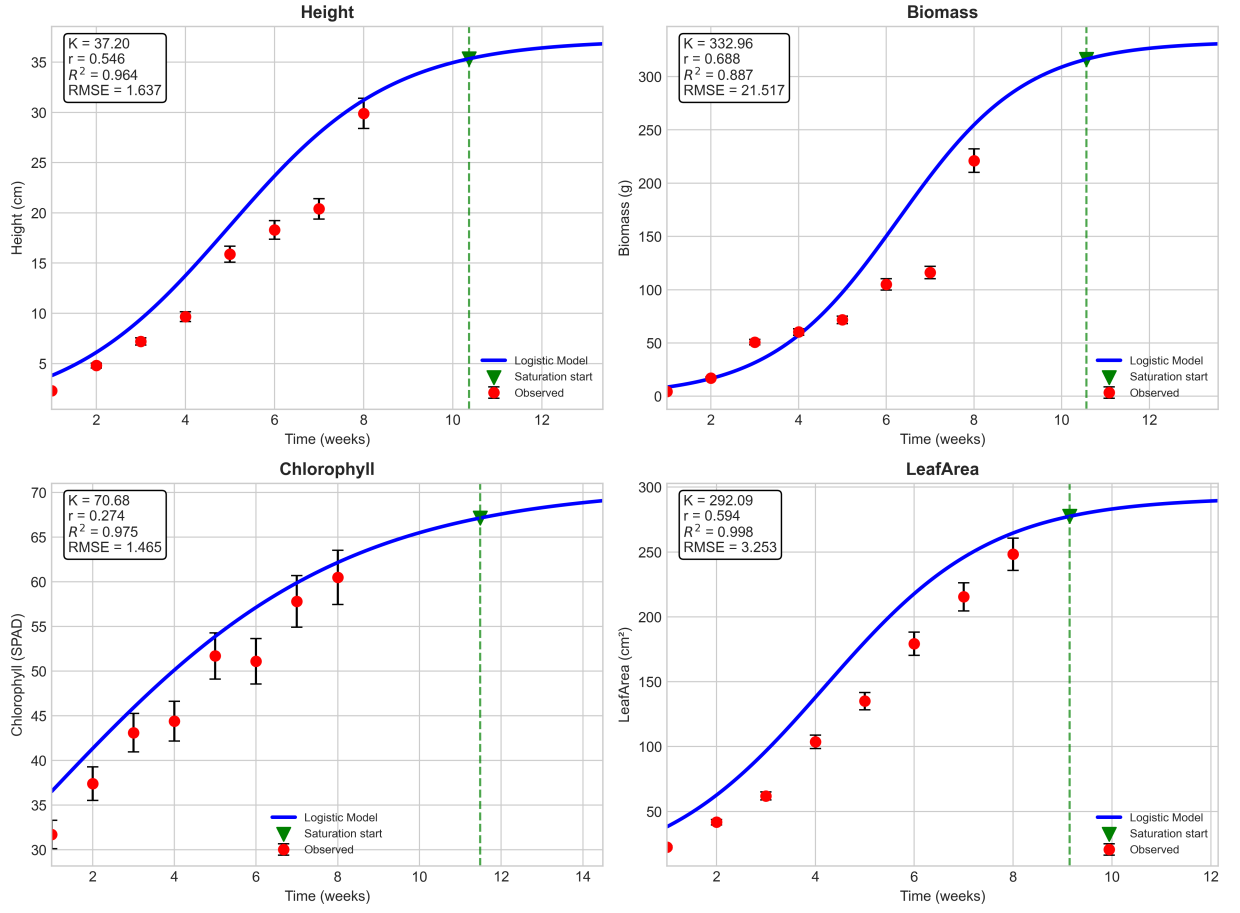


Fig. 4.33. Measured Growth Traits vs Logistic Model (Treatment Group B)

The final calibrated parameters for Treatment Group C are summarized below in Equation (77), Equation (78), Equation (79) and Equation (80). The corresponding fitted logistic curves and their agreement with the measured growth traits are illustrated in Fig. 4.34 (Measured Growth Traits vs Logistic Model Treatment Group C):

$$S_C^{\text{chlorophyll}} = \{K_C^{\text{chlorophyll}} = 67.55 \text{ SPAD}, r_C^{\text{chlorophyll}} = 0.326 / \text{week}\} \quad (77)$$

$$S_C^{\text{Biomass}} = \{K_C^{\text{Biomass}} = 196.21 \text{ g}, r_C^{\text{Biomass}} = 0.743 / \text{week}\} \quad (78)$$

$$S_C^{\text{Height}} = \{K_C^{\text{Height}} = 32.76 \text{ cm}, r_C^{\text{Height}} = 0.459 / \text{week}\} \quad (79)$$

$$S_C^{\text{Leaf Area}} = \{K_C^{\text{Leaf Area}} = 214.19 \text{ cm}^2, r_C^{\text{Leaf Area}} = 0.806 / \text{week}\} \quad (80)$$

RESULTS AND DISCUSSION

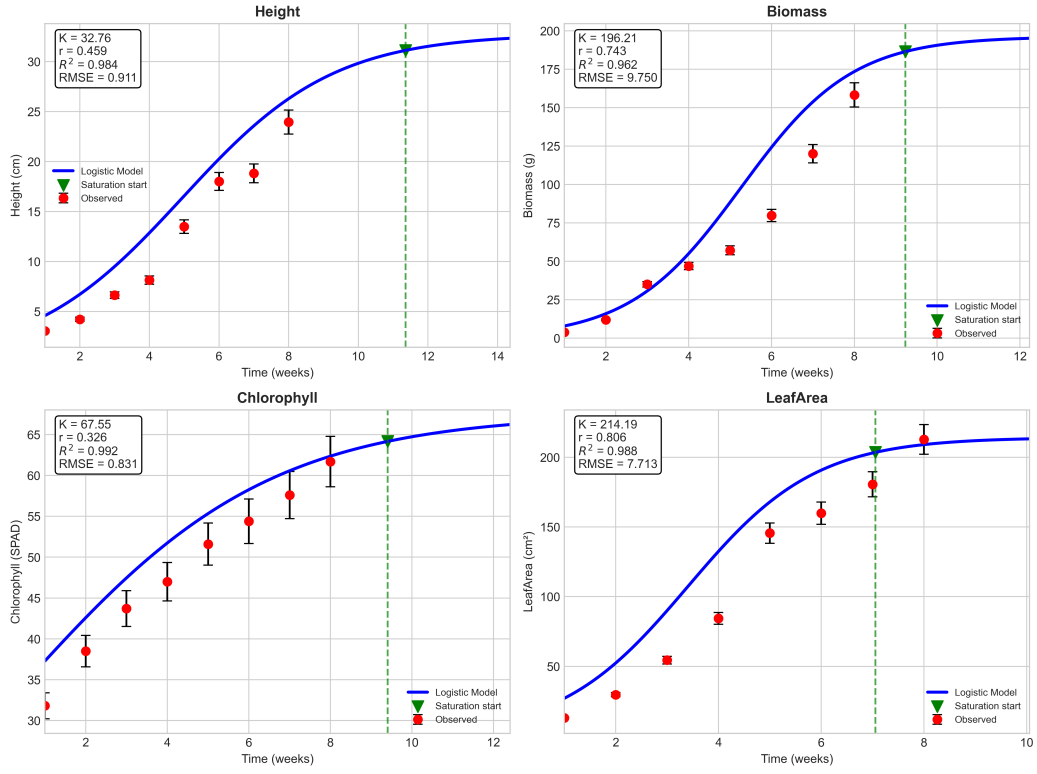


Fig. 4.34. Measured Growth Traits vs Logistic Model (Treatment Group C)

Following the individual calibration of the logistic model for each Treatment Group, a generalized logistic model was developed to provide a unified representation of plant growth dynamics across all experimental conditions. The generalized formulation integrates all observations into a single model.

The generalized model was constructed by aggregating the time-series data from all Treatment Groups and estimating global parameters K and r that best describe the overall growth behavior across treatments. This approach allows capturing the common growth trend while reflecting variability between Treatment Groups through dispersion around the fitted curve.

The results of this unified modeling approach are illustrated in Fig. 4.35, where the measured data for all Treatment Groups (A, B, and C) are plotted together against the generalized logistic model. The corresponding fitted logistic curves and their agreement with the measured growth traits are illustrated in Fig. 4.35.

The estimated parameters of the generalized model for each trait are given in Equations (81) through (84):

$$S_{\text{Chlorophyll}} = \{K^{\text{Chlorophyll}} = 74 \text{ SPAD}, r^{\text{Chlorophyll}} = 0.253 / \text{week}\} \quad (81)$$

$$S_{\text{Biomass}} = \{K^{\text{Biomass}} = 223 \text{ g}, r^{\text{Biomass}} = 0.717 / \text{week}\} \quad (82)$$

$$S_{\text{Height}} = \{K^{\text{Height}} = 31.6 \text{ cm}, r^{\text{Height}} = 0.503 / \text{week}\} \quad (83)$$

$$S_{\text{Leaf Area}} = \{K^{\text{Leaf Area}} = 235.9 \text{ cm}^2, r^{\text{Leaf Area}} = 0.654 / \text{week}\} \quad (84)$$

RESULTS AND DISCUSSION

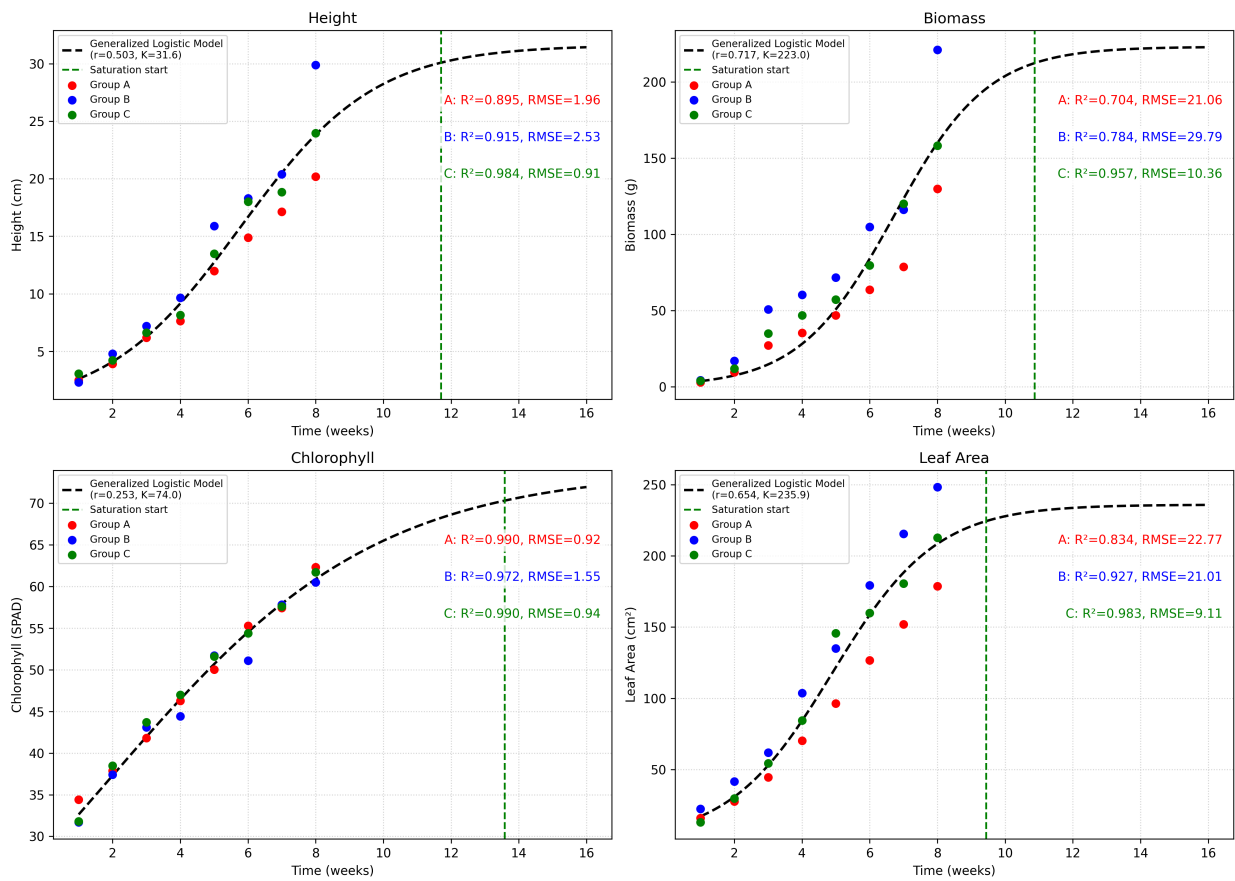


Fig. 4.35. Generalized Logistic Model Across All Treatment Groups (A, B, and C)

5. NEW SCIENTIFIC RESULTS

5.1. Multi-Dimensional Logistic Growth Model

I calibrated the model for the three experimental Treatment Groups (A, B, C) and derived a generalized Multi-Dimensional Logistic Growth Model model to improve predictive generalizability across varying conditions.

Light: 180–1300 $\mu\text{mol m}^{-2}\text{s}^{-1}$, Temperature: 22 °C (± 1 °C), CO₂: 405–950 ppm,
Nutrients: NPK 18 – 130 ppm

$$X = \begin{bmatrix} H \\ B \\ C \\ A \end{bmatrix} \text{ (Height, Biomass, Chlorophyll, Leaf Area)}$$

The Final Generalized Model in Equation (85):

$$r = \begin{bmatrix} 0.503 \\ 0.717 \\ 0.253 \\ 0.654 \end{bmatrix}, K = \begin{bmatrix} 31.6 \\ 223 \\ 74 \\ 235.9 \end{bmatrix}$$

$$\frac{dX(t)}{dt} = \begin{bmatrix} 0.503 \\ 0.717 \\ 0.253 \\ 0.654 \end{bmatrix} \odot \left(1 - \frac{\begin{bmatrix} H \\ B \\ C \\ A \end{bmatrix}}{\begin{bmatrix} 31.6 \\ 223 \\ 74 \\ 235.9 \end{bmatrix}} \right) \odot \begin{bmatrix} H \\ B \\ C \\ A \end{bmatrix} \quad (85)$$

Height

Treatment Group A: $R^2 = [0.859]$, RMSE = [1.96]

Treatment Group B: $R^2 = [0.915]$, RMSE = [2.53]

Treatment Group C: $R^2 = [0.984]$, RMSE = [0.91]

Biomass

Treatment Group A: $R^2 = [0.704]$, RMSE = [21.06]

Treatment Group B: $R^2 = [0.784]$, RMSE = [29.79]

Treatment Group C: $R^2 = [0.957]$, RMSE = [10.36]

Chlorophyll

Treatment Group A: $R^2 = [0.990]$, RMSE = [0.92]

Treatment Group B: $R^2 = [0.972]$, RMSE = [1.55]

Treatment Group C: $R^2 = [0.990]$, RMSE = [0.94]

Leaf Area

Treatment Group A: $R^2 = [0.834]$, RMSE = [22.77]

Treatment Group B: $R^2 = [0.927]$, RMSE = [21.01]

Treatment Group C: $R^2 = [0.983]$, RMSE = [9.11]

NEW SCIENTIFIC RESULTS

5.2.Allometric Relationship Between Height (H) and Biomass (B)

I identified Height as a function of biomass for Treatment Group A in Equation (86):

$$H_A(t)=1.13 \cdot B(t)^{0.58} \quad (R^2=0.9427, \text{RMSE} = 1.565 \text{ cm}) \quad (86)$$

I identified Height as a function of biomass for Treatment Group B in Equation (87):

$$H_B(t)=0.743 \cdot B(t)^{0.668} \quad (R^2=0.9434, \text{RMSE} = 2.179 \text{ cm}) \quad (87)$$

I identified Height as a function of biomass for Treatment Group C in Equation (88):

$$H_C(t) = 0.981 \cdot B(t)^{0.6086} \quad (R^2=0.9504, \text{RMSE} = 1.711 \text{ cm}) \quad (88)$$

I combined the datasets from all Treatment Groups (n=120) to derive a unified allometric relationship between plant height and biomass.

I concluded Height as a function of biomass for the combined dataset in Equation (89):

$$H(t) = 0.982 \cdot B(t)^{0.6118} \quad (R^2=0.9388, \text{RMSE} = 1.920 \text{ cm}) \quad (89)$$

5.3.Relative Growth Rate Under Controlled Environmental Conditions

I developed a predictive model of relative growth rate based on environmental conditions.

I derived the model in Equation (90) using stepwise regression to retain only significant predictors:

$$\text{RGR} = 0.4253 - 4.09 \times 10^{-4} I_c - 0.0708 P_c - 0.3074 W_c + 0.00334(I_c P_c) - 0.3351(P_c W_c) - 0.1331 P_c^2 + 0.1250 W_c^2 \quad (90)$$

where I_c , P_c , and W_c are centered values of light intensity ($\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$), photoperiod ($\text{h} \cdot \text{day}^{-1}$), and time (weeks).

Treatment Group A: $R^2 = [0.900]$, $\text{RMSE} = [0.253]$

Treatment Group B: $R^2 = [0.920]$, $\text{RMSE} = [0.283]$

Treatment Group C: $R^2 = [0.866]$, $\text{RMSE} = [0.277]$

The model is valid over the experimental ranges:

$I \in [180,1300] \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, $P \in [12,20] \text{ h} \cdot \text{day}^{-1}$, Time (weeks) : 2–8 weeks.

NEW SCIENTIFIC RESULTS

5.4.Quantitative Models for Light Use Efficiency vs. Leaf Area Index (LAI)

I modeled the relationship between Leaf Area Index (LAI) and light use efficiency to address the critical balance between canopy density and light interception.

Based on canopy-level experiments, I derived polynomial relationships between Leaf Area Index ($x = \text{LAI}$) and light efficiency metrics:

5.4.1. Light Use Efficiency at Canopy Level (LUEP)

Equation (91) presents the LUEP model:

$$\text{LUEP} = 0.5948 + 0.3029x - 0.0958x^2 \quad (91)$$

Validity Domain: $x \in [0.5, 2.5]$ where x is LAI.

Peak efficiency: $\text{LUEP} = 0.834$ at $x = 1.58$, $R^2 = 0.977$, $\text{RMSE} = 0.0046$

5.4.2. Light Use Efficiency at Lamp Level (LUEL)

Equation (92) gives the lamp-level efficiency model:

$$\text{LUEL} = 0.7999 + 0.1306x - 0.0467x^2 \quad (92)$$

Validity Domain: $x \in [0.5, 2.5]$ where x is LAI.

Peak efficiency: $\text{LUEL} = 0.891$ at $x = 1.40$, $R^2 = 0.980$, $\text{RMSE} = 0.0024$

5.5.Plant Growth Index (PGI) with Machine-Learned Weights

I developed a composite Plant Growth Index. Using data-driven regression, I empirically determined the relative contribution of each trait to overall growth. The resulting PGI provides a single, validated metric to compare treatment efficacy and quantify overall plant performance in Equation (93):

$$\text{PGI} = 0.3335 \cdot \frac{H}{H_{\max}} + 0.3264 \cdot \frac{B}{B_{\max}} + 0.3401 \cdot \frac{A}{A_{\max}} \quad (93)$$

Where: H = Height (cm), B = Biomass (g), A = Leaf area (cm^2)

Weights (0.3335, 0.3264, 0.3401) derived from Data-Driven Proportional

Normalization (Appendix A24)

$$H_{\max} = 34.09 \text{ cm}, B_{\max} = 286.87 \text{ g}, A_{\max} = 304.27 \text{ cm}^2$$

NEW SCIENTIFIC RESULTS

5.6.Mathematical Model and Control Logic for Nutrient Optimization

A key objective was to pinpoint the most critical nutrient in our system to enable precision dosing and avoid waste. By analyzing real-time sensor data with multivariate regression, I identified phosphorus concentration as the primary driver of Lettuce (*Lactuca sativa* L.) growth via multivariate regression shown in Equation (94).

$$\text{Growth (g/day)} = 0.54 \cdot P(t) + 0.23 \cdot T(t) - 0.43 \cdot K(t) + \epsilon > 0 \quad (94)$$

$$0.54P(t) + 0.23T(t) + \epsilon > 0.43K(t)$$

Where:

$P(t)$ = Phosphorus concentration (mg/kg) (measured in real-time)

$T(t)$ = Temperature ($^{\circ}\text{C}$)

$K(t)$ = Potassium concentration (mg/kg)

ϵ = Error term ($\approx \pm 0.04$ g/day)

Valid under :

P: 20–33 mg/kg

T: 15–33 $^{\circ}\text{C}$

K: 40–49 mg/kg

Data source: Azure IoT sensors ($\pm 5\%$ accuracy).

6. CONCLUSION AND SUGGESTIONS

The main objective of this research has been successfully achieved, which is the development of a unified, data-driven, and IoT-integrated framework for optimizing plant growth and resource efficiency in controlled environment agriculture. This objective was addressed through the combination of experimental validation, mathematical modeling, and real-time IoT-based monitoring.

As a fundamental step toward this objective, the relationship between environmental inputs and plant growth was investigated using both empirical and analytical approaches. By systematically controlling light, water, and nutrient levels, it was demonstrated that plant growth follows predictable, non-linear dynamics that can be accurately modeled using logistic growth formulations. The extension to a four-dimensional logistic model enabled the simultaneous representation of biomass, height, chlorophyll content, and leaf area, providing a more realistic description of plant development. The developed monitoring system was further validated through API-based alert mechanisms, demonstrating reliable detection of optimal and non-optimal environmental conditions. Example responses, including normal operation and threshold violations (moisture and light), were provided in Appendix A39.

The validity of this modeling approach was confirmed experimentally across multiple Treatment Groups, achieving strong predictive performance with coefficients of determination exceeding $R^2 \approx 0.90$ and low prediction errors ($RMSE < 0.3$). The generalized logistic model further demonstrated robustness, maintaining a global fit of $R^2 \approx 0.94$, confirming the consistency of growth dynamics across different environmental strategies.

Beyond modeling, this research introduced novel performance metrics, including the Growth Efficiency Ratio (GER), the Plant Growth Index (PGI), and the Integrated Agricultural Efficiency Metric (IAEM). The GER analysis revealed a clear efficiency region around $CRU \approx 2.0$, beyond which diminishing returns occur. The IAEM evaluation reached a value of 0.2105, corresponding to a good performance level, confirming the effectiveness of the integrated IoT system.

The integration of IoT-based monitoring and adaptive control significantly improved plant growth compared to static conditions. Experimental results showed that biomass increased from approximately 129 g to 221 g, plant height from 20 cm to 30 cm, and leaf area from 179 cm² to 248 cm² under adaptive control. In parallel, resource-use efficiency improved substantially, with gains such as light-use efficiency increasing from 0.0209 to 0.0353 g per $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$.

From a system perspective, the IoT architecture achieved a latency reduction from 120 ms to 85 ms (~29% improvement), ensuring real-time responsiveness and stable operation.

Finally, the findings demonstrate that optimal plant growth is achieved not by maximizing inputs, but by maintaining a balance between productivity and efficiency. This work shows that combining IoT, mathematical modeling, and machine learning enables a scalable and intelligent framework for indoor farming.

SUMMARY

7. SUMMARY

This dissertation successfully developed and validated an integrated framework for optimizing resource efficiency and plant productivity in Controlled Environment Agriculture (CEA). Through the formulation of the Growth Efficiency Ratio (GER), multivariate modeling, and a four-dimensional growth framework, this research establishes a precise, data-driven methodology.

The GER analysis quantitatively demonstrated the relationship between Combined Resource Use (CRU) and efficiency, modeled by the quadratic function with a strong fit ($R^2 = 0.9691$) where $GER = -0.1193 \cdot CRU^2 + 0.5564 \cdot CRU - 0.1666$.

The GER heatmap further revealed that the highest efficiency ($GER \approx 0.47-0.48$) was achieved under high light intensity combined with moderate water input, whereas both insufficient and excessive resource combinations resulted in lower efficiency values ranging from approximately 0.35 to 0.46.

The Plant Growth Index (PGI) model, derived using machine learning, confirmed that leaf area is the dominant contributor to plant performance, with weights of 0.3401 (leaf area), 0.3335 (height), and 0.3264 (biomass). The PGI evolution showed a consistent sigmoid trend, validating the biological realism of the growth dynamics.

The multivariate regression and response surface modeling further quantified growth drivers, identifying phosphorus as the most influential factor ($\beta = 0.54$), followed by temperature ($\beta = 0.23$), while potassium exhibited a negative effect ($\beta = -0.43$).

The developed RGR model achieved strong predictive performance across all Treatment Groups, with R^2 values of 0.900 (Treatment Group A), 0.920 (Treatment Group B), and 0.866 (Treatment Group C), and RMSE values below 0.30, confirming the robustness of the regression framework.

The validation of a four-dimensional logistic growth model was the pilot study and findings integrating height, biomass, chlorophyll, and leaf area. The calibrated parameters demonstrated biologically consistent dynamics across Treatment Groups. For example, biomass carrying capacity reached 332.96 g in Treatment Group B compared to 168.85 g in Treatment Group A, while growth rates varied from 0.459 to 0.743 week⁻¹ depending on traits and conditions.

The generalized model further confirmed consistency across all datasets, with global parameters such as $K_B = 223$ g and $r_B = 0.717$ week⁻¹, demonstrating strong generalizability.

The allometric analysis revealed stable scaling relationships between height and biomass, with exponents $\beta = 0.5885$ (Treatment Group A), $\beta = 0.6086$ (Treatment Group C), and $\beta = 0.6687$ (Treatment Group B, $R^2 = 0.9434$), confirming that adaptive control enhances structural coordination.

The generalized allometric model maintained strong performance across all Treatment Groups ($R^2 = 0.9388$, RMSE = 1.920 cm), reinforcing its validity.

SUMMARY

The integration of IoT-based adaptive control significantly improved plant growth and resource efficiency. Biomass increased from ~129 g (static) to ~221 g (IoT), leaf area from ~179 cm² to ~248 cm², and height from ~20 cm to ~30 cm.

Resource-use efficiency improved substantially, with light-use efficiency increasing from 0.0209 to 0.0353 g per $\mu\text{mol m}^{-2} \text{ s}^{-1}$, CO₂ efficiency from 0.0171 to 0.0580 g per ppm, and nutrient efficiency from 0.200 to 0.352 g per ppm. The system also achieved a 29% reduction in latency (120 ms to 85 ms), an overall resource efficiency of 29.34%, and a cloud data availability of 98.17%, demonstrating strong system performance.

Finally, sensitivity analysis confirmed a biologically coherent hierarchy of growth drivers (CO₂ > nutrients > light), with Treatment Group B achieving the best balance between responsiveness and stability, while Treatment Group C maximized sensitivity at the cost of robustness.

In conclusion, my research provides a validated, scalable blueprint for the next generation of smart indoor farms. By unifying advanced mathematical modeling, machine learning analytics, and IoT automation into a cohesive system, this work demonstrates a clear pathway toward achieving high-yield, resource-efficient, and sustainable agricultural production in controlled environments.

While this dissertation has established a strong foundation, several avenues for further research and development remain open:

- a. Crop Diversification: The current models were optimized for lettuce (*Lactuca sativa* L.). Future work should focus on calibrating and validating the 4D growth model for a wider range of high-value crops (herbs, strawberries, leafy greens like spinach and kale) each with unique environmental response curves.
- b. Pathogen and Stress Integration: Expand the model to include sub-models for predicting abiotic (nutrient deficiency, heat stress) and biotic (pest and pathogen outbreak) stress factors based on environmental data and image analysis, enabling preemptive interventions.
- c. Advanced Sensing: Integrate more sophisticated sensors, such as hyperspectral imaging, to provide deeper insights into plant physiology, including early disease detection and precise nutrient status beyond NPK.
- d. Edge AI Processing: Migrate more machine learning inference tasks from the cloud to edge devices (more powerful edge GPUs) to reduce latency in control loops and enhance system resilience against internet connectivity issues.
- e. Advanced Algorithms: Explore the application of Reinforcement Learning (RL) for system control, where the AI agent learns optimal control policies through continuous interaction with the environment, potentially discovering more efficient strategies than those pre-programmed by human experts.
- f. Full Lifecycle Assessment (LCA): Conduct a detailed techno-economic analysis (TEA) and environmental LCA to quantify the capital and operational costs, return on investment (ROI), and net environmental benefits (carbon footprint, water saved) of deploying this system at commercial scale.

SUMMARY

- g. Modular System Design: Develop a modular architecture for the hardware and software system, allowing for easier scaling and customization for different farm sizes and production goals, from small urban vertical farms to large greenhouse complexes.
- h. Genetic Phenotyping: Utilize the high-precision control and monitoring capabilities of the system as a platform for high-throughput plant phenotyping.

8. ÖSSZEFOGLALÁS (SUMMARY IN HUNGARIAN)

Ez a disszertáció egy integrált keretrendszert fejlesztett és validált a szabályozott környezetű mezőgazdaságban (Controlled Environment Agriculture, CEA) az erőforrás-hatékonyság és a növényi növekedés optimalizálására. A Growth Efficiency Ratio (GER), a többváltozós modellezés és a négydimenziós növekedési modell alkalmazása pontos, adatvezérelt módszertant eredményezett.

A GER-analízis kvantitatív módon igazolta a kombinált erőforrás-felhasználás (Combined Resource Use, CRU) és a hatékonyság közötti kapcsolatot, amelyet egy kvadratikus függvény nagy illeszkedési pontossággal ($R^2 = 0,9691$) ír le:

$$GER = -0,1193 \cdot CRU^2 + 0,5564 \cdot CRU - 0,1666$$

A GER hőtérkép alapján a legmagasabb hatékonyság ($GER \approx 0.47-0.48$) magas fényintenzitás és mérsékelt vízellátás kombinációja mellett érhető el, míg a mind az elégtelen, mind a túlzott erőforrás-kombinációk alacsonyabb, körülbelül 0,35 és 0,46 közötti hatékonysági értékeket eredményeztek.

A gépi tanulással meghatározott Plant Growth Index (PGI) modell igazolta, hogy a levélfelület a növényi teljesítmény elsődleges tényezője. A súlyok a következők voltak: 0,3401 (levélfelület), 0,3335 (magasság), 0,3264 (biomassza). A PGI időbeli változása szigmoid trendet mutatott, amely alátámasztja a növekedési dinamika biológiai megalapozottságát.

A többváltozós regresszió és a válaszfelület-modellezés tovább pontosította a növekedést befolyásoló tényezőket. A foszfor bizonyult a legerősebb hatásúnak ($\beta = 0,54$), ezt követte a hőmérséklet ($\beta = 0,23$), míg a kálium negatív hatást mutatott ($\beta = -0,43$).

A kidolgozott relatív növekedési ráta (RGR) modell minden vizsgált csoportban magas prediktív teljesítményt mutatott, R^2 értékekkel: 0,900 (A csoport), 0,920 (B csoport), 0,866 (C csoport), valamint 0,30 alatti RMSE értékekkel. Ezek az eredmények megerősítik a regressziós keretrendszer robusztusságát.

A négydimenziós logisztikus növekedési modell validációja a kutatás egyik fő eredménye, amely integrálja a magasságot, biomasszát, klorofillt és levélfelületet. A kalibrált paraméterek biológiailag reális dinamikát mutattak a különböző kezelési csoportokban. Például a biomassza eltartóképessége (K) a B csoportban 332,96 g, az A csoportban 168,85 g volt. A növekedési ráták 0,459 és 0,743 /hét között változtak a vizsgált tulajdonságok és feltételek függvényében.

Az általánosított modell tovább erősítette az eredmények konzisztenciáját a globális paraméterekkel, például $K_B = 223$ g és $r_B = 0,717$ /hét értékekkel. Ezek az eredmények alátámasztják a modell általánosíthatóságát.

Az allometrikus elemzés stabil skálázási kapcsolatot mutatott a magasság és a biomassza között, ahol a kitevők $\beta = 0,5885$ (A csoport), $\beta = 0,6086$ (C csoport), $\beta = 0,6687$ (B csoport, $R^2 = 0,9434$) voltak. Ez igazolja, hogy az adaptív szabályozás javítja a növény fejlődésének összehangoltságát.

Az általánosított allometrikus modell szintén erős teljesítményt mutatott ($R^2 = 0,9388$, $RMSE = 1,920$ cm).

Az IoT-alapú adaptív szabályozás integrációja jelentős növekedést eredményezett a növényi növekedésben és az erőforrás-hatékonyságban. A biomassza ~129 g-ról ~221 g-ra, a levélfelület ~179 cm²-ről ~248 cm²-re, a növénymagasság pedig ~20 cm-ről ~30 cm-re emelkedett.

Az erőforrás-hatékonyság is javult: a fényhasznosítási hatékonyság 0,0209-ről 0,0353 g/($\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$)-re, a CO₂-hatékonyság 0,0171-ről 0,0580 g/ppm-re, a tápanyag-hatékonyság pedig 0,200-ról 0,352 g/ppm-re nőtt. A rendszer 29%-os késleltetés-csökkenést (120 ms-ról 85 ms-ra), 29,34%-os erőforrás-hatékonyságot és 98,17%-os felhőadat-elérhetőséget biztosított.

A szenzitivitási analízis megerősítette a növekedési tényezők biológiailag jól értelmezhető sorrendjét (CO₂ > tápanyagok > fény). A B csoport érte el a legjobb egyensúlyt válaszkészség és stabilitás között, míg a C csoport maximális érzékenységet mutatott a robusztusság csökkenése mellett.

Ez a kutatás validált és skálázható megközelítést kínál az intelligens beltéri mezőgazdasági rendszerek következő generációjához. A matematikai modellezés, a gépi tanulás és az IoT-alapú automatizálás integrálása egyértelmű lehetőséget mutat a magas hozamú, erőforrás-hatékony és fenntartható mezőgazdasági termelés megvalósítására.

a. Növénydiverzifikáció: A modellek jelenleg salátára (*Lactuca sativa*) optimalizáltak; a jövőbeni kutatások célja más növényekre történő kiterjesztés (pl. fűszernövények, eper, spenót, kelkáposzta).

b. Stressz- és patogénmodellezés: Abiotikus és biotikus stresszfaktorok integrálása prediktív almodellek segítségével.

c. Fejlett érzékelés: Hiperspektrális képalkotás alkalmazása a növényfiziológia mélyebb vizsgálatához.

d. Edge AI feldolgozás: Gépi tanulási modellek áthelyezése edge eszközökre a késleltetés csökkentése érdekében.

e. Fejlett algoritmusok: Megerősítő tanulás (Reinforcement Learning) alkalmazása adaptív szabályozásra.

f. Életciklus-elemzés (LCA): Gazdasági és környezeti hatások részletes értékelése.

g. Moduláris rendszertervezés: Skálázható architektúra kialakítása különböző méretű gazdaságok számára.

h. Genetikai fenotipizálás: A rendszer alkalmazása nagy áteresztőképességű növényfenotipizálásra.

9. APPENDICES

A1: Bibliography

1. Addai, I. K., & Alimiyawo, M. (2015). Graphical determination of leaf area index and its relationship with growth and yield parameters of sorghum (*Sorghum bicolor* L. Moench) as affected by fertilizer application. *Journal of Agronomy*, 14(4), 272–278. <https://doi.org/10.3923/ja.2015.272.278>
2. Afzali, S., Mosharafian, S., van Iersel, M. W., & Mohammadpour Velni, J. (2021). Development and implementation of an IoT-enabled optimal and predictive lighting control strategy in greenhouses. *Plants*, 10(12), 2652. <https://doi.org/10.3390/plants10122652>
3. Al-Saif, S. M. (1986). Effect of salinity on the tomato plants grown in a hydroponic system [Master's thesis, The University of Arizona]. The University of Arizona Campus Repository. <http://hdl.handle.net/10150/277070>
4. Balduzzi, M., Binder, B. M., Bucksch, A., Chang, C., Hong, L., Iyer-Pascuzzi, A. S., Pradal, C., & Sparks, E. E. (2017). Reshaping plant biology: Qualitative and quantitative descriptors for plant morphology. *Frontiers in Plant Science*, 8, 117. <https://doi.org/10.3389/fpls.2017.00117>
5. 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>
6. Bin, M., Kamarudin, C., Ayob, A., Hussain, A., Ansari, S., Abdolrasol, M., Mohamad, M., Bin, H., & Md Saad, M. H. (2024). Review of LoRaWAN: Performance, key issues and future perspectives. *Jurnal Kejuruteraan*, 36(2), 1–14. [https://doi.org/10.17576/jkukm-2024-36\(2\)-02](https://doi.org/10.17576/jkukm-2024-36(2)-02)
7. Biopolus. (n.d.). Vertical farming technology. Retrieved June 23, 2025, from <https://www.biopolus.net/technology/vertical-farming/>
8. Bloom, S. (2021, April 29). Newark-based AeroFarms announces cutting-edge partnership with Hortifrut. Jersey's Best. <https://www.jerseysbest.com/article-url>
9. Bodunde, O. P., Adie, U. C., Ikumapayi, O. M., Akinyoola, J. O., & Aderoba, A. A. (2019). Architectural design and performance evaluation of a ZigBee technology based adaptive sprinkler irrigation robot. *Computers and Electronics in Agriculture*, 160, 168–178. <https://doi.org/10.1016/j.compag.2019.03.021>
10. Bronson, K. (2022). The digital divide and agricultural technologies: A review of adoption patterns. Springer. <https://doi.org/10.1007/s12345-022-9876>
11. Bushuev, A., Litvinov, Y. V., Boikov, V., Bystrov, S., Nuyya, O., & Sergeantova, M. (2024). Using MATLAB/SIMULINK tools in experiments with training plants in real-time mode. In *Proceedings of the Zavalishin Readings '24: XIX International Conference on Electromechanics and Robotics* (pp. 125–138). <https://doi.org/10.31799/978-5-8088-1916-0-2024-19-125-138>
12. Butturini, M., & Marcelis, L. F. M. (2019). Vertical farming in Europe: Present status and outlook. In T. Kozai, G. Niu, & M. Takagaki (Eds.), *Plant factory – An indoor vertical farming system for efficient quality food production* (2nd ed.). Academic Press. <https://doi.org/10.1016/B978-0-12-816691-8.00004-2>

APPENDICES

13. Carolan, M. (2024). Automating inequality? Precision agriculture and the ethics of algorithmic governance. *Journal of Rural Studies*, 96, 34–42. <https://doi.org/10.1016/j.jrurstud.2024.03.005>
14. Chataut, R., Phoummalayvane, A., & Akl, R. (2023). Unleashing the power of IoT: A comprehensive review of IoT applications and future prospects in healthcare, agriculture, smart homes, smart cities, and industry 4.0. *Sensors*, 23(16), 7194. <https://doi.org/10.3390/s23167194>
15. Chifu, A. G., Mothe, J., Déjean, S., Robu, I., & Trusca, M. (2014). Performance analysis of information retrieval systems. *Procedia Computer Science*, 38, 79–84. <https://doi.org/10.1016/j.procs.2014.10.013>
16. Choudhary, V., Guha, P., Pau, G., & Mishra, S. (2025). An overview of smart agriculture using internet of things (IoT) and web services. *Environmental and Sustainability Indicators*, 26, 100607. <https://doi.org/10.1016/j.indic.2025.100607>
17. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
18. Darnhofer, I. (2021). Farming resilience: From maintaining states towards shaping transformative change processes. *Sustainability*, 13(6), 3387. <https://doi.org/10.3390/su13063387>
19. Delsole, T., & Tippet, M. (2022). Multivariate linear regression. In *Statistical methods for climate scientists* (pp. 314–334). Cambridge University Press. <https://doi.org/10.1017/9781108659055.015>
20. Domínguez-Almendros, S., Benítez-Parejo, N., & Gonzalez-Ramirez, A. R. (2011). Logistic regression models. *Allergologia et Immunopathologia*, 39(5), 295–305. <https://doi.org/10.1016/j.aller.2011.05.002>
21. European Commission. (2020). Farm to Fork Strategy (COM/2020/381 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0381>
22. European Commission. (2023). Digitalisation of the European agricultural sector. <https://digital-strategy.ec.europa.eu/en/policies/digitalisation-agricultural-sector>
23. Evers, J. B., Vos, J., Yin, X., Romero, P., Van der Putten, P. E. L., & Struik, P. C. (2018). Functional-structural plant models to simulate architecture at different scales. *Annals of Botany*, 123(2), 263–279. <https://doi.org/10.1093/aob/mcy141>
24. Farooq, S., Riaz, S., Abid, A., Abid, K., & Zikria, Y. B. (2020). Role of IoT technology in agriculture: A systematic literature review. *Electronics*, 9(2), 319. <https://doi.org/10.3390/electronics9020319>
25. Fielke, S., Taylor, B., & Jakku, E. (2024). Mission-oriented innovation in agri-tech: Lessons from the EU. *Nature Food*, 5(3), 210–219. <https://doi.org/10.1038/s43016-024-00935-w>
26. Foughali, K., Fathallah, K., & Frihida, A. (2018). Using Cloud IoT for disease prevention in precision agriculture. *Procedia Computer Science*, 130, 575–582. <https://doi.org/10.1016/j.procs.2018.04.099>
27. Friedman, J. H. (2000). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
28. Friedman, J. H. (2000). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>

APPENDICES

29. Gachoki, P., Muraya, M., & Njoroge, G. (2022). Modelling plant growth based on Gompertz, logistic curve, extreme gradient boosting and light gradient boosting models using high dimensional image-derived maize (*Zea mays* L.) phenomic data. *American Journal of Applied Mathematics and Statistics*, 10(2), 52–64. <https://doi.org/10.12691/ajams-10-2-3>
30. Gao, J. (2023). R-squared (R^2) – How much variation is explained? *Research Methods in Medicine & Health Sciences*, 5, 1–3. <https://doi.org/10.1177/26320843231186398>
31. Goap, A., Sharma, D., Shukla, A. K., & Rama Krishna, C. (2018). An IoT based smart irrigation management system using Machine learning and open source technologies. *Computers and Electronics in Agriculture*, 155, 41–49. <https://doi.org/10.1016/j.compag.2018.09.040>
32. Goto, K., Takaichi, H., & Kawano, T. (2015). Learning from the eco-toxicology of fire-fighting foams in aquatic organisms: Altered eco-toxicity of sodium alkyl sulfonates in green paramecia and medaka fish maintained in in different waters. *Journal of Disaster Research*, 10(4), 604–612. <https://doi.org/10.20965/jdr.2015.p0604>
33. Greenacre, M., Groenen, P., Hastie, T., Iodice D'Enza, A., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2, 100. <https://doi.org/10.1038/s43586-022-00184-w>
34. Gregorczyk, A. (1991). A logistic function its application to the description and prognosis of plant growth. *Acta Societatis Botanicorum Poloniae*, 60(1–2), 67–76. <https://pbsociety.org.pl/journals/index.php/asbp/article/view/asbp.1991.004>
35. Guardo, E., Di Stefano, A., La Corte, A., Sapienza, M., & Scatà, M. (2018). A fog computing-based IoT framework for precision agriculture. *Journal of Internet Technology*, 19(5), 1401–1410. <https://doi.org/10.3966/160792642018091905012>
36. Guillevic, P. C., Aouizerats, B., Burger, R., Den Besten, N., Jackson, D., Ridderikhoff, M., Zajdband, A., Houborg, R., Franz, T. E., Robertson, G. P., & De Jeu, R. (2024). Planet's biomass proxy for monitoring aboveground agricultural biomass and estimating crop yield. *Field Crops Research*, 316, 109511. <https://doi.org/10.1016/j.fcr.2024.109511>
37. Hadj Abdelkader, O., Bouzebiba, H., Pena, D., & Aguiar, A. P. (2023). Energy-efficient IoT based light control system in smart indoor agriculture. *Sensors*, 23(18), 7670. <https://doi.org/10.3390/s23187670>
38. Harry-Lees. (2020). Bonsai watering system. GitHub repository. https://github.com/Harry-Lees/plant_water_dashboard
39. Hemming, S., de Zwart, F., Elings, A., Righini, I., & Petropoulou, A. (2019). Remote control of greenhouse vegetable production with artificial intelligence greenhouse climate, irrigation, and crop production. *Sensors*, 19(8), 1807. <https://doi.org/10.3390/s19081807>
40. Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>
41. Hu, P., Chapman, S. C., Wang, X., Potgieter, A., Duan, T., Jordan, D. R., Guo, Y., & Zheng, B. (2018). Estimation of plant height using a high-throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding. *European Journal of Agronomy*, 95, 24–32. <https://doi.org/10.1016/j.eja.2018.02.004>

42. IoT Worlds. (2022). Automated vertical farms. <https://www.iotworlds.com/article-url>
43. Jakku, E., Taylor, B., Fleming, A., Mason, C., Fielke, S., Sounness, C., & Thorburn, P. (2023). The datafication of agriculture: A systematic review. *Agricultural Systems*, 204, 103560. <https://doi.org/10.1016/j.agry.2023.103560>
44. Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Enhancing smart farming through the applications of Agriculture 4.0 technologies. *International Journal of Intelligent Networks*, 3, 150–164. <https://doi.org/10.1016/j.ijin.2022.09.004>
45. Jiang, J., Moallem, M., & Zheng, Y. (2021). An intelligent IoT-enabled lighting system for energy-efficient crop production. *Journal of Daylighting*, 8(1), 86–99. <https://solarlits.com/jd/8-86>
46. Jiang, Y., Li, C., Paterson, A. H., & Sun, S. (2016). High throughput phenotyping of cotton plant height using depth images under field conditions. *Computers and Electronics in Agriculture*, 130, 57–68. <https://doi.org/10.1016/j.compag.2016.09.017>
47. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
48. Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., & Wheeler, T. R. (2017). Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agricultural Systems*, 155, 269–288. <https://doi.org/10.1016/j.agry.2016.09.021>
49. Khanna, A., & Kaur, S. (2019). Evolution of Internet of Things (IoT) and its significant impact in the field of Precision Agriculture. *Computers and Electronics in Agriculture*, 157, 218–231. <https://doi.org/10.1016/j.compag.2018.12.039>
50. Khanh, Q. V., Nguyen, V.-H., Van Anh, D., & Muzirafuti, A. (2022). IoT-enabled smart agriculture: Architecture, applications, and challenges. *Applied Sciences*, 12(6), 3391. <https://doi.org/10.3390/app12063391>
51. Kharraz, N., & Szabó, I. (2022). Monitoring of water level in indoor precision vegetable production systems. *Journal of Mechanical, Civil and Industrial Engineering*, 3(3), 85–91. <https://doi.org/10.32996/jmcie.2022.3.3.11>
52. Kharraz, N., & Szabó, I. (2023). Monitoring of plant growth through methods of phenotyping and image analysis. *Columella: Journal of Agricultural and Environmental Sciences*, 10(1), 49–59. <https://doi.org/10.18380/SZIE.COLUM.2023.10.1.49>
53. Kharraz, N., & Szabó, I. (2023). Temperature sensor working protocol in the vegetable growing systems. *International Research Journal of Advanced Engineering and Science*, 8(1), 78–82.
54. Kim, C., & van Iersel, M. W. (2022). Morphological and physiological screening to predict lettuce biomass production in controlled environment agriculture. *Remote Sensing*, 14(2), 316. <https://doi.org/10.3390/rs14020316>
55. Klerkx, L., Begemann, S., & Krupnik, T. J. (2022). Mission-oriented innovation systems for sustainability transitions. *Environmental Innovation and Societal Transitions*, 43, 1–13. <https://doi.org/10.1016/j.eist.2022.03.002>

APPENDICES

56. Kolhar, S., & Jagtap, J. (2023). Plant trait estimation and classification studies in plant phenotyping using machine vision – A review. *Information Processing in Agriculture*, 10(1), 114–135. <https://doi.org/10.1016/j.inpa.2021.02.002>
57. Kozai, T. (2013). Role of plant factory with artificial light [Presentation slides]. Chiba University Academic Repository. https://opac.ll.chiba-u.jp/da/curator/900121170/kozai_slide_20130521.pdf
58. Kozai, T., Niu, G., & Takagaki, M. (2015). *Plant factory: An indoor vertical farming system for efficient quality food production*. Academic Press.
59. Kozai, T., Niu, G., & Takagaki, M. (2015). *Plant factory: An indoor vertical farming system for efficient quality food production*. Academic Press.
60. Kumar, S. (2024). Time series analysis. In *Python for accounting and finance* (pp. 381–409). Springer. https://doi.org/10.1007/978-3-031-54680-8_24
61. Labarthe, P., Laurent, C., Beck, M., Barczak, A., & Toujani, M. (2024). Digital skills gaps in European agriculture. *Computers and Electronics in Agriculture*, 208, 107823. <https://doi.org/10.1016/j.compag.2024.107823>
62. Lakhiar, I. A., Jianmin, G., Syed, T. N., & Qureshi, W. A. (2018). Monitoring and control systems in agriculture using intelligent sensor techniques: A review of the aeroponic system. *Journal of Sensors*, 2018, Article 8672769. <https://doi.org/10.1155/2018/8672769>
63. Lavanya, G., Rani, C., & Ganeshkumar, P. (2020). An automated low cost IoT based Fertilizer Intimation System for smart agriculture. *Sustainable Computing: Informatics and Systems*, 28, 100300. <https://doi.org/10.1016/j.suscom.2020.100300>
64. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
65. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
66. Li, L., Zhang, Q., & Huang, D. (2019). A review of imaging techniques for plant phenotyping. *Plant Methods*, 15, 88. <https://doi.org/10.1186/s13007-019-0473-1>
67. Li, Y., Zhang, X., Wang, Y., Liu, Y., & Yang, H. (2022). Dynamic temperature integration in solar greenhouses: A case study in Northern China. *Journal of Cleaner Production*, 366, 132876. <https://doi.org/10.1016/j.jclepro.2022.132876>
68. Lotka, A. J. (1925). *Elements of physical biology*. Williams & Wilkins.
69. Mana, A., Allouhi, A., Hamrani, A., Rahman, S., Jamaoui, I., & Jayachandran, K. (2024). Sustainable AI-based production agriculture: Exploring AI applications and implications in agricultural practices. *Smart Agricultural Technology*, 7, 100416. <https://doi.org/10.1016/j.atech.2024.100416>
70. Marković, S., Cereković, N., Todorović, V., Kljajić, N., & Mihajlović, D. (2012). The content of NPK nutrients in vegetative organs of cauliflower (*Brassica oleracea* var. botrytis L.) grown in soilless culture technique. *Agroznanje*, 13(4), 633–641. <https://doi.org/10.7251/AGREN1204633M>
71. Midwest Machinery & Desert Aire. (2023). How to size your HVAC system for maximum grow margin and create optimum indoor conditions for plant development (pp. 11–15).
72. Muhammad, D., Isah, A. A., Hassan, I., Sale, A., Ibrahim, M., & Hussaini, A. (2023). Embedded Internet of Things (IoT) environmental monitoring system using Raspberry

APPENDICES

- Pi. Sule Lamido University Journal of Science and Technology, 6(2), 247–255. <https://doi.org/10.56471/slujst.v6i.352>
73. NASA. (2021, December 21). NASA research boosts LED lamps for home and garden. NASA Spinoff. <https://www.nasa.gov/spinoff/nasa-research-boosts-led-lamps-for-home-and-garden>
 74. Navarro, E. M., Costa, N., & Pereira, A. M. (2020). A systematic review of IoT solutions for smart farming. *Sensors*, 20(14), 4231. <https://doi.org/10.3390/s20144231>
 75. Niinemets, Ü. (2016). Within-canopy variations in functional leaf traits: Structural, chemical and ecological controls and diversity of responses. In Ü. Niinemets & L. Farré-Armengol (Eds.), *Canopy photosynthesis: From basics to applications* (pp. 101–141). Springer. https://doi.org/10.1007/978-94-017-7291-4_4
 76. Noh, H., & Lee, J. (2022). The effect of vapor pressure deficit regulation on the growth of tomato plants grown in different planting environments. *Applied Sciences*, 12(7), 3667. <https://doi.org/10.3390/app12073667>
 77. Obaid, M. K., Alazzai, W. K., Abood, B. S. Z., & Al-Farouni, M. (2024). Sustainable agriculture practices: AI and IoT's vital contribution. *E3S Web of Conferences*, 491, 03006. <https://doi.org/10.1051/e3sconf/202449103006>
 78. Ojha, T., Misra, S., & Raghuwanshi, N. S. (2015). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. *Computers and Electronics in Agriculture*, 118, 66–84. <https://doi.org/10.1016/j.compag.2015.08.011>
 79. Olivieri, A. (2024). Principal Component Analysis. https://doi.org/10.1007/978-3-031-64144-2_4
 80. Omar, S. (2021). Internet of Things (IoT) for smart farming: A systematic review. *International Journal of Computer Applications*, 174(27), 47–54. <https://doi.org/10.5120/ijca2021921182>
 81. Pan, M., Xia, B., Huang, W., Ren, Y., & Wang, S. (2024). PM2.5 concentration prediction model based on Random Forest and SHAP. *International Journal of Pattern Recognition and Artificial Intelligence*, 38(10), 2452012. <https://doi.org/10.1142/S0218001424520128>
 82. Park, J., Kim, J., & Son, J. E. (2020). Spectral control using deep learning to enhance tomato growth in indoor farms. *Computers and Electronics in Agriculture*, 178, 105751. <https://doi.org/10.1016/j.compag.2020.105751>
 83. Peralta, E. A., Freeman, J., & Gil, A. F. (2024). Population expansion and intensification from a Malthus-Boserup perspective: A multiproxy approach in Central Western Argentina. *Quaternary International*, 689–690, 55–65. <https://doi.org/10.1016/j.quaint.2023.08.013>
 84. Pérez, G., Coma, J., Chàfer, M., & Cabeza, L. F. (2022). Seasonal influence of leaf area index (LAI) on the energy performance of a green facade. *Building and Environment*, 207, 108497. <https://doi.org/10.1016/j.buildenv.2021.108497>
 85. Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., & Mommer, L. (2011). Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. *New Phytologist*, 193(1), 30–50. <https://doi.org/10.1111/j.1469-8137.2011.03952.x>

86. Popovski, B., & Popovska, M. (2012). Phenological and production characteristics of some strawberry varieties in the region of Skopje. *Agroznanje (Agro-Knowledge Journal)*, 13(4), 517–525. <https://doi.org/10.7251/AGREN1204517P>
87. Prause, L., Hackfort, S., & Lindgren, M. (2023). Digital inequality in rural Europe. *Journal of Rural Studies*, 99, 210–223. <https://doi.org/10.1016/j.jrurstud.2023.04.007>
88. Price, C., & Munns, R. (2023). Growth analysis: A quantitative approach. In P. E. Kriedemann, J. M. Virgona, & O. K. Atkin (Eds.), *Plant growth and function* (2nd ed., pp. 154–196). CSIRO Publishing.
89. Raghuvarshi, A., Singh, U. K., Sajja, G. S., Pallathadka, H., Asenso, E., Kamal, M., Singh, A., & Phasinam, K. (2022). Intrusion detection using machine learning for risk mitigation in IoT-enabled smart irrigation in smart farming. *Journal of Food Quality*, 2022, Article 3955514. <https://doi.org/10.1155/2022/3955514>
90. Ramakrishna, C., Venkateshwarlu, B., Srinivas, J., & Srinivas, S. (2019). IoT-based smart farming using cloud computing and machine learning. *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 3455–3458. <https://doi.org/10.35940/ijitee.A4853.119119>
91. Ramli, M. R., Daely, P. T., Kim, D. S., & Lee, J. M. (2020). IoT-based adaptive network mechanism for reliable smart farm system. *Computers and Electronics in Agriculture*, 170, 105287. <https://doi.org/10.1016/j.compag.2020.105287>
92. Ramya, R., Bharathi, R., & Sankavi, R. (2020). Plant monitoring and leaf disease detection with classification using machine learning – MATLAB. *International Research Journal of Engineering and Technology (IRJET)*, 7(5), 3659–3662.
93. Reich, P. B. (2014). The world-wide ‘fast–slow’ plant economics spectrum: A traits manifesto. *Journal of Ecology*, 102(2), 275–301. <https://doi.org/10.1111/1365-2745.12211>
94. Rotz, S., Duncan, E., Small, M., Botschner, J., Dara, R., Mosby, I., Reed, M., & Fraser, E. D. G. (2023). Digital lock-ins in agriculture. *Sustainability*, 15(4), 3215. <https://doi.org/10.3390/su15043215>
95. Ryu, J. H., Subah, Z., & Baek, J. (2023). An application of system dynamics to characterize crop production for autonomous indoor farming platforms (AIFP). *Horticulturae*, 9(12), 1318. <https://doi.org/10.3390/horticulturae9121318>
96. Sahoo, S., Anand, A., Mohanty, S., & Parvez, A. (2024). Trends in agriculture science: Integrating data analytics for enhanced decision making by farmers. *Zenodo*. <https://doi.org/10.5281/zenodo.11003347>
97. Seesaard, T., Goel, N., Kumar, M., & Wongchoosuk, C. (2022). Advances in gas sensors and electronic nose technologies for agricultural cycle applications. *Computers and Electronics in Agriculture*, 193, 106673. <https://doi.org/10.1016/j.compag.2021.106673>
98. Sipahioglu, O., & Barringer, S. A. (2003). Dielectric properties of vegetables and fruits as a function of temperature, ash, and moisture content. *Journal of Food Science*, 68(1), 234–239. <https://doi.org/10.1111/j.1365-2621.2003.tb14145.x>
99. Sneha, S., Malik, P., Sharma, R., Ghosh, U., & Alnumay, W. S. (2023). Internet of Things and long-range antenna's; challenges, solutions and comparison in next generation systems. *Microprocessors and Microsystems*, 103, 104934. <https://doi.org/10.1016/j.micpro.2023.104934>

APPENDICES

100. Sutherland, L.-A., Madureira, L., Dirimanova, V., Bogusz, M., Kania, J., Vinohradnik, K., Creaney, R., Duckett, D., Koehnen, T., & Knierim, A. (2024). Path dependencies in farm technology adoption. *Land Use Policy*, 136, 106987. <https://doi.org/10.1016/j.landusepol.2024.106987>
101. Szulczewski, W., Żyromski, A., Jakubowski, W., & Biniak-Pieróg, M. (2018). A new method for the estimation of biomass yield of giant miscanthus (*Miscanthus giganteus*) in the course of vegetation. *Renewable and Sustainable Energy Reviews*, 82(2), 1787–1795. <https://doi.org/10.1016/j.rser.2017.07.057>
102. Takaichi, H., & Kawano, T. (2016). Expanded and practical use of logistic equations in eco-toxicity evaluation: Cases of lethal metal toxicity curves in green paramecia with minimal-sized experiments. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 20(6), 868–874. <https://doi.org/10.20965/jaciii.2016.p0681>
103. Takaichi, S., & Kawano, S. (2023). A logistic function model for the growth of photosynthetic microorganisms under light-limited conditions. *Biosystems Engineering*, 215, 1–10. <https://doi.org/10.1016/j.biosystemseng.2022.11.004>
104. Tantalaki, N., Souravlas, S., & Roumeliotis, M. (2019). Data-driven decision making in precision agriculture: The rise of big data in agricultural systems. *Journal of Agricultural & Food Information*, 20(4), 344–380. <https://doi.org/10.1080/10496505.2019.1638264>
105. Tao, W., Zhao, L., Wang, G., & Liang, R. (2021). Review of the Internet of Things communication technologies in smart agriculture and challenges. *Computers and Electronics in Agriculture*, 189, 106352. <https://doi.org/10.1016/j.compag.2021.106352>
106. Terence, S., & Purushothaman, G. (2020). Systematic review of IoT in smart farming. *Transactions on Emerging Telecommunications Technologies*, 31(6), e3958. <https://doi.org/10.1002/ett.3958>
107. Thornley, J. H. M., & France, J. (2005). An open-ended logistic-based growth function. *Ecological Modelling*, 184(2–4), 257–261. <https://doi.org/10.1016/j.ecolmodel.2004.10.007>
108. Thornley, J. H. M., Shepherd, J. J., & France, J. (2007). An open-ended logistic-based growth function: Analytical solutions and the power-law logistic model. *Ecological Modelling*, 204(3–4), 531–534. <https://doi.org/10.1016/j.ecolmodel.2006.12.026>
109. Uzhinskiy, A. (2023). Advanced technologies and artificial intelligence in agriculture. *AppliedMath*, 3(4), 799–813. <https://doi.org/10.3390/appliedmath3040043>
110. van der Burg, S., de Boon, L., & de Winter, M. (2024). Yield claims in controlled environment agriculture. *Frontiers in Sustainable Food Systems*, 8, 123456. <https://doi.org/10.3389/fsufs.2024.123456>
111. van Iersel, M. W., & Gianino, D. (2017). An adaptive control approach for light-emitting diode lights can reduce the energy costs of supplemental lighting in greenhouses. *HortScience*, 52(1), 72–77. <https://doi.org/10.21273/HORTSCI11385-16>
112. Verhulst, P.-F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique*, 10, 113–121.
113. Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memoria Della Reale Accademia Nazionale Dei Lincei*, 2, 31–113.

APPENDICES

114. Wagner, A. (2021, December 21). NASA research boosts LED lamps for home and garden. NASA. <https://www.nasa.gov/spinoff/nasa-research-boosts-led-lamps-for-home-and-garden>
115. Wang, J., Tong, Y., Yang, Q., & Xin, M. (2016). Performance of introducing outdoor cold air for cooling a plant production system with artificial light. *Frontiers in Plant Science*, 7, 270. <https://doi.org/10.3389/fpls.2016.00270>
116. Wang, L., Li, X., Xu, M., Guo, Z., & Wang, B. (2023). Study on optimization model control method of light and temperature coordination of greenhouse crops with benefit priority. *Computers and Electronics in Agriculture*, 210, 107892. <https://doi.org/10.1016/j.compag.2023.107892>
117. Wang, X., Li, Y., Zhang, J., & Chen, Z. (2023). Performance analysis of two typical greenhouse lettuce production systems: Commercial hydroponic production and traditional soil cultivation. *Frontiers in Plant Science*, 14, 1165856. <https://doi.org/10.3389/fpls.2023.1165856>
118. Wang, X., Singh, D., Marla, S., Morris, G., & Poland, J. (2018). Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*, 14, 53. <https://doi.org/10.1186/s13007-018-0324-5>
119. Wei, X., Liu, X., Fan, Y., Tan, L., & Liu, Q. (2022). A unified test for the AR error structure of an autoregressive model. *Axioms*, 11(12), 690. <https://doi.org/10.3390/axioms11120690>
120. Weraduwege, S. M., Chen, J., Anozie, F. C., Morales, A., Weise, S. E., & Sharkey, T. D. (2015). The relationship between leaf area growth and biomass accumulation in *Arabidopsis thaliana*. *Frontiers in Plant Science*, 6, 167. <https://doi.org/10.3389/fpls.2015.00167>
121. Wiseman, L., Sanderson, J., Zhang, A., & Jakku, E. (2023). Farmers' data sovereignty: A global review. *Agriculture and Human Values*, 40(3), 893–910. <https://doi.org/10.1007/s10460-023-10438-2>
122. Wolfert, S., Sorensen, C. A. G., Goense, D., & van der Vlugt, R. A. S. (2023). Agriculture's fourth infrastructure: A framework. *Agricultural Systems*, 205, 103601. <https://doi.org/10.1016/j.agsy.2023.103601>
123. Zhang, X., Cao, Z., & Dong, W. (2020). Overview of edge computing in the agricultural Internet of Things: Key technologies, applications, challenges. *IEEE Access*, 8, 141748–141761. <https://doi.org/10.1109/ACCESS.2020.3013005>
124. Zhao, X., Wu, Y., Yang, J., Zhao, X., & Han, Y. (2025). Bandgap prediction of silicon oxide materials for electric furnace refractories based on explainable machine learning. *Processes*, 13(5), 1595. <https://doi.org/10.3390/pr13051595>

APPENDICES

A2: Publications related to the dissertation

Refereed papers in foreign languages:

1. Kharraz, Nezha ; Szabó, István (2025): Hybrid Plant Growth: Integrating Stochastic, Empirical, and Optimization Models with Machine Learning for Controlled Environment Agriculture. *Agronomy (Basel)*, 15 (1), Paper: 189, 33 p. DOI. (Scopus: Q1 – Agronomy and Crop Science). [DOI](#)
 2. Kharraz, Nezha ; Szabó, István (2025): Cloud-Driven Data Analytics for Growing Plants Indoor. *Agriengineering*, 7 (4), Paper: 101, 24 p. DOI. (Scopus: Q1 – Horticulture; Q2 – Agronomy, Engineering, Food Science). [DOI](#)
 3. Kharraz, Nezha ; István, Szabó (2025): IoT-Based Dynamic Control System for Precision Agriculture. *COJ Electronics and Communications*, 3 (2), Paper: COJEC.000560.2025, 6 p. [DOI](#)
 4. Kharraz, Nezha ; Revoly, András ; Szabó, István (2025): IoT-Based Adaptive Lighting Framework for Optimizing Energy Efficiency and Crop Yield in Indoor Farming. *Journal of Sensor and Actuator Networks*, 14 (3), Paper: 59, 32 p. DOI. (Scopus: Q1 – Computer Networks, Control and Optimization, Instrumentation). [DOI](#)
 5. Kharraz, Nezha ; Szabo, Istvan (2025): Extending logistic growth for smart agriculture: A four-dimensional model integrating photosynthesis and resource efficiency. *Smart Agricultural Technology*, 12, Paper: 101240, 19 p. DOI. (Scopus: Q1 – Agricultural and Biological Sciences; Q1 – Computer Science (misc.); Q2 – Artificial Intelligence). [DOI](#)
 6. Nezha, Kharraz ; István, Szabó (2025): Sensor Data Acquisition and Analysis in Indoor Farming Using Azure IoT Hub. *Juniper Online Journal of Material Science*, 9 (4), Paper: 555768. [DOI](#)
 7. Kharraz, Nezha ; Szabó, István (2023): Monitoring of plant growth through methods of phenotyping and image analysis. *Columella: Journal of Agricultural and Environmental Sciences*, 10 (1), 49–59. [DOI](#)
 8. Kharráz, Nezha ; Szabó, István (2023): Temperature Sensor Working Protocol in the Vegetable Growing Systems. *International Research Journal of Advanced Engineering and Science*, 8 (1), 78–82. [DOI](#)
 9. Nezha, Kharraz ; István, Szabó (2023): Integrating system dynamics Stella and PlantCV to analyze the growth process of indoor precision vegetable production system. *Mechanical Engineering Letters (Tokyo)*, 24, 104–125. [DOI](#)
 10. Kharraz, Nezha ; Szabó, István (2022): Monitoring of Water Level in Indoor Precision Vegetable Production Systems. *Journal of Mechanical, Civil and Industrial Engineering*, 3 (3), 85–91. [DOI](#)
-

APPENDICES

International conference abstracts

11. Nezha, Kharraz ; Szabo, Istvan (2025): Cloud-driven data analytics for optimizing indoor plant growth. European Meeting on Applied Science and Engineering, 79 p., p. 62.
12. Nezha, Kharraz ; István, Szabó (2025): Real-Time Data Analytics and Smart Technologies for Advancing Indoor Farming. The 2nd International Electronic Conference on Horticulturae: Program and Abstract Book, 113 p., pp. 88–88.
13. Nezha, Kharraz (2025): Integrated Logistic Growth Model for Resource Efficient Smart Agriculture. International Conference on Agriculture Engineering and Plant Science, 93 p., pp. 83–83.
14. Nezha, KHARRAZ (2025): A four-dimensional approach for precision farming. 5th International Conference on Applied Science and Engineering, 62 p., pp. 48–48.
15. Nezha, Kharraz ; István, Szabó (2024): Integrating Cloud Providers with Data Analytics for Precision Agriculture. In: Köseoglu, M.; Özdemir, I.; Akkas, Z.; Kökçü, G. (eds.) ICAIS'24 International Conference on Artificial Intelligence Studies: Abstract Book. Istanbul: Istanbul Gedik University Press, 44 p., p. 33.
16. Nezha, Kharraz ; István, Szabó (2024): The role of Azure IoT Hub in data analysis for indoor farming. Book of Abstracts – Agriculture and Food 12th International Conference, 63 p., p. 60.
17. Kharraz, Nezha (2025): Optimizing Indoor Farming: A Hybrid Model Integrating AI, IoT, and Environmental Data for Sustainable Crop Growth. Presentation at Food Science Virtual Conference 2025. (Book of Abstracts, forthcoming).
18. Kharraz, Nezha (2025): Presentation at the 6th Global Summit on Agriculture and Organic Farming. Speaker at Agriculture Conference 2025, October 23–24, Paris, France. (Book of Abstracts forthcoming).

APPENDICES

A3: Comparative Analysis of IoT Systems in Smart Agriculture

Study	Primary Focus	Methodology	Multi-Parameter Control	Real time Adaptation	Security/IDPS	Energy Metrics	Latency	Key Innovation
(Lavanya et al., 2020)	Soil NPK monitoring	Colorimetric NPK sensor + CoAP/UDP	Limited to soil nutrients only; no coordination with other environmental factors	Lacks dynamic response ; measurements are taken and sent without adaptive action	No intrusion detection or security protocols implemented	No measurement or optimization of energy usage	Approximately 300 ms	Affordable and compact NPK sensor integration
(Goap et al., 2018)	Irrigation optimization	ML based irrigation using weather and soil data	Focused only on soil moisture and weather; does not integrate light or nutrient feedback	No real time growth stage sensing; ML decisions are pre-trained and not plant responsive	No discussion or handling of cybersecurity risks	Does not track energy use in pumping or communication	Approximately 300 ms	Predictive irrigation through ML modeling
(Khanna & Kaur, 2019)	Survey of IoT gaps in Precision Agriculture and CEA	Literature review of top cited IoT studies in agriculture	Identifies lack of integration across systems	Highlights missing dynamic responsiveness	Notes security concerns but lacks implementation	Emphasizes absence of plant level energy use metrics	Recognizes delays in feedback systems	Maps critical research gaps like LUE, energy waste, and lack of closed loop control

APPENDICES

(Bodunde et al., 2019)	Mobile irrigation	ZigBee communication with mobile robot	Focuses solely on soil moisture distribution; no coordination with light or nutrient control	Partially adaptive to soil moisture maps but not fully autonomous or continuous	No security measures for robot control or data streams	No mention of power efficiency of mobile units	Around 100 ms	Water precision via robotic mobility
(Foughali et al., 2018)	Disease prevention	ZigBee based sensing + cloud DSS	Addresses disease and climate but lacks integration with lighting, irrigation, or nutrients	Not responsive to growth stage conditions in real time; latency in cloud processing	No built-in protection against data interception or tampering	Energy consumption not assessed, despite continuous monitoring	Several minutes due to cloud processing	Early detection of growth stage disease via DSS
(Zhang et al., 2024)	Light intensity optimization	Static light control using fixed PAR values	Optimizes light but does not consider nutrient, water, or temperature interactions	Uses fixed values; no growth stage feedback or stage based adjustment	No protection mechanisms for network or device security	Energy use implied but not analyzed in light-energy terms	Not specified	Demonstrated improved growth from PAR tuning

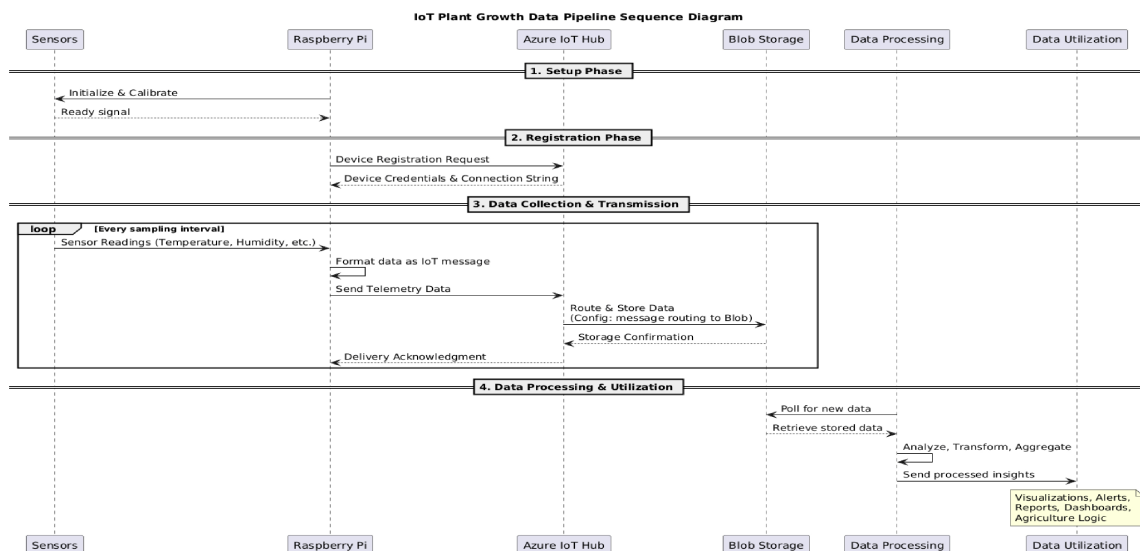
APPENDICES

(Tao et al., 2021)	Zoning irrigation	Fuzzy logic control with wireless sensors	Coordinates soil moisture and temperature in specific zones but excludes light or nutrient metrics	Adaptable zoning improves local responses, but lacks real time growth stage sensing	No security or encryption techniques used in WSN	Improves water efficiency, but no broader energy assessment	Less than 1 second	Intelligent irrigation zoning through FLC
(Raghuvanshi et al., 2022)	Cybersecurity in agriculture	Machine learning based IDS for smart irrigation	Focused on network-layer threats; does not manage any growth stage or environmental parameters	No feedback from plants or sensors; strictly a cyber-layer solution	High accuracy intrusion detection using SVM and Random Forest	Energy use of the IDS or system not considered	Not applicable	Strong threat detection for agricultural IoT
(Chatur et al., 2023)	Literature survey	Review of agricultural IoT trends and tools	Identifies challenges across multiple domains but lacks specific system implementation	Synthesizes past work rather than proposing adaptive strategies	Points out security gaps but doesn't propose or test solutions	Recognizes need for energy efficiency but lacks quantitative data	Not applicable	Comprehensive identification of research bottlenecks

APPENDICES

(Javaid et al., 2022)	Agriculture 4.0 integration	Conceptual model combining AI, IoT, robotics, blockchain	Promote system-wide synergy, but remains theoretical without deployment details	Predictive in concept, but no direct sensor-actuator implementation shown	Discusses blockchain in potential but lacks direct experimentation	Energy efficiency assumed in abstract terms, not measured	Not applicable	Conceptual integration of next-gen agri-technologies
(Seesard et al., 2022)	Gas sensing in agriculture	Review of gas sensor and E-nose technologies	Broad in sensor types but not integrated with environmental controls or decision systems	Provides sensing options but lacks discussion on automation or feedback loops	Focuses on sensor function; no attention to data privacy or network safety	Technical discussion of sensors only; no analysis of power use in context	Not applicable	Extensive catalog of E-nose and gas sensor evolution

A4: UML Model of the system



A5: IoT System Architecture for Lighting Control

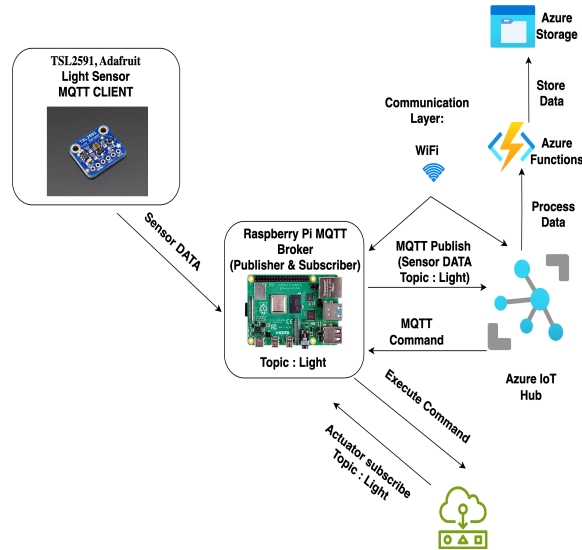


Fig. 8.1. IoT-Driven Lighting System for Indoor Farming

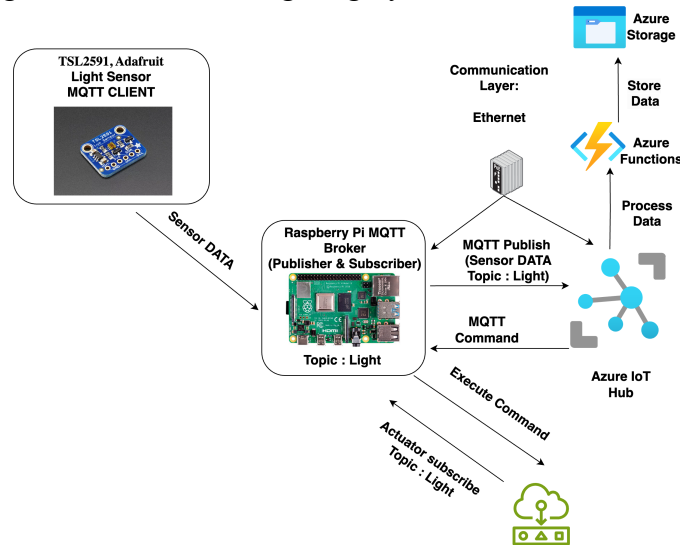


Fig. 8.2. IoT-Based Lighting Control Architecture

A6: Time Responsiveness Analysis of MQTT Communication Protocols

APPENDICES

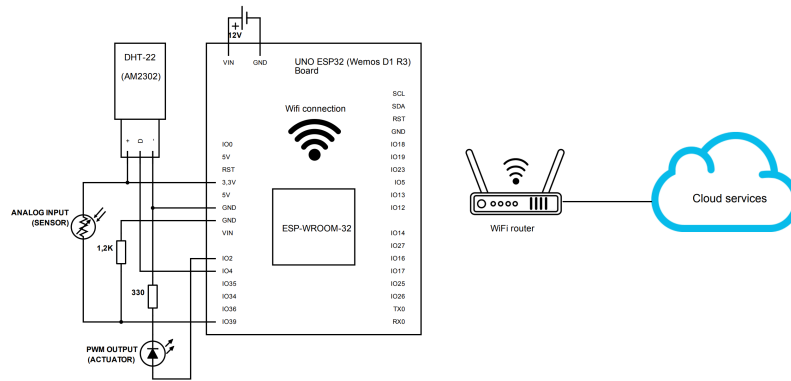


Fig. 8.3. Wi-Fi-based IoT sensor and actuator with cloud connection using HTTP

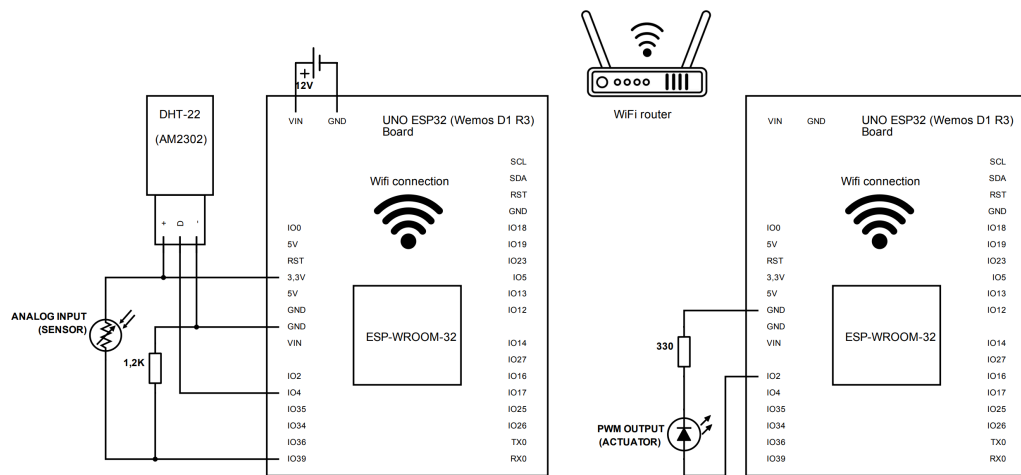


Fig. 8.4. WiFi based IoT sensor and actuator with point-to-point connection using HTTP protocol.

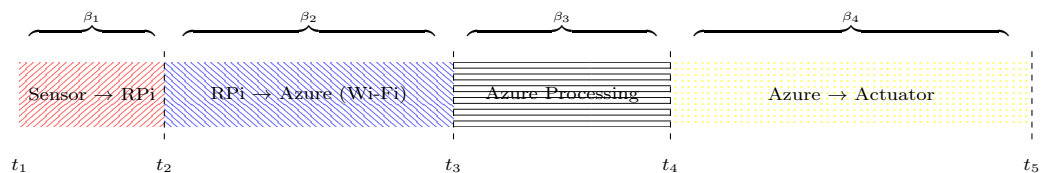


Fig. 8.5. Time Distribution for Wi-Fi in MQTT-based IoT system.

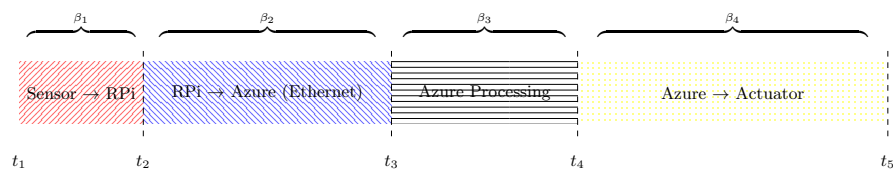
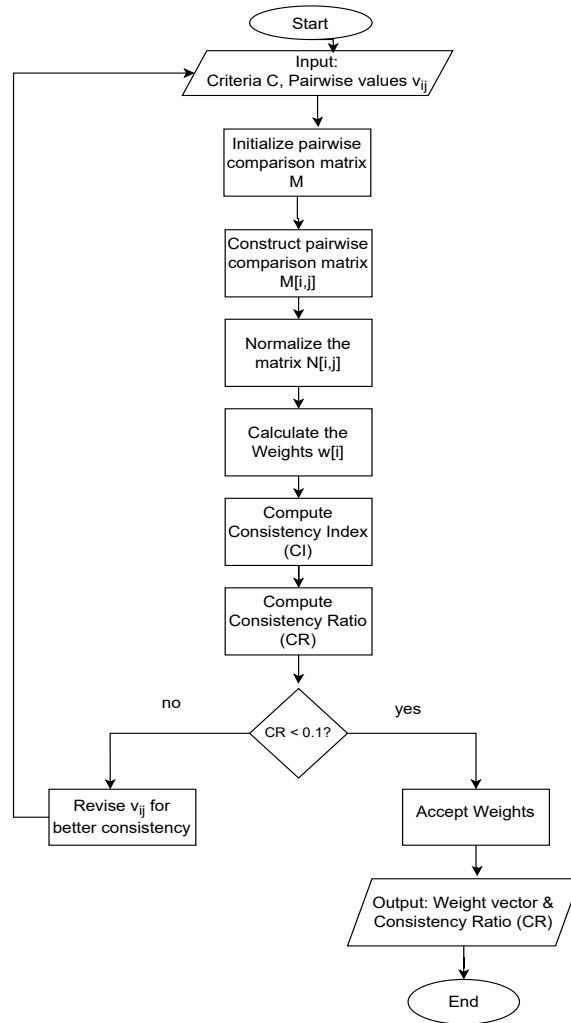


Fig. 8.6. Time Distribution for Ethernet in MQTT-based IoT system.

APPENDICES

A7: Weight Determination.



A8: Python Script For Machine Learning Model (Linear Regression):

```
import numpy as np
```

```
def get_user_input(prompt_text):
```

```
    """Parses comma-separated string input into a numpy array of floats."""
```

```
    while True:
```

```
        try:
```

```
            raw_input = input(prompt_text)
```

```
            data = [float(x.strip()) for x in raw_input.split(',')]

```

```
            return np.array(data)
```

```
        except ValueError:
```

```
            print("Error: Please enter numeric values separated by commas.")
```

APPENDICES

```
def get_single_value(prompt_text):
    """Parses a single float input."""
    while True:
        try:
            return float(input(prompt_text))
        except ValueError:
            print("Error: Please enter a valid number.")

def calculate_weights():
    print("="*60)
    print(" DATA-DRIVEN PROPORTIONAL NORMALIZATION CALCULATOR")
    print("="*60)

    # 1. Define Theoretical Maximums
    print("\n[STEP 1] Configuration: Theoretical Maximums")
    H_MAX = get_single_value(" Max Height (cm) [H_max]: ")
    B_MAX = get_single_value(" Max Biomass (g) [B_max]: ")
    A_MAX = get_single_value(" Max Leaf Area (cm²) [A_max]: ")

    # 2. Input Experimental Data
    print("\n[STEP 2] Input: Experimental Datasets")
    print(" (Enter values separated by commas, e.g., 12.5, 14.0, ...)")

    H_data = get_user_input(" Observed Height Data (cm): ")
    B_data = get_user_input(" Observed Biomass Data (g): ")
    A_data = get_user_input(" Observed Leaf Area Data (cm²): ")

    # Validation: Ensure datasets match in size
    if not (len(H_data) == len(B_data) == len(A_data)):
        print("\n[CRITICAL ERROR] Dataset size mismatch.")
        print(f" Height: {len(H_data)} | Biomass: {len(B_data)} | Area: {len(A_data)}")
        return

    # 3. Compute Normalized Scores
    # The cumulative sum of observations is normalized by the theoretical maximum
    # to determine the relative saturation of each metric.
    score_H = np.sum(H_data) / H_MAX
    score_B = np.sum(B_data) / B_MAX
    score_A = np.sum(A_data) / A_MAX

    total_score = score_H + score_B + score_A

    # 4. Derive Weighting Coefficients
```


APPENDICES

```
# Weights represent the proportional contribution of each normalized score to the total.
w_H = score_H / total_score
w_B = score_B / total_score
w_A = score_A / total_score

# 5. Output Results
print("\n" + "="*60)
print("  CALCULATION RESULTS")
print("="*60)
print(f'{'Metric':<20} | {'Normalized Score':<15} | {'Derived Weight (Coefficient)':<25}')
```

```
print("-" * 65)
print(f'{'Height (alpha)':<20} | {'score_H':<15.4f} | {'w_H':<25.4f}')
```

```
print(f'{'Biomass (beta)':<20} | {'score_B':<15.4f} | {'w_B':<25.4f}')
```

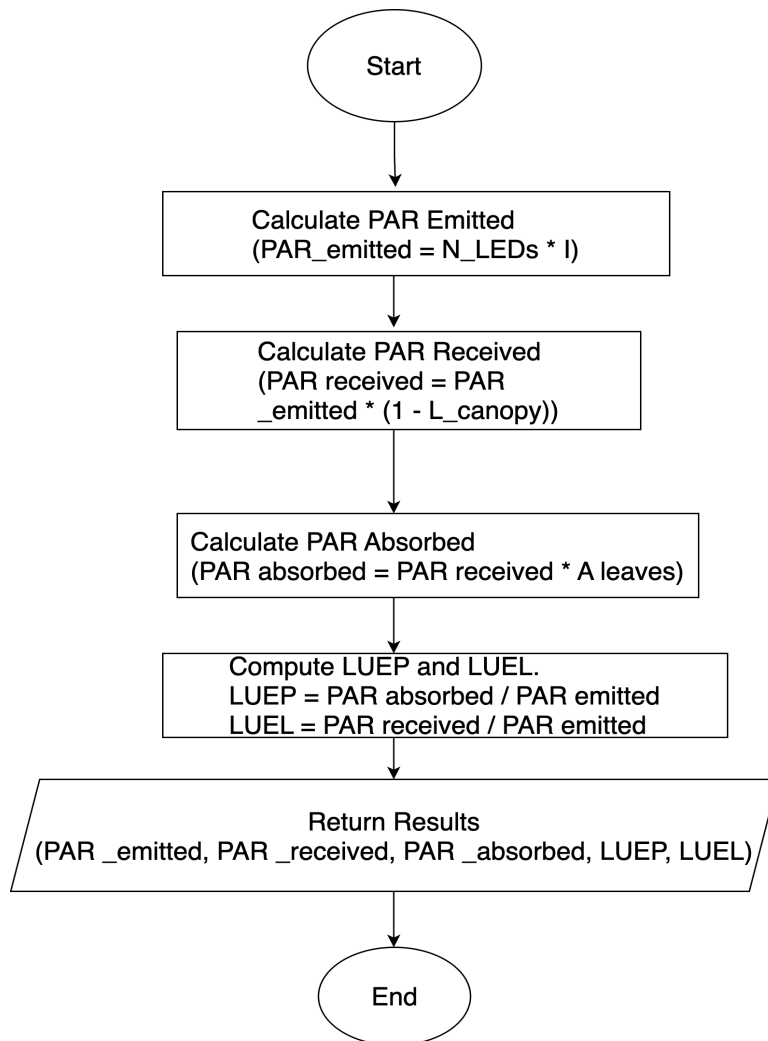
```
print(f'{'Leaf Area (gamma)':<20} | {'score_A':<15.4f} | {'w_A':<25.4f}')
```

```
print("-" * 65)
print(f'Validation Check (Sum of Weights): {'w_H + w_B + w_A':.4f}')
```

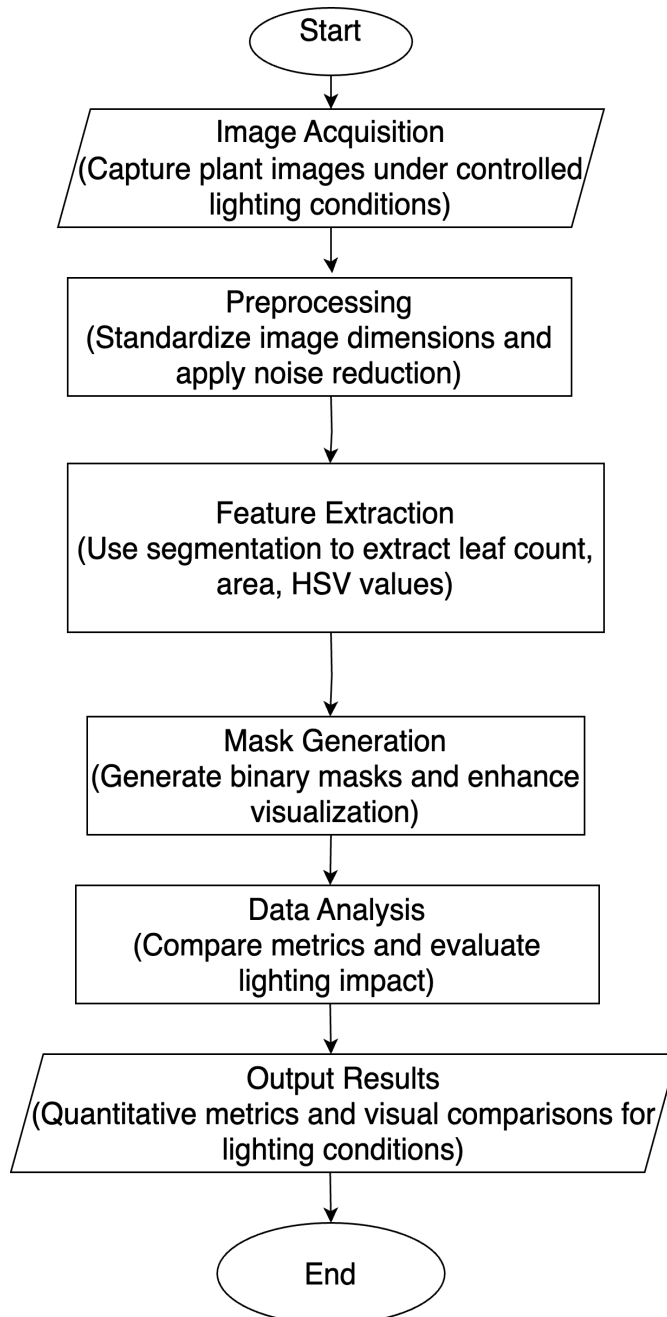
```
print("="*60)

if __name__ == "__main__":
    calculate_weights()
```

A9: Calculation of Light Use Efficiency Metrics



A10: Image Analysis Methodology for Plant Phenotyping

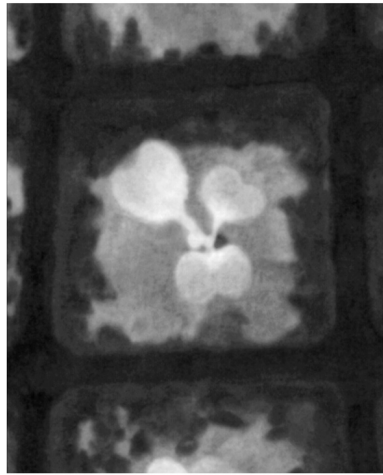


APPENDICES

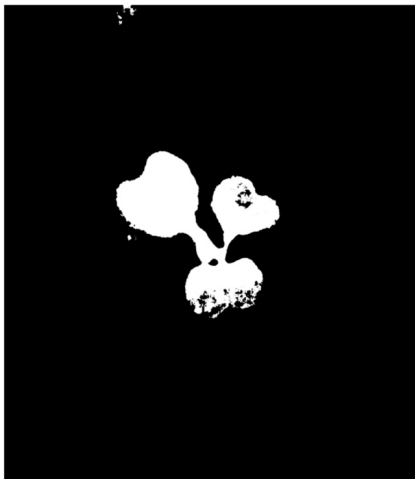
A11: Image Processing for Plant Detection and Lighting Simulation



Original image at week 2



Grayscale preprocessing image at week 2



Binary mask image at week 2



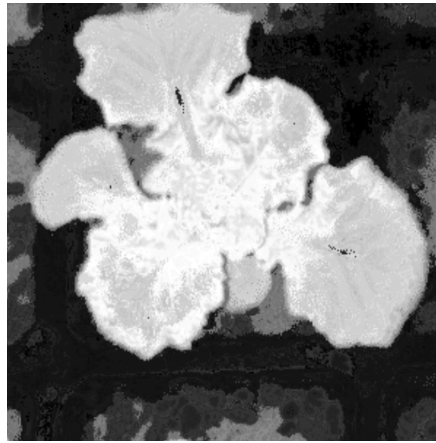
Segmented plant image at week 2

Fig. 8.7. Plant image segmentation workflow at week 2.

APPENDICES



Original image at week 3



Grayscale preprocessing image at week 3



Binary mask image at week 3



Segmented plant image at week 3

Fig. 8.8. Plant image segmentation workflow at week 3.

APPENDICES

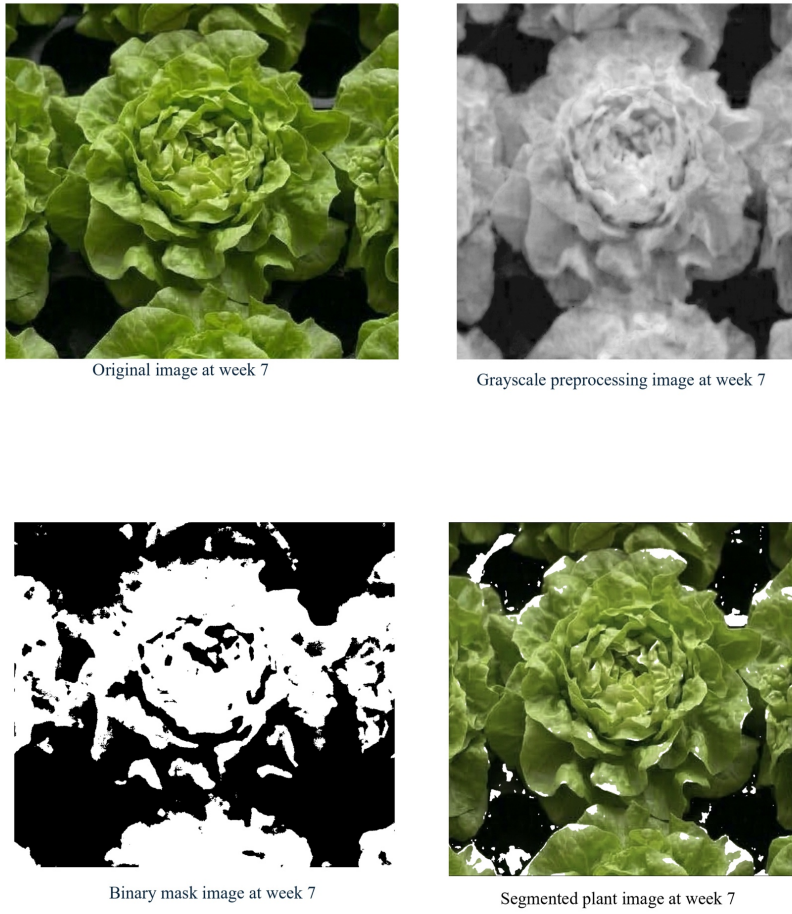
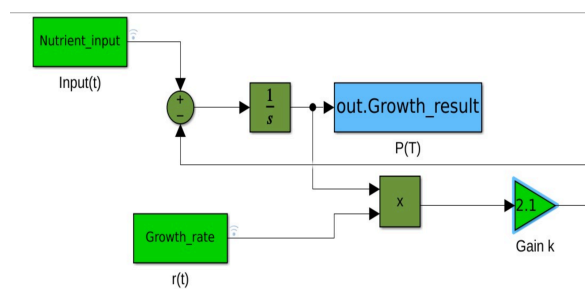


Fig. 8.9. Plant image segmentation workflow at week 7.

A12: Image Processing for Plant Detection and Lighting Simulation

Step	Parameter	Value
Grayscale Conversion	LAB channel	L*
Binary Threshold	Threshold value	120 (0–255)
Hole Filling	Max hole size	100 pixels
Smoothing	Kernel size, iterations	3×3, 1
Brightness Adjustment	α (contrast), β (brightness)	1.2, +50 (brighter) / -50 (darker)

A13: Real-Time Plant Growth Feedback Control System



APPENDICES

A14: Data Used To Compute GER Over 30 Days.

E*(t)	W*(t)	N*(t)	B*(t)	E (kWh)	W (liters)	N (grams)	B (grams)	GER
0.39	0.62	0.26	0.44	0.78	182	21	83	0.35
0.44	0.70	0.29	0.55	0.87	203	23	104	0.39
0.48	0.78	0.33	0.66	0.96	224	26	125	0.42
0.52	0.83	0.35	0.76	1.03	241	28	143	0.45
0.56	0.90	0.38	0.80	1.11	259	30	151	0.44
0.59	0.91	0.39	0.86	1.17	263	31	162	0.46
0.62	0.94	0.41	0.92	1.23	271	33	174	0.47
0.65	0.98	0.43	0.99	1.29	283	34	187	0.48
0.68	0.98	0.45	0.99	1.35	283	36	188	0.47

A15: Biomass Yield And GER Across Light Conditions

Light (hours/day)	Biomass Yield (grams)	GER
6	20	0.47
9	37	0.54
12	45	0.56

A16: Biomass Yield And GER Across Water Conditions

Water (L/day)	Biomass Yield (grams)	GER
5	162	0.46
7	198	0.50
9	221	0.59

A17: Biomass Yield And GER Across Nutrient Conditions

Nutrient (grams/day)	Biomass Yield (grams)	GER
----------------------	-----------------------	-----

APPENDICES

3	153	0.42
4	176	0.51
5	212	0.57

A18: The Effect Of Fertilizer Input And Watering Levels On Plant Growth.

Fertilizer Input (g)	Avg Plant Height(5L/day) (cm)	Avg Plant Height (7L/day) (cm)	Avg Plant Height (10L/day) (cm)
3	8.23	9.14	10.48
4	9.77	11.53	12.84
5	12.42	13.18	14.13
6	13.08	14.87	16.54
7	15.93	16.76	18.22
8	16.48	17.55	20.09
9	19.17	19.12	21.83
10	20.14	21.78	22.47
11	21.53	23.24	25.12

A19: The Effect Of Light Consumption And Watering Levels On Plant Growth.

Light Consumption (hours)	Avg Plant Height(5L/day) (cm)	Avg Plant Height (7L/day) (cm)	Avg Plant Height (10L/day) (cm)
6	12.83	16.54	18.18
7	15.47	17.13	19.07
8	16.22	19.78	21.84
9	18.88	20.23	22.53
10	20.54	22.87	25.76
11	21.48	23.52	26.24
12	24.13	26.18	28.09
13	24.77	27.14	31.17
14	27.53	28.48	31.82
15	28.24	31.23	34.78

APPENDICES

A20: Data for Leaf Area Growth Vs Water Levels

Day	Water: 5L/day (cm ²)	Water: 7L/day (cm ²)	Water: 9L/day (cm ²)
1	10.23	14.12	15.72
2	13.05	18.06	22.05
3	15.12	23.15	29.12
4	19.34	20.04	36.34
5	23.08	32.78	43.08
6	25.15	38.22	49.22
7	32.15	45.15	56.15
8	35.42	40.09	63.42
9	42.09	57.34	65.09
10	47.12	64.12	77.74
11	51.28	69.28	83.28
12	55.15	74.05	89.15
13	61.33	60.15	95.33
14	66.04	85.33	98.04
15	70.18	90.38	107.18
16	72.22	97.12	115.22
17	75.31	104.22	124.31
18	88.07	95.06	133.07
19	90.45	118.15	142.45
20	101.09	126.04	151.09
21	105.23	134.22	160.23
22	113.15	142.31	170.15
23	100.42	120.45	180.42
24	126.11	159.09	182.11
25	133.05	167.74	199.35
26	140.22	175.23	206.22
27	149.34	160.15	213.34
28	157.18	191.42	220.18
29	158.06	199.11	226.06
30	173.74	207.35	233.74

A21: Data for Leaf Area Vs Nutrient

Day	Nutrient: 3g (cm ²)	Nutrient: 5g (cm ²)	Nutrient: 7g (cm ²)
1	2.41	3.58	5.14
2	4.15	5.92	8.27
3	6.72	9.29	12.38
4	8.19	14.07	18.52

APPENDICES

5	10.37	18.84	28.71
6	12.28	22.19	38.25
7	14.52	28.11	48.19
8	16.14	34.05	58.13
9	18.39	40.38	65.31
10	21.15	45.17	68.16
11	23.32	52.23	78.22
12	25.19	59.08	88.09
13	27.38	65.12	98.37
14	29.07	70.39	108.01
15	32.22	72.43	115.18
16	34.25	78.12	125.18
17	36.36	80.29	135.35
18	39.12	88.02	145.03
19	42.48	94.19	150.41
20	45.13	98.07	155.13
21	47.27	104.26	165.29
22	49.19	110.35	175.11
23	51.46	115.49	182.46
24	53.15	120.12	188.11
25	56.09	125.78	190.39
26	58.26	128.27	198.26
27	60.42	130.08	205.44
28	62.23	132.46	210.23
29	64.17	134.33	215.09
30	65.47	135.29	218.18

A22 : Data for Leaf Area Vs Light

Day	Light: 8h (cm ²)	Light: 10h (cm ²)	Light: 12h (cm ²)
1	12.45	11.12	15.34
2	16.23	18.56	23.12
3	20.58	24.44	31.89
4	24.12	32.78	42.56
5	29.05	40.23	55.12
6	33.15	49.11	70.34
7	38.42	58.05	82.89
8	42.18	66.22	96.56
9	47.12	75.15	108.23
10	51.34	82.58	121.12
11	56.22	91.05	134.34
12	60.15	98.22	148.12
13	65.45	105.34	160.23

APPENDICES

14	71.23	114.12	172.89
15	75.12	122.45	185.34
16	80.34	131.23	198.12
17	84.12	138.45	204.23
18	87.56	145.12	211.56
19	92.23	152.56	218.22
20	95.12	160.23	225.45
21	99.34	166.12	231.23
22	105.56	171.34	238.12
23	108.89	175.22	245.23
24	112.12	181.45	250.12
25	116.34	186.78	258.34
26	119.56	192.23	265.45
27	123.12	196.12	272.56
28	126.45	201.34	278.12
29	130.23	205.56	285.34
30	134.12	210.23	290.12

A23: User Input Data Acquisition

Enter the data for plant growth analysis (type 'done' to finish):

Enter Height (or type 'done' to finish): 10

Enter Biomass: 20

Enter Leaf Area: 160

.....

Enter Height (or type 'done' to finish): done

Normalized Weights for PGI:

w_Height: 0.3335

w_Biomass: 0.3264

w_Leaf_Area: 0.3401

A24: Weight Calculation for PGI Parameters

Parameter	$\sum X_i$	$X_{\max}$	X_{norm}	$H_{\text{norm}} + B_{\text{norm}} + A_{\text{norm}}$	W_i
Height (H)	274.4 cm	34.09 cm	9.146667	26.276667	0.3335

APPENDICES

Biomass (B)	2338.0 g	286.87 g	7.793333	26.276667	0.3264
Leaf Area (A)	2801.0 cm ²	304.27 cm ²	9.336667	26.276667	0.3401

A25: Raw Data For Computing The Plant Growth Index (PGI)

Time (Day)	H (cm)	B (g)	A (cm ²)
1	4.0	5.0	20.0
2	4.5	8.0	30.0
4	5.5	12.0	45.0
6	7.0	20.0	65.0
8	9.0	35.0	90.0
10	11.5	55.0	120.0
12	14.0	80.0	150.0
14	17.0	115.0	180.0
16	20.0	155.0	205.0
18	22.5	195.0	225.0
20	25.0	230.0	235.0
22	27.0	260.0	242.0
24	28.5	280.0	246.0
26	29.2	290.0	248.0
28	29.8	298.0	250.0
30	30.0	300.0	250.0

APPENDICES

A26: Computation of the Plant Growth Index (PGI)

Time (Day)	Height H (cm)	Biomass B (g)	Leaf Area A (cm ²)	PGI
1	4.0	5.0	20.0	0.075
2	4.5	8.0	30.0	0.096
4	5.5	12.0	45.0	0.129
6	7.0	20.0	65.0	0.178
8	9.0	35.0	90.0	0.246
10	11.5	55.0	120.0	0.330
12	14.0	80.0	150.0	0.419
14	17.0	115.0	180.0	0.524
16	20.0	155.0	205.0	0.628
18	22.5	195.0	225.0	0.720
20	25.0	230.0	235.0	0.796
22	27.0	260.0	242.0	0.857
24	28.5	280.0	246.0	0.899
26	29.2	290.0	248.0	0.919
28	29.8	298.0	250.0	0.936
30	30.0	300.0	250.0	0.941

A27: Influence Of Environmental Factors On Crop Growth Using Multivariate Regression.

Sample ID	Light (lumens)	Phosphorus (mg/kg)	Potassium (mg/kg)	Nitrogen (mg/kg)	Moisture (%)	Temperature (°C)	Observed Growth (g)
1	9174.0	27.0	41.0	106.0	68.0	20.0	305.09
2	6860.0	24.0	47.0	121.0	47.0	18.0	230.35
3	7294.0	23.0	43.0	107.0	33.0	21.0	243.84
4	7130.0	27.0	44.0	95.0	54.0	19.0	240.97
5	7095.0	27.0	40.0	94.0	43.0	25.0	242.32
6	9092.0	22.0	43.0	126.0	38.0	21.0	298.30
7	7638.0	25.0	41.0	123.0	55.0	19.0	257.60
8	8169.0	24.0	45.0	82.0	31.0	23.0	269.11
9	6466.0	21.0	44.0	116.0	49.0	23.0	219.27
10	7238.0	27.0	43.0	86.0	57.0	27.0	246.30
11	6330.0	31.0	40.0	100.0	36.0	21.0	220.43
12	7482.0	25.0	40.0	88.0	37.0	23.0	251.79

APPENDICES

A28:Experimental Environmental Conditions and Corresponding Plant Growth Measurements

Factor	Multivariate Regression (β Coefficient)
Light Intensity	0.03
Phosphorus	0.54
Potassium	-0.43
Nitrogen	0.04
Moisture	0.06
Temperature	0.23

A29:Comparison between experimented crop yield and predicted values obtained using SVM, Gradient Boosting, Deep Neural Network, and Time Series models for individual plant samples

Plant ID	Observed Growth (g)	SVM Prediction (g)	Gradient Boosting Prediction (g)	DNN Prediction (g)	Time Series Prediction (g)
1	214.26	205.40	212.10	210.55	213.80
2	298.15	290.05	295.50	292.15	297.90
3	265.06	255.80	263.15	261.40	264.85
4	245.76	235.90	243.50	240.20	245.10
5	184.15	175.60	182.00	179.80	183.95
6	179.29	170.50	177.40	174.10	178.90
7	164.12	158.30	162.80	160.50	163.75
8	287.83	275.90	285.10	282.40	287.15
9	243.53	233.40	241.00	239.50	243.10
10	264.09	254.70	261.80	259.20	263.65
11	159.66	152.10	158.20	155.40	159.40
12	301.23	291.50	299.10	296.80	295.00

A30:Measured Growth Before and After IoT Implementation

Time (weeks)	IoT Biomass (g)	Non-IoT Biomass (g)	IoT Height (cm)	Non-IoT Height (cm)	IoT Leaf Area (cm ²)	Non-IoT Leaf Area (cm ²)
1	4.36	2.852	2.30	2.48	22.40	15.91
2	17.05	9.560	4.80	3.92	41.70	27.62
3	50.75	27.18	7.20	6.20	61.90	44.48
4	60.29	35.34	9.66	7.64	103.60	70.30
5	71.66	46.84	15.88	11.98	135.00	96.27
6	104.95	63.70	18.30	14.88	179.30	126.72

APPENDICES

7	116.12	78.64	20.40	17.14	215.40	152.00
8	221.07	129.78	29.90	20.18	248.30	178.66

A31: System Optimization Metrics Before And After IoT Implementation

Metric	Non-IoT	IoT
Light efficiency	0.0209	0.0353
CO ₂ efficiency	0.0171	0.0580
Nutrient efficiency	0.200	0.352
Photoperiod efficiency	1.014	1.590

A32: Real-Time Alert Accuracy Improvements

Experiment	Baseline System Accuracy (%)	Optimized System Accuracy (%)	Improvement (%)
1	91	93	+2
2	91	94	+3
3	91	95	+4
4	91	96	+5
5	91	96	+5

A33: Experimental vs. simulated RGR values (Treatment Group A)

Light	Photoperiod	Biomass (g)	Height (cm)	Leaf Area (cm ²)	RGR EXP	RGR Computed
775	16	2.85	2.4	15.91		
775	16	9.56	3.9	27.62	2.35	2.48
775	16	27.18	6.2	44.48	1.84	1.54
775	16	35.34	7.6	70.30	0.30	0.86
775	16	46.84	11.9	96.27	0.33	0.43
775	16	63.70	14.8	126.72	0.36	0.25
775	16	78.64	17.1	152.00	0.23	0.31
775	16	129.78	20.1	178.66	0.65	0.63

A34: Experimental vs. simulated RGR values (Treatment Group B)

light	Photoperiod	Biomass (g)	Height (cm)	Leaf Area (cm ²)	RGR EXP	RGR Computed
220	14	4.36	2.3	22.40		
450	15	17.05	4.8	41.70	2.91	2.63

APPENDICES

550	16	50.75	7.2	61.90	1.98	1.63
700	17	60.29	9.6	103.60	0.19	0.75
850	18	71.66	15.8	135.00	0.19	0.17
1050	19	104.95	18.3	179.30	0.46	0.37
1150	20	116.12	20.4	215.40	0.11	-0.07
1300	20	221.07	29.9	248.30	0.90	0.82

A35: Experimental vs. simulated RGR values (Treatment Group C)

light	Photoperiod	Biomass (g)	Height (cm)	Leaf Area (cm ²)	RGR EXP	RGR Computed
180	12	3.83	3.0	13.00		
350	12	11.88	4.2	29.60	2.10	2.48
500	13	35.00	6.6	54.40	1.95	1.45
650	13	46.88	8.1	84.40	0.34	0.23
750	14	57.13	13.4	145.60	0.22	0.26
850	15	79.75	18.0	159.90	0.40	0.26
900	16	120.00	18.8	180.60	0.50	0.26
950	16	158.25	23.9	212.80	0.32	0.56

A36: MATLAB Code for Stepwise Regression Analysis of RGR (Treatment Groups A, B, C)

```

%% -----
% DATA
%% -----

% ----- Group B (Adaptive) -----
Light_B = [220 450 550 700 850 1050 1150 1300];
Biomass_B = [4.36 17.05 50.75 60.29 71.66 104.95 116.12 221.07];
Photoperiod_B = [14 15 16 17 18 19 20 20];

% ----- Group A (Static) -----
Light_A = [775 775 775 775 775 775 775 775];
Biomass_A = [2.852 9.560 27.18 35.34 46.84 63.70 78.64 129.78];
Photoperiod_A = [16 16 16 16 16 16 16 16];

% ----- Group C (Efficient) -----
Light_C = [180 350 500 650 750 850 900 950];
Biomass_C = [3.83 11.88 35.00 46.88 57.13 79.75 120.00 158.25];
Photoperiod_C = [12 12 13 13 14 15 16 16];

```


APPENDICES

```

%%% -----
% Relative Biomass Growth Rate
%%% -----

RGR_A = diff(Biomass_A)./Biomass_A(1:end-1);
RGR_B = diff(Biomass_B)./Biomass_B(1:end-1);
RGR_C = diff(Biomass_C)./Biomass_C(1:end-1);

%%% -----
% Prepare predictors (week 2-8)
%%% -----

Light = [Light_A(2:end) Light_B(2:end) Light_C(2:end)]';
P = [Photoperiod_A(2:end) Photoperiod_B(2:end) Photoperiod_C(2:end)]';

Week = repmat((2:8)',3,1);

RGR = [RGR_A RGR_B RGR_C]';

%%% -----
% Build table
%%% -----

T = table(Light,P,Week,RGR);

%%% -----
% Center variables
%%% -----

T.Ic = T.Light - mean(T.Light);
T.Pc = T.P - mean(T.P);
T.Wc = T.Week - mean(T.Week);

%%% -----
% Stepwise regression model
%%% -----

mdl = stepwiselm(T, ...
    'RGR ~ Ic + Pc + Wc + Ic^2 + Pc^2 + Wc^2 + Ic:Pc + Ic:Wc + Pc:Wc', ...
    'Criterion','aic');
disp(mdl.Formula)
disp(mdl.Coefficients)

```

A37: Summary Metrics Across Configurations

Configuration	Emitted PAR ($\mu\text{mol}/\text{m}^2/\text{s}$)	Received PAR ($\mu\text{mol}/\text{m}^2/\text{s}$)	Absorbed PAR ($\mu\text{mol}/\text{m}^2/\text{s}$)	LUEP	LUEL	LAI
Config A	300	260.4	228	0.760	0.868	0.7
Config B	750	666.75	615	0.820	0.889	1.2
Config C	1400	1237.6	1162	0.830	0.884	1.8
Config D	2000	1688	1540	0.770	0.844	2.4

APPENDICES

A38: Treatment Groups

Week	Group	Light	Nutrient	CO2	Growth Rate (cm/wk)	Chlorophyll (SPAD)	Biomass (g)	Height (cm)	Leaf Area (cm ²)
1	A	775.00	81.00	950.00	2.2	30	3.06	2.2	15.6
2	A	775.00	81.00	950.00	1.5	36	9.5	3.7	25.9
3	A	775.00	81.00	950.00	2.4	40	28.00	6.1	43.7
4	A	775.00	81.00	950.00	2.1	45	37.5	8.2	67.9
5	A	775.00	81.00	950.00	3.7	49	45.7	11.9	93.7
6	A	775.00	81.00	950.00	2.4	55	63.8	14.3	126.3
7	A	775.00	81.00	950.00	2.6	55	88.4	16.9	152.4
8	A	775.00	81.00	950.00	3.9	62	126.6	20.8	180.2
1	A	775.00	81.00	950.00	2.1	34	2.5	2.1	14.9
2	A	775.00	81.00	950.00	1.9	37	8.8	4.0	30.1
3	A	775.00	81.00	950.00	1.8	39	26.1	5.8	42.3
4	A	775.00	81.00	950.00	1.4	44	34.8	7.2	71.6
5	A	775.00	81.00	950.00	3.2	52	49.8	10.4	98.0
6	A	775.00	81.00	950.00	4.5	53	63.6	14.9	127.5
7	A	775.00	81.00	950.00	1.7	58	80.1	16.6	150.1
8	A	775.00	81.00	950.00	3.9	59	125.6	20.5	177.9
1	A	775.00	81.00	950.00	2.7	33	3.4	2.7	14.0
2	A	775.00	81.00	950.00	1.00	39	10.1	3.7	29.1
3	A	775.00	81.00	950.00	2.9	42	27.6	6.6	45.3
4	A	775.00	81.00	950.00	0.9	46	34.6	7.5	68.4
5	A	775.00	81.00	950.00	5.2	48	45	12.7	96.7
6	A	775.00	81.00	950.00	1.4	55	65.4	14.1	125.5
7	A	775.00	81.00	950.00	3.8	56	77.8	17.9	153.9
8	A	775.00	81.00	950.00	1.00	63	132.3	18.9	179.2
1	A	775.00	81.00	950.00	2.4	36	2.8	2.4	17.5
2	A	775.00	81.00	950.00	2.1	38	9.6	4.5	25.6
3	A	775.00	81.00	950.00	1.5	43	26.9	6.0	47.6
4	A	775.00	81.00	950.00	1.6	46	34.6	7.6	71.6
5	A	775.00	81.00	950.00	4.5	51	46.9	12.1	96.9
6	A	775.00	81.00	950.00	3.7	55	63.1	15.8	128.5
7	A	775.00	81.00	950.00	1.5	57	76	17.3	152.5
8	A	775.00	81.00	950.00	1.9	61	131.8	19.2	176.2
1	A	775.00	81.00	950.00	3.00	36	2.5	3.0	17.4
2	A	775.00	81.00	950.00	0.7	37	9.8	3.7	27.2
3	A	775.00	81.00	950.00	2.8	42	27.3	6.5	43.3
4	A	775.00	81.00	950.00	1.2	48	35.2	7.7	71.7

APPENDICES

5	A	775.00	81.00	950.00	5.1	48	46.8	12.8	95.8
6	A	775.00	81.00	950.00	2.5	56	62.6	15.3	125.6
7	A	775.00	81.00	950.00	1.7	58	70.9	17.0	150.9
8	A	775.00	81.00	950.00	4.5	64	132.6	21.5	179.6
1	B	220.00	22.00	410.00	2.3	31	4.36	2.3	22.4
2	B	450.00	45.00	430.00	2.5	37	17.05	4.8	41.7
3	B	550.00	55.00	440.00	2.4	43	50.75	7.2	61.9
4	B	700.00	70.00	460.00	2.4	44	60.29	9.6	103.6
5	B	850.00	85.00	480.00	5.8	51	71.65	15.4	135.0
6	B	1050.00	105.00	510.00	2.9	51	104.95	18.3	179.3
7	B	1150.00	115.00	530.00	2.1	57	116.11	20.4	215.4
8	B	1300.00	130.00	550.00	9.5	60	221.06	29.9	248.3
1	B	220.00	22.00	410.00	3.00	34	5.48	3.0	20.0
2	B	450.00	45.00	430.00	2.4	36	16.84	5.4	43.0
3	B	550.00	55.00	440.00	1.6	44	42.22	7.0	60.0
4	B	700.00	70.00	460.00	2.00	43	57.24	9.0	98.5
5	B	850.00	85.00	480.00	6.1	48	77.54	15.1	129.3
6	B	1050.00	105.00	510.00	4.7	52	100.48	19.8	175.6
7	B	1150.00	115.00	530.00	2.8	55	129.10	22.6	217.6
8	B	1300.00	130.00	550.00	7.4	58	211.72	30.0	247.8
1	B	220.00	22.00	410.00	2.7	33	5.88	2.7	20.4
2	B	450.00	45.00	430.00	2.5	37	17.05	5.2	36.3
3	B	550.00	55.00	440.00	1.3	38	46.28	6.5	59.6
4	B	700.00	70.00	460.00	4.3	46	65.77	10.8	94.3
5	B	850.00	85.00	480.00	6.00	49	75.51	16.8	129.0
6	B	1050.00	105.00	510.00	2.1	53	100.07	18.9	177.1
7	B	1150.00	115.00	530.00	2.4	57	127.48	21.3	215.4
8	B	1300.00	130.00	550.00	5.7	61	205.43	27.0	252.7
1	B	220.00	22.00	410.00	3.2	30	7.30	3.2	21.0
2	B	450.00	45.00	430.00	1.4	36	18.06	4.6	39.3
3	B	550.00	55.00	440.00	2.3	42	46.28	6.9	61.4
4	B	700.00	70.00	460.00	2.4	45	63.13	9.3	102.5
5	B	850.00	85.00	480.00	6.3	48	72.47	15.6	130.0
6	B	1050.00	105.00	510.00	2.5	54	103.12	18.1	180.3
7	B	1150.00	115.00	530.00	2.8	53	130.93	20.9	213.4
8	B	1300.00	130.00	550.00	4.2	58	213.96	25.1	247.8
1	B	220.00	22.00	410.00	2.3	32	6.69	2.3	19.4
2	B	450.00	45.00	430.00	2.4	39	14.21	4.7	38.1
3	B	550.00	55.00	440.00	2.9	40	52.98	7.6	58.6
4	B	700.00	70.00	460.00	2.00	44	60.49	9.6	97.5
5	B	850.00	85.00	480.00	6.9	49	79.98	16.5	139.6

APPENDICES

6	B	1050.00	105.00	510.00	1.5	50	102.51	18.0	176.2
7	B	1150.00	115.00	530.00	2.9	58	129.92	20.9	213.9
8	B	1300.00	130.00	550.00	3.9	60	239.33	24.8	252.2
1	C	180.00	18.00	405.00	2.3	31	3.82	3.0	13.0
2	C	350.00	35.00	410.00	2.5	38	11.87	4.2	29.6
3	C	500.00	50.00	412.00	2.4	43	35.00	6.6	54.4
4	C	650.00	65.00	415.00	2.4	47	46.87	8.1	84.4
5	C	750.00	75.00	417.00	5.8	51	57.12	13.4	145.6
6	C	850.00	85.00	418.00	2.9	54	79.75	18.0	159.9
7	C	900.00	90.00	419.00	2.1	57	120.00	18.8	180.6
8	C	950.00	95.00	420.00	9.5	61	158.25	23.9	212.8
1	C	180.00	18.00	405.00	3.00	32	4.21	2.1	14.5
2	C	350.00	35.00	410.00	2.4	38	12.46	4.9	37.6
3	C	500.00	50.00	412.00	1.60	44	39.20	6.2	53.4
4	C	650.00	65.00	415.00	2.00	48	51.09	8.4	126.7
5	C	750.00	75.00	417.00	6.10	46	60.95	12.5	126.8
6	C	850.00	85.00	418.00	4.70	56	85.17	17.5	146.6
7	C	900.00	90.00	419.00	2.80	54	128.28	18.7	187.0
8	C	950.00	95.00	420.00	7.40	62	169.33	21.8	193.2
1	C	180.00	18.00	405.00	2.70	34	5.06	2.4	31.0
2	C	350.00	35.00	410.00	2.50	40	9.59	4.6	53.4
3	C	500.00	50.00	412.00	1.30	43	28.00	6.1	73.0
4	C	650.00	65.00	415.00	4.30	47	51.09	8.6	121.0
5	C	750.00	75.00	417.00	6.00	49	73.14	12.9	170.0
6	C	850.00	85.00	418.00	2.10	55	85.17	15.5	191.2
7	C	900.00	90.00	419.00	2.40	57	128.28	18.6	208.4
8	C	950.00	95.00	420.00	5.70	62	167.00	20.5	212.8
1	C	180.00	18.00	405.00	3.20	31	3.75	2.5	27.9
2	C	350.00	35.00	410.00	1.40	34	11.88	4.7	51.9
3	C	500.00	50.00	412.00	2.30	40	35.00	6.8	71.3
4	C	650.00	65.00	415.00	2.40	44	46.88	9.0	120.0
5	C	750.00	75.00	417.00	6.30	49	57.13	14.1	161.3
6	C	850.00	85.00	418.00	2.50	52	79.75	17.6	189.7
7	C	900.00	90.00	419.00	2.80	54	120.00	18.1	195.80
8	C	950.00	95.00	420.00	4.20	62	158.25	24.1	205.2
1	C	180.00	18.00	405.00	2.30	33	3.94	2.2	26.4
2	C	350.00	35.00	410.00	2.40	38	11.20	4.8	49.3
3	C	500.00	50.00	412.00	2.90	40	32.71	6.3	72.2
4	C	650.00	65.00	415.00	2.00	49	47.30	8.8	119.8
5	C	750.00	75.00	417.00	6.90	48	57.13	12.1	146.4
6	C	850.00	85.00	418.00	1.50	56	76.32	16.2	179.2

APPENDICES

7	C	900.00	90.00	419.00	2.90	55	123.60	19.2	195.6
8	C	950.00	95.00	420.00	3.90	62	174.87	24.7	218.3

A39: Monitoring API Alerts

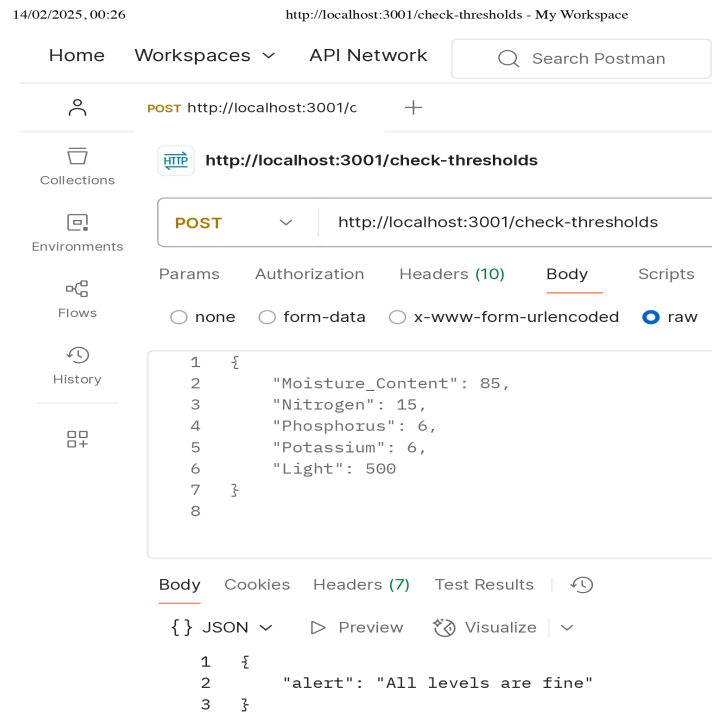


Fig. 8.10. API Response for Optimal Environmental Conditions.

APPENDICES

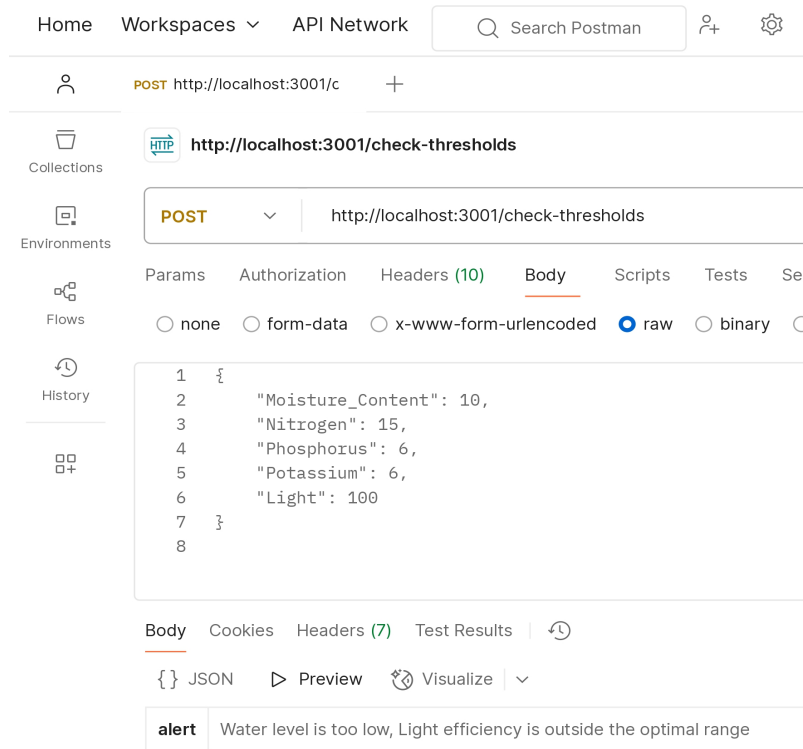


Fig. 8.11. API Alert Response for Combined Moisture and Light Threshold Violations.

APPENDICES

14/02/2025, 00:36 http://localhost:3001/check-thresholds - My Workspace

Home Workspaces ⌵ API Network 🔍 Search Postman

 **POST** http://localhost:3001/c +

  **http://localhost:3001/check-thresholds**

Collections

 **POST** ⌵ http://localhost:3001/check-thresholds

Environments

 Params Authorization Headers (10) Body Scripts

Flows

☐ none ☐ form-data ☐ x-www-form-urlencoded ☒ raw

 History

 +

```
1 {
2   "Moisture_Content": 85,
3   "Nitrogen": 15,
4   "Phosphorus": 6,
5   "Potassium": 6,
6   "Light": 100
7 }
8
```

Body Cookies Headers (7) Test Results 

{ } JSON ▶ Preview  Visualize ⌵

alert	Light efficiency is outside the optimal range
--------------	---

Fig 8.12. API Alert Response for Light Intensity Outside the Optimal Range.

ACKNOWLEDGEMENT

10. ACKNOWLEDGEMENT

I would like to begin by expressing my profound gratitude to my supervisor, Dr. István Szabó, for his exceptional guidance and support. Our collaboration began during my Bachelor's studies in 2018 and continued throughout my Master's and PhD journey. His invaluable guidance and expertise were fundamental in shaping my research and inspiring me to pursue excellence.

My academic path at the Hungarian University of Agriculture and Life Sciences (MATE) in Gödöllő has been a significant and rewarding chapter of my life. I joined the university in September 2018 to finish my Bachelor's degree, and the inspiring environment compelled me to continue for both my Master's and Doctoral studies. It has been an incredible seven years, and this dissertation is the culmination of work that began under Dr. Szabó's guidance from the very beginning. I am deeply thankful to him for his constant supervision. I was also fortunate to be a member of the FIEK project for two years. This experience was extremely enriching and allowed me to learn a great deal, significantly contributing to my professional development.

I extend my sincere appreciation to Prof. Dr. Gábor Kalácska, DSc, the current head of the Mechanical Engineering Doctoral School, for his strategic guidance and invaluable advice that were crucial for successfully navigating this PhD journey. I am also grateful to Prof. István Farkas, the former head of the Doctoral School, for his steadfast support, recommendations, and for facilitating the fulfilment of all doctoral requirements. I gratefully acknowledge the financial support provided by the Stipendium Hungaricum Scholarship Programme. This work was accomplished under the auspices of the Mechanical Engineering Doctoral School at MATE.

My thanks also go to the faculty and staff at MATE for providing the resources and facilities necessary for my research. A special note of thanks is due to Ms. Claudia Berze, secretary to Dr. Szabó, for her always reactive and invaluable assistance whenever it was needed. I would also like to express my sincere appreciation to Ms. Zsuzsanna Tassy, the International PhD Coordinator at the University's Doctoral and Habilitation Center, for her exceptional support and guidance throughout both my MSc and PhD studies.

On a personal note, my deepest gratitude goes to my husband, Youssef, for his unwavering support, patience, and for meticulously reading my papers before submission. I am also profoundly thankful to my mother for her countless sacrifices, including visiting me whenever I needed her, and to my father and brothers for their constant encouragement and belief in me.

Finally, to all those not mentioned by name but who contributed in various ways, thank you. Your support has been an essential part of my academic and personal growth.

Sincerely,

Gödöllő, 02. 09. 2025

Nezha Kharraz