Cairo University

Faculty of Computers and Artificial Intelligence: Operations Research & Decision Support Department

<u>Irrigation water management</u>

This Graduation Project is Submitted to
The Faculty of Computers and Artificial Intelligence
Cairo University
In Partial Fulfillment of the Requirements
for the bachelor's degree In

Operations Research and Decision Support

Presented By:

Abdelrahman Sameh Fahem	**20200285**
Alaa Saeed Ahmed	**20201023**
Marwa Osama Emam	**20200515**
Menna Mahmoud Sayed	**20200560**

Under Supervision of:
Dr. Sally Saad Kassem

CAIRO UNIVERSITY

July/2024

ABSTRACT

In contemporary agriculture, the efficient management of irrigation water is critical for enhancing crop yield while conserving water resources. This challenge centers on designing an irrigation strategy that carefully considers multiple variables without compromising crop productivity. Optimizing the quantity of the water to be used in irrigation and the interval between each irrigation (plant irrigation needs) is a complex process for all the farmers, as it depends on various factors some of those factors are controllable and some are uncontrollable.

Water scarcity, waterlogging, erosion, pollution, and climate change are some of the reasons that make the determination of the optimal quantity of irrigation water to be used during a season a very challenging mission. Based on the previous information, this project aims to build a mathematical model that can help solve the problem of knowing the optimal quantity of irrigation water to be used during a season using linear programming and machine learning.

Precision irrigation technologies offer a powerful solution for optimizing water use. Sensors that monitor soil moisture, weather conditions, and plant health data provide real-time insights into crop water needs. This allows farmers to deliver water directly to the root zone at the most crucial times, minimizing evaporation and runoff losses. Automation systems coupled with these sensors can further enhance efficiency by scheduling irrigation events based on real-time data, ensuring crops receive the precise amount of water they require throughout the growing season.

Applying such an optimization model in the Egyptian farms and agricultural lands will help in conserving water resources by reducing water consumption while also enhancing crop productivity as the model takes into consideration the yield.

Beyond the direct benefits of water conservation and yield improvement, this model offers significant economic advantages for Egyptian agriculture. By optimizing irrigation practices, farmers can drastically reduce water waste, translating to lower pumping and operation costs. Additionally, increased crop yields due to optimal water management can lead to higher profits. This economic efficiency can incentivize broader adoption of the model, creating a positive ripple effect throughout the agricultural sector.

DECLARATION

We hereby declare that our dissertation is entirely our work and genuine / original. We understand that in case of discovery of any PLAGIARISM at any stage, our group will be assigned an F (FAIL) grade, and it may result in withdrawal of our bachelor's degree.

Group members:

Name	Signature
Abdelrahman Sameh Fahem	____________________
Alaa Saeed Ahmed	____________________
Marwa Osama Emam	____________________
Menna Mahmoud Sayed	____________________

Plagiarism certificate

This is to certify that the project entitled “Irrigation Water Management”, which is being submitted here for the award of the “**Bachelor of Computer and Artificial Intelligence Degree**” in “**Operations Research and Decision Support”**. This is the result of the original work by **Abdelrhman Sameh, Menna Mahmoud, Alaa Saeed, and Marwa Osama** under my supervision and guidance. The work embodied in this project has not been done earlier for the basis of award of any degree or compatible certificate or similar tile of this for any other diploma/examining body or university to the best of my knowledge and belief.

Turnitin Originality Report
Processed on 4-July-2024 00:14 PKT
ID: 300502964
Word Count: 10597

Similarity Index
10%
Similarity by Source
Internet Sources: 06%
Publications: 0 %
Student Papers: 08%

Date: 04/07/2024

Dr. Sally Kassem (Supervisor)

DEDICATION

We dedicate this project to the resilient farmers who toil tirelessly to feed the world. In the face of unpredictable variables and the complexities of irrigation water management, your dedication to agriculture fuels our commitment to finding sustainable solutions.

To the agricultural community, who faces the challenge of balancing controllable and uncontrollable factors in irrigation, this work provides a comprehensive strategy. May the outcomes of this project contribute to increased crop yield and the conservation of precious water resources.

Our heartfelt gratitude goes to the farmers who generously shared their insights and experiences, helping us grasp the intricacies of irrigation water needs. Your wisdom has been the cornerstone of our pursuit for an effective and adaptable irrigation strategy.

We extend our appreciation to the researchers and practitioners in the field of agriculture, whose dedication inspires us. Your tireless efforts in developing solutions for efficient irrigation management have a lasting impact on global food security.

Special thanks to the mentors and advisors who guided us through the complexities of agricultural water management. Your expertise has been invaluable in shaping this project and its goals.

ACKNOWLEDGEMENT

This project has been a journey shaped by the collective efforts and support of individuals and institutions committed to addressing the challenges of irrigation water management in contemporary agriculture.

To our esteemed advisor, [**Dr. Sally Kassem**], [**TA. Yousra Ayman**] your unwavering guidance and insightful mentorship have been instrumental in shaping our understanding of this complex field. Your passion for research and commitment to excellence have inspired us to pursue innovative solutions in irrigation water management.

A heartfelt acknowledgment is extended to our mentors:

- **Dr. Hazem Sayed Mhawd:** Head of the Institute of Agricultural Engineering at the Agricultural Research Center (He is credited with allowing us to visit the institute's farm and learn about agricultural concepts).
- **Dr. Samar Mahmoud:** Senior researcher at the Institute of Agricultural Engineering (She has credited us with clarifying the concepts of machine learning, she also helped us by providing the data used in the project).
- **Dr. Aml Abo Elmagd:** Senior researcher at the Institute of Agricultural Engineering (She explained to us some of the most complex concepts related to agriculture in general and water management in particular. She also helped us understand the data and provided some really valuable insights).
- **Dr. Wafaa Mahmoud**: Researcher at the Institute of Agricultural Engineering. (She explained to the us the practical side of the data)
- **Eng. Abdelrahman:** Researcher at the Institute of Agricultural Engineering.
- **Dr. Engy Mosalam:** Researcher at the Institute of Agricultural Engineering (She helped us understand the general division of agricultural land, agricultural transactions, different types of irrigation, and most of the concepts related to agriculture and water management).

Their guidance has been instrumental in navigating us through the complex concepts of agricultural water management. Your expertise has illuminated our path and enriched the quality of our work.

To our colleagues and fellow researchers, your collaboration and exchange of ideas have been essential in refining the methodologies and approaches applied in this study.

We would like to recognize the researchers and practitioners in the field of agriculture for their dedicated efforts in advancing the understanding of irrigation strategies. Your contributions have provided a solid foundation for our exploration and innovation.

Table Of Contents

List Of Figures

Chapter 1 – Introduction

1.1 What is Irrigation

Irrigation is the artificial application of water to land to help grow crops, landscape plants, and lawns. It is a crucial agricultural practice that has been used for centuries to increase crop yields and support sustainable agriculture in regions with insufficient rainfall.

Irrigation's significance extends beyond mere water supplementation; it serves as a critical tool for mitigating climate-induced challenges, such as droughts and erratic rainfall patterns. As the world grapples with a growing population and the impacts of climate change, irrigation will undoubtedly remain an indispensable tool in safeguarding global food security.

Irrigation serves several vital purposes in agriculture and landscaping as; Increased Crop Yields, Supplemental Water Supply, Expanded Crop Cultivation, Improved Soil Quality, Reduced Soil Erosion, and many more other purposes.

1.2 What is Irrigation Management?

Irrigation management is the process of applying water to land to meet the water needs of crops and other agricultural plants. It is a critical component of agriculture, especially in areas with limited rainfall. Proper irrigation management can help to increase crop yields, improve water use efficiency, and reduce environmental impacts.

There are various irrigation methods employed to deliver water to crops and landscapes like Surface Irrigation, Sprinkler Irrigation, Drip Irrigation, Subsurface Irrigation. Our case study follows Drip Irrigation which is a water-efficient irrigation system that delivers water to plants directly at the root zone through emitters or drippers. This method helps to conserve water and nutrients by minimizing evaporation and runoff. Drip irrigation systems are commonly used in agriculture, horticulture, and landscaping.

1.2.1 What are Irrigation Management Goals?

Optimizing irrigation water is often an important goal of irrigation management. Optimizing irrigation water helps to achieve a number of benefits, including Increased water use efficiency, Reduced water costs, Reduced environmental impacts, Improved crop yields. In our study we will focus directly on Increasing water use efficiency and Reducing water costs.

Increased water use efficiency: Irrigation water is a valuable resource, and optimizing irrigation water can help to ensure that this resource is used efficiently. This can be achieved by using precision irrigation techniques, such as drip irrigation or subsurface drip irrigation, which deliver water directly to the root zone of plants and reduce evaporation and runoff.

Reduced water costs: The cost of water can be a significant expense for farmers and other water users. Optimizing irrigation water can help to reduce water costs by reducing the amount of water that is used. This can be achieved by using water-efficient irrigation techniques and by scheduling irrigation carefully to avoid watering during periods of peak demand.

1.3 Approaches for Optimizing irrigation water Problem:

To address inefficient use of irrigation water in Egypt, this graduation project aims to optimize irrigation water used while maintaining crop yield in Egypt in a specific Farm in El Nobaria, using real data from it. Following a certain approach by employing a Mathematical optimization techniques.

1.3.1 Approaches based on Mathematical optimization.

The choice of approach depends on the problem context, available data, and specific goals.

linear Programming (LP): Linear programming to optimize a linear objective function subject to linear constraints. Improving the amount of irrigation water for crops.

Dynamic Programming (DP): For sequential decision-making over time. Optimal irrigation scheduling across multiple seasons.

Goal Programming: For balancing multiple conflicting goals to arrive at satisfactory solutions. Balancing water conservation and crop productivity.

Stochastic Programming: For factors like weather conditions affect outcomes. Irrigation planning considering uncertain rainfall.

Mixed-Integer Linear Programming (MILP): For Handling both continuous and discrete decision variables.

1.3.2 Approaches based on Data Analysis.

Data preprocessing

Data Cleaning: For Detecting and fixing inaccurate, missing, or irrelevant data points. Remove duplicates, handle outliers, and address inconsistencies.

Feature engineering: We create new features that follow the same distribution of existing features by analyzing their charts to improve the model's performance.

simple regression models.

Based on the nature of the data, its fit to the data structure and Simplicity and Interpretability for understanding the correlation between the actual yield and each parameter in the actual data.

Multiple regression model (The chosen model)

For investigating, the relationship between three key agricultural factors and crop yield. The explanatory variables include dew point, irrigation amount (Irri), and water content in the effective root zone.

1.4 Problem Description:

Egypt, the cradle of civilization, is a land renowned for its fertile agricultural heritage. However, the country's arid climate poses significant challenges to sustainable agriculture, particularly in terms of water resource management. According to the **Fanack Water website**, updated in July 2023, the agricultural sector in Egypt requires an estimated 10 billion cubic meters (BCM) of water per year to meet demand. This figure is expected to rise to 13 BCM by 2025. **Past and future trends of Egypt's water consumption and its sources**: This research article published in Nature in 2021 states that Egypt's total water consumption is about 88 billion m^3 annually with agriculture accounting for 85% of this consumption. The inefficient use of irrigation water, coupled with the increasing demand for water from various sectors, has led to a critical water scarcity situation in Egypt. This alarming situation necessitates the development of innovative and sustainable irrigation water management strategies.

Chapter 2: Background & Existing work

2.1. Background introduction

Agriculture is a cornerstone of Egypt's economy, employing over 50% of the population and contributing significantly to the country's gross domestic product (GDP). With over 95% of Egypt's agriculture relying on irrigation, efficient water management is crucial for maximizing crop yield while conserving this precious resource. Despite its importance, Egyptian agriculture faces significant challenges in irrigation water management. Traditional methods, often based on estimations and historical practices, lead to inefficient water usage, resulting in water wastage and reduced crop productivity.

2.2 The problem of water scarcity and irrigation inefficiency:

Freshwater is a finite resource, and agriculture is the biggest user, consuming around 70% globally. This puts immense pressure on water supplies, especially with a growing population and changing climate. Inefficient irrigation practices can lead to water waste, salinization of soil, and depletion of groundwater reserves. This threatens the long-term sustainability of agricultural systems. Egypt is facing an annual water deficit of around 7 billion cubic meters. In fact, United Nations is already warning that Egypt could run out of water by the year 2025. According to **My Custom Essay expert**.

Traditional irrigation methods like flood irrigation, where water flows freely across fields, are a major contributor to water waste. These methods often suffer from:

- **Inefficient application**: Large amounts of water can evaporate before reaching the plant roots, leading to significant waste.
- **Uneven distribution**: Water may not reach all parts of the field uniformly, resulting in overwatering some areas and underwatering others.
- **Runoff and drainage**: Excess water can seep into the ground, potentially salinizing soil, or run off into waterways, causing environmental damage.

These inefficiencies translate to a significant loss of precious water resources. Estimates suggest that traditional irrigation methods can waste anywhere from **30% to 70%** of the applied water. This not only impacts crop yields but also contributes to environmental degradation and depletion of groundwater reserves.

Nowadays, Egypt's irrigation network draws almost entirely from the Aswan High Dam, which regulates more than 18,000 miles of canals and sub-canals that push out into the country's farmlands adjacent to the river. This system is highly inefficient, losing as much as 3 billion cubic meters of Nile water per year through evaporation and could be detrimental by not only intensifying water and water stress but also creating unemployment.

Water waste in agriculture isn't just a lost resource; it has a domino effect impacting various aspects of our food system and environment. Here's a breakdown of how water waste creates problems:

- **Soil Health and Salinization:** When large amounts of water are applied, they can percolate down through the soil profile. This can lead to a rise in the water table, bringing dissolved salts closer to the surface. As water evaporates from the soil surface, these salts concentrate, leading to a condition called salinization. Salinity disrupts the soil's delicate ecosystem, harming beneficial microorganisms and hindering plant nutrient uptake. This can lead to stunted plant growth, decreased crop yields, and eventually, soil becoming unfit for agriculture.
- **Crop yields:** When water is wasted, less is available for plant growth. Plants which are under water stress experience stunted growth, reduced photosynthesis, and ultimately produce lower yields.
- **Environmental sustainability:** Excessive water use depletes freshwater resources, impacting rivers, lakes, and groundwater reserves. This can have cascading effects on aquatic ecosystems and human water security. Excess irrigation water can leach essential nutrients like nitrogen and phosphorus from the soil. These can travel to waterways, causing algal blooms and disrupting aquatic ecosystems.

2.3 Existing work:

2.3.1 A new method based on machine learning to forecast fruit yield using spectrometric data: analysis in a fruit supply chain context:

A study published in **Precision Agriculture** (August 2022) which aims to propose a new method that leverages machine learning and spectrometric data to predict fruit yield. This is significant for the fruit supply chain as it allows for improved planning as many stages require advanced notice of at least two months.

2.3.2 Improvement of Irrigation Water Management Using Simulation Models and Artificial Intelligence Under Dry Environment Conditions in Egypt:

A research paper published in June 2023, the research provides insights into the potential benefits of leveraging simulation models and artificial intelligence in improving irrigation water management, addressing water scarcity challenges, and enhancing agricultural sustainability in Egypt. The paper's findings that these two new technologies can help with irrigation and water conservation. Farmers will benefit from the technology because it will help them achieve higher yields and a more consistent seasonal crop.

2.3.3 Optimizing Irrigation Water Management Practices to Improve Water Productivity:

A study published in the **Journal of Soil Sciences and Agricultural Engineering** (May 2007) investigated current knowledge on water resource conservation through modern irrigation systems and moisture regimes, and their impact on water productivity. The research identified that, for two consecutive growing seasons, the highest crop yields were achieved under drip irrigation with a 10% soil moisture depletion from available water. In contrast, furrow irrigation systems yielded the best results with a 25% soil moisture depletion level.

2.3.4 Optimization of Irrigation Water Allocation Framework Based on Genetic Algorithm Approach:

A study published in **Journal of Water Resource and Protection** (December 2020) developed a water allocation model using GA to equitably allocation available water to the various sectors in Kano River. The model has proven to be an effective tool in decision making tool for effective water allocation as the water allocation model yielded an optimal as well as equitable water release with a 96.44% demand met and an average relative supply of 0.94 was obtained indicating that there was even supply of water at all the sectors.

2.3.5 Linear Optimization Model for Efficient Use of Irrigation Water:

A research paper published in the **International Journal of Agronomy** (2017) acknowledges the challenge of efficiently using irrigation water in agriculture. This paper introduces a linear programming model as a tool to optimize water use. This model considers factors like crop water needs, precipitation, and available water supply. The model aims to allocate irrigation water strategically, potentially maximizing crop production or minimizing water waste.

2.3.6 Prediction of Irrigation Water Requirements for Green Beans-Based Machine Learning Algorithm Models in Arid Region:

A study published in **Water Resources Management** (March 2023) investigates the potential of machine learning for predicting irrigation water needs for green beans in dry areas. The research aims to assess how well machine learning models can forecast irrigation water requirements (IWR) for green beans. It explores different scenarios using various input factors, including weather data, crop characteristics, and soil properties. The findings identify the most effective combination of input factors and the best performing machine learning model for predicting green bean irrigation water needs in arid regions.

2.3.7 Using artificial neural networks to predict the reference evapotranspiration:

A study published in **Journal of Water and Land Development** (2022) where artificial neural network models (ANNs) were used to predict reference evapotranspiration (ETo) using climatic data from the meteorological station at the test station in Kafr El-Sheikh Governorate as inputs and reference evaporation values computed using the Penman-Monteith (PM) equation. These datasets were used to train and test seven different ANN models that included different combinations of the five diurnal meteorological variables used in this study, namely, maximum and minimum air temperature (T max and T min), dew point temperature (T dw), wind speed (u), and precipitation (P), how well artificial neural networks could predict ETo values.

2.3.8 USING ARTIFICIAL NEURAL NETWORKS MODELS FOR PREDICTING WHEAT YIELD PRODUCTIVITY:

A study published in **Arab Universities Journal of Agricultural Sciences** (August 2020) aimed to use three models of artificial neural networks (Feed Forward Neural Network (FFNN), Generalized Regression Neural Network (GRNN) and Radial-Basis Neural Network (RBNN)) in the field of wheat yield prediction. 27-year data for the period (1986-2012) were utilized to improve the models and four-year data (2013 and 2016) were used to estimate the models, to compare their outputs with the measured data.

Chapter 3 – Data & Case study

3.1 Why is data important?

Data collection is one of the most important and essential phases that must be considered. Operations research relies heavily on quantitative analysis using mathematical models and techniques. Without data, you essentially have no foundation to build your models on. A core part of operations research is identifying inefficiencies or areas for improvement in a system. Good data collection helps you pinpoint these problem areas and understand their scope.

The ultimate goal of operations research is to generate actionable insights that improve decision-making. Without data, you can't measure the impact of your proposed solutions or even know if they're truly addressing the problem. In simpler terms, good data collection is like having a strong foundation for your house. Without it, your entire project – the analysis, models, and recommendations – could be built on shaky ground.

3.2 Activities:

To gain a comprehensive understanding of core agricultural concepts and contribute meaningfully to this expansive field, we actively sought out expertise. This included engaging with agricultural professionals, researchers, and data collectors. Through these interactions, we were able to significantly deepen our knowledge and insights into the agricultural sector.

By fostering these relationships with industry leaders, we not only gained valuable knowledge, but also established a network of resources that will prove invaluable as we continue to explore and contribute to the field of agriculture. This ongoing collaboration ensures we stay abreast of the latest advancements and best practices, allowing us to make informed decisions and deliver impactful contributions.

3.2.1 Activity 1:

In the initial stages of our graduation project, we recognized the importance of real-world application. To ensure our project idea addressed a relevant agricultural need, we sought out expert feedback. We strategically chose to visit the **Agricultural Research Center in Giza**, a leading research institution in the Middle East. During our visit, we engaged in productive discussions with professors and researchers from various departments and institutes such as **Plant Protection Research Institute (PPRI), Agricultural Economics Research Institute (AERI), Agricultural Genetic Engineering Research Institute (AGERI),** and many other institutes.

These interactions at the **Agricultural Research Center** proved to be invaluable. The insights and perspectives gleaned from our conversations with the on-site experts significantly enhanced the development of our project proposal. Their guidance helped us refine our research question and ensure its practicality within the agricultural context. We believe this initial engagement with leading agricultural professionals positioned our project for success.

3.2.2 Activity 2:

At the end of the first activity, the researchers guided us to the **Agricultural Engineering Research Institute (AENRI)** and said that they are willing to help us as our project idea is very close to their research area. We went directly to **Dr. Hazem Mhawd (Head of the Agricultural Engineering Research Institute)** as he welcomed us warmly and we started to have a deep conversation with him about our idea and our objectives. He listened to us and started to give us great insights about what we already understand and what we need to study more, he helped us to formulate our project idea more effectively, set our boundaries and identify our scope of work.

Our brainstorm session with **Dr. Hazem** ended up in the best way for us as we were ready to make our goal specific, implementable, and reliable.

Then we continued our discussion with some of the most knowledgeable researchers in the institute; **Dr. Aml Abo Elmagd (Senior Researcher) and Eng. Abdelrhman (Researcher),** They were very helpful as they had great knowledge in the agricultural field, and provided us with a foundational understanding of key concepts critical to the successful execution of our project. This initial knowledge transfer ensured we possessed a strong base upon which to build our research endeavors. **Dr. Aml** took the responsibility of making sure that these concepts are now clear to us. At the end of this enriching activity, **Dr. Aml** suggested to us some research papers to read and study and come back to her once we understand these papers.

3.2.3 Activity 3:

Dr. Aml's provided materials proved instrumental in establishing a thorough understanding of the relevant agricultural concepts and definitions. We dedicated a significant amount of time, approximately one month, to diligently studying these resources. This in-depth exploration ensured we possessed a strong foundation in the theoretical underpinnings of our project's scope of work. Then we took another visit to the **Agricultural Engineering Research Institute (AENRI)** and met with **Dr. Aml** and with another respected researcher **Dr. Wafaa Mahmoud (Researher),** Following a productive exchange of ideas, we transitioned to the crucial stage of data acquisition.

Through collaborative discussions, a focus on **olive** cultivation emerged. Recognizing our project's alignment with their expertise, **Dr. Aml** generously offered to share a dataset pertaining to olive crops from a functioning farm managed by the institute. Leveraging a real-world dataset significantly enhanced the feasibility and potential impact of our project. The data provided by **Dr. Aml** offered a unique opportunity to analyze and interpret information directly tied to agricultural practices. This access to real-world data allowed us to move beyond theoretical concepts and delve into the practical applications within the olive cultivation domain.

3.2.4 Activity 4:

Having established a comprehensive understanding of our project scope and readiness to receive the olive crop dataset, we expressed our desire to gain firsthand experience with real-world data collection. This included exploring the process of data acquisition and delving deeper into the significance of each parameter crucial to our forthcoming analyses. So, **Dr. Hazem** settled an appointment for us to visit an olive farm which was owned by **Al Salam International for Development & Agricultural Investment** and the institute was already running a field experiment in that farm. Our farm visit facilitated extensive discussions with the data collection team **(Dr. Wafaa Mahmoud & Dr. Engy Mosalam)**. We actively engaged by posing a series of insightful questions, including the farm's irrigation system, the specific sensor technology employed for climatic data acquisition, and unique characteristics of the olive cultivar being studied.

This interactive experience not only equipped us with essential knowledge but also served to validate our existing understanding of the subject matter. Furthermore, by witnessing the data collection process firsthand, we gained valuable insight into the practical considerations and potential challenges involved. This on-site observation allowed us to appreciate the complexities of agricultural data acquisition and the meticulous attention to detail required for accurate measurements. This newfound appreciation strengthens the foundation of our project and fosters a deeper understanding of the real-world implications of our research. At the end of this visit, we had the opportunity to meet **Dr. Mohammed Mahmoud** (Owner of Al Salam Internation for Development & Agricultural Investment) who is a very knowledgeable and experienced man as he appreciated our efforts and gave us some extra information about the nature of olive cultivation, he also imparted invaluable wisdom to his peers and successors. Emphasizing the ethos of humility and gratitude, he advocated for a culture of sharing knowledge and welcoming opportunities to uplift others.

3.2.5 Activity 5:

After the olive farm visit, we had a call from a senior researcher at the **Agricultural Engineering Research Institute (AENRI) Dr. Samar Mahmoud** who wanted to meet with us in order to discuss some thoughts about the application of our graduation project on the Olive crop. We settled an appointment with her in the presence of **Dr. Aml Abo Elmagd** and we started to explain our project to **Dr. Samar** as she wanted to hear it directly from us. Dr. Samar was deeply involved in the field of computer science and the integration between computer science and agriculture, so she gave us more insights about the applicability of our project as she was convinced that olive is a really difficult crop to be studied as it adapts to very difficult circumstances, and we wouldn't be able to get trends from any dataset concerned with the olive crop. She elaborated with more agricultural details why olive crop is one of the most difficult crops to be addressed and also why would our results be useless if we used a dataset about olive, so she suggested another crop that we can use in our study which will help us more to reach our objectives. She offered us a dataset containing data about a tomato farm located in **Nubaria** where Dr. Samar was doing a field experiment for about 17 years as she was collecting and studying real data about tomato. We accepted her suggestion and we started to study more about the nature of tomato crop and spent a few weeks trying to understand the nature of tomato crop.

3.2.6 Activity 6:

Following receipt of the tomato dataset from **Dr. Samar**, we embarked on a thorough exploration of the data. This in-depth analysis aimed to achieve a comprehensive understanding of the included parameters and to extract valuable insights that would facilitate our effective data utilization. To further enhance our comprehension of specific data points and to identify potential trends within the dataset, we conducted a series of online meetings with **Dr. Aml**. These collaborative discussions proved instrumental in maximizing the utility of the provided data. Through this meticulous data exploration process, we were able to identify potential areas for data cleaning and pre-processing. This initial data preparation stage ensured the data's quality and consistency, ultimately paving the way for robust and reliable statistical analyses. Dr. Aml's guidance during this phase proved invaluable, as her expertise helped us navigate the intricacies of data manipulation and prepare the dataset for subsequent phases of our project.

3.3 Case study:

3.3.1 Introduction:

Our case study focuses on the Castle Rock hybrid tomato crop cultivated in an agricultural land situated in **Nubaria**. The research spans over a period of 17 years, from 2000 to 2016, with data collection specifically conducted during the summer season. The chosen planting date for the Castle Rock hybrid tomatoes is mid-March (3/15) each year.

3.3.2 Location:

The application was applied in an agricultural land situated in **Nubaria**, (lat: 30.67, long: 30.08, alt: 38) known for its unique climatic and soil conditions, serves as the backdrop for our study.

3.3.3 Soil texture:

Our soil texture is ideal for agriculture because it is Sand Clay Loam, which is a mixture of sand and clay, is considered ideal soil for agriculture as it combines the benefits of each soil type: good drainage and excellent nutrient retention.

3.3.4 Crop variety:

The primary focus of our investigation is on the **Castle Rock hybrid tomato**, a cultivar recognized for its resilience, its resilience can be attributed to their genetic makeup, which includes traits that make them less susceptible to common tomato diseases and pests. This hybrid has gained popularity among farmers for its promising features.

3.3.5 Data collection methodology:

Data collection commenced in the year 2000 and continued annually until 2016. The meticulous recording of various parameters includes growth patterns, yield per plant, disease resistance, and overall crop health. Planting activities were consistently initiated on the 15th of March each year, aligning with the onset of the summer season in **Nubaria**.

3.3.6 Irrigation system:

The chosen irrigation system for the Castle Rock hybrid tomato crop in the Nubaria agricultural land is **drip irrigation system**. This method involves the controlled delivery of water directly to the root zone of plants, minimizing water wastage and optimizing water use efficiency. Drip irrigation is known for its precision and adaptability to various crop types, making it a suitable choice for the Castle Rock hybrid in the specific climatic and soil conditions of Nubaria.

3.3.7 Tomato season duration:

The period from planting to harvest, known as the tomato season, spans 165 days (S = 165). This duration serves as a crucial factor in determining the irrigation schedule and water requirements throughout the crop's growth cycle.

3.3.8 Period between irrigations (T):

The time interval between successive irrigations ranges from 1 to 4 days (T = 1-4). This variable allows for flexibility in adjusting the irrigation schedule based on the specific needs of the Castle Rock hybrid tomatoes, considering factors such as weather conditions, soil moisture levels, and growth stages. The variability in the watering period ensures adaptability to the dynamic environmental conditions in **Nubaria**, promoting efficient water management throughout the tomato season.

3.3.9 Average Amount of Irrigation Water per Watering Day (h):

The irrigation system provides an average amount of water ranging from 2 to 20 mm (the depth of water applied) on each watering day (h = 20). This variable allows for flexibility in adjusting water supply based on the crop's specific needs and environmental conditions.

3.3.10 Drip irrigation system specifications:

- Leaching Requirements (LR): The leaching requirements for the drip irrigation system are set at LR = 1.2. Leaching helps prevent the accumulation of salts in the root zone, ensuring a healthy and productive growth environment for the Castle Rock hybrid tomatoes.

-Irrigation Efficiency (IE): The irrigation efficiency of the drip system is established at IE = 0.9. This parameter represents the system's effectiveness in delivering water to the plants' root zones, considering losses due to evaporation, runoff, and other factors.

3.4 Parameters & Columns:

Our comprehensive case study on the Castle Rock hybrid tomato crop in Nubaria is structured into five main sections, each meticulously curated to capture and analyze the various factors influencing the target crop's performance.

3.4.1 Climatic parameters:

1- Temperature (°C): Recording the daily (Max & Min) temperatures to understand the thermal conditions experienced by the Castle Rock hybrid tomatoes throughout their growth cycle.

2- Dew Point: Monitoring the dew point, which provides insights into humidity levels and potential moisture stress on the crops.

3- Wind (m/s): Measuring wind speed to assess its impact on transpiration, evaporation, and overall crop health.

4- Eto (mm/day): Calculating reference evapotranspiration to gauge the water demand from the atmosphere, aiding in irrigation planning.

Location	lat	long	alt	YEAR	MO	DA	Temp_max [C°]	Temp_min [C°]	Dew Point (°C)	Wind (m/s)	Eto [mm/day]
Nobaria	30.67	30.08	38	2000	3	15	21	10	7	1.7	3.1
				2000	3	16	24	11	7	3.1	4.1
				2000	3	17	32	12	2	3.9	7.8
				2000	3	18	30	17	3	4.4	7.4

Figure 1: Climatic data

3.4.2 Soil parameters:

1- Soil Texture: Identifying the specific texture of the soil in the cultivation area, influencing water retention, drainage, and nutrient availability.

2- F.C % (Field Capacity): Determining the field capacity percentage, indicating the maximum amount of water the soil can hold against gravity after excess water has drained away.

3- Available Water %: Assessing the percentage of water available to plants within the soil, a critical factor in determining irrigation needs.

Depth	Soil texture			Class	Bulk Density	F.C. %	P.W.P %	Available water %
[cm]	Clay %	Silt %	Sand %		[g/ cm^3]			
0-10	27.19	20.53	52.28	Sandy clay loam	1.48	19.52	9.43	10.90
10-20	23.50	21.22	55.28	Sandy clay loam	1.48	19.86	10.15	9.71
20-30	25.3	22.22	52.48	Sandy clay loam	1.48	21.56	12.23	9.3

Figure 2 : Soil data

3.4.3 Soil-Water content:

1- Days after planting (DAP): Tracking the progression of the Castle Rock hybrid tomatoes' growth by recording the number of days since planting.

2- Irri (mm): Documenting the irrigation applied in millimeters, offering insights into the water management practices employed.

3- Actual Evapotranspiration (ET) [mm]: Quantifying the actual water consumption by the crops, considering both transpiration and evaporation.

4- Water content in the effective root zone (Wr) [mm]: Measuring the moisture content within the root zone, a key indicator of the water availability for the Castle Rock hybrid tomatoes.

Day	Month	Year	DAP	Rain [mm]	Irri [mm]	Actual evapotranspiration (ET) [mm]	Water content in the effective root zone (Wr) [mm]
15	3	2000	1	0	2	1	33.4
16	3	2000	2	0	0	0.8	31.9
17	3	2000	3	0	0	1.2	30.6
18	3	2000	4	0	3	2.2	31.4
19	3	2000	5	0	0	1.2	30.2
20	3	2000	6	0	0	0.8	29.4

Figure 3:Soil Water content data

3.4.4 Irrigation:

Depth (mm): Specifying the depth of irrigation applied, providing essential information for understanding water penetration and distribution in the soil.

Day	Depth (mm)
1	2
4	3
8	3
11	4
15	4
18	5
22	5
25	6
29	6
32	7

Figure 4:Irrigation data

3.4.5 Actual Yield:

Documenting the actual yield of Castle Rock hybrid tomatoes from the year 2000 to 2016, serving as the ultimate metric to evaluate the success and productivity of the cultivation practices over the study period.

Year	Actual [ton/ha]
2000	2.54
2001	2.98
2002	2.81
2003	2.55
2004	2.52
2005	2.48
2006	2.31
2007	2.26
2008	2.24
2009	2.24
2010	2.19
2011	2.16
2012	2.11
2013	2.11
2014	2.01
2015	1.91

Figure 5 :yield data

3.5 Data cleaning:

Is the process of identifying and correcting (or removing) errors and inconsistencies in data to improve its quality as the quality of the data directly impacts the quality of the analysis and the results derived from it. This process is essential in ensuring that data is accurate, complete, and reliable for analysis and decision-making.

The data cleaning process for this project involved the following steps:

1. Removing Duplicates: Duplicates can distort analysis and lead to incorrect conclusions. We identified and removed duplicate records from the dataset to ensure each entry was unique.

2. Handling Missing Data: Missing data can bias the results if not handled properly. We addressed missing values through:

Imputation: Filling in missing values with mean, median, or mode.

Removal: Deleting records with missing critical data points.

Prediction: Using algorithms to predict and fill in missing values where appropriate.

3. Correcting Errors: Errors such as typos, misspellings, and incorrect formatting were identified and corrected. This involved:

4. Standardizing Data: To ensure consistency, data was standardized to follow uniform formats. This included:

Text Formatting: Ensuring consistent capitalization and removing unwanted characters.

Units of Measurement: Standardizing units for numerical data.

5. Filtering Outliers: Outliers can skew results and affect the validity of the analysis. We identified and handled outliers through:

Statistical Methods: Using statistical tests to identify outliers.

Domain Knowledge: Leveraging domain expertise to decide whether to remove or adjust outliers.

6. Validating Data: Data validation ensures the dataset adheres to defined rules and constraints. We validated data by:

Checking Data Types: Ensuring data types were consistent with expectations.

Range Validation: Confirming that values fell within acceptable ranges.

Uniqueness Constraints: Verifying that unique identifiers were indeed unique.

7. Reformatting Data: Data was transformed into a suitable format for analysis. This included:

Parsing Dates: Converting date strings into date objects.

Splitting/Merging Columns: Splitting composite columns into individual elements or merging related columns.

Converting Data Types: Ensuring numerical values were stored in the correct numeric format.

Chapter 4: Mathematical Model development

4.1 Introduction to Regression

Regression modeling Regression modeling is a statistical technique used to examine the relationship between a dependent variable (often called the outcome or response variable) and one or more independent variables (predictors or explanatory variables). The primary goal is to model this relationship so that the dependent variable can be predicted based on the values of the independent variables.

4.1.1 Applications of Regression Modeling:

- Predictive Analysis: Forecasting future trends, such as sales, stock prices, or weather.
- Identifying Relationships: Understanding the impact of several factors on an outcome.
- Optimization: Finding the best conditions for a desired outcome, like maximizing efficiency or minimizing costs.
- Risk Management: Assessing the likelihood of risks and their potential impact.

4.1.2 Key Types of Regression Models:

Linear Regression, Non-Linear Regression, Logistic Regression, Ridge and Lasso Regression, Polynomial Regression, and Stepwise Regression.

Linear Regression:

- Simple Linear Regression: Involves one dependent variable and one independent variable. It fits a straight line (linear relationship) to the data.
- Multiple Linear Regression: Involves one dependent variable and multiple independent variables. It fits a linear plane to the data.

4.2 Why Regression approach?

4.2.1 Nature of the Data and Problem:

Relationship Exploration: Explain that regression modeling is well-suited for understanding and quantifying the relationships between dependent and independent variables. One of our main purposes was to find out how different factors influence the yield.

Predictive Capability: Regression models are effective for making predictions.

4.2.2 Suitability for the Data Structure:

Data Characteristics: Both dependent and predictor variables are continuous, while looking for linear relationships, regression models is appropriate. Mention if the dependent variable is continuous and if the predictors are either continuous or categorical.

4.2.3 Simplicity and Interpretability:

Ease of Interpretation: Regression models, especially linear regression, provide coefficients that are easy to interpret, making it straightforward to understand the impact of each predictor on the outcome.

Transparency: Compared to more complex models like neural networks, regression models are more transparent and easier to explain to stakeholders.

4.3. The Regression Model

4.3.1 Simple regression models

Steps followed to build the simple regression models:

1- Compute the correlation between the explanatory variable and the dependent variable.
2- A scatterplot between the explanatory variable and the dependent variable.
3- Check the following assumptions: **Linearity – Normality of errors – Homoscedasticity – Interdependence of errors**.
4- Build the regression model.
5- Check the significance of the model .

4.3.1.1 Temp_Max and Actual yield:

1- Correlation = 0.2 (weak positive correlation).

2-

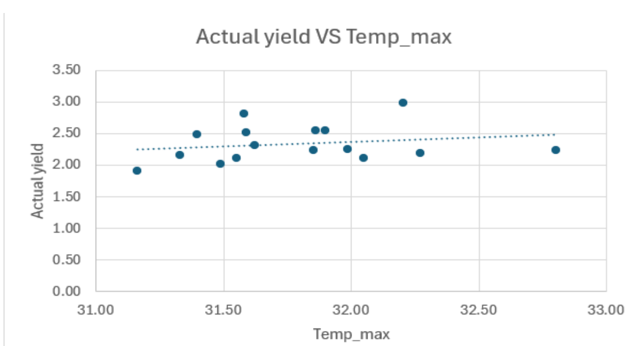

Figure 6 : Temp_Max and Actual yield Scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.204706073							
R Square	0.041904576							
Adjusted R Square	-0.026530811							
Standard Error	0.291369417							
Observations	16							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	0.051983876	0.051983876	0.612323216	0.446948814			
Residual	14	1.188545922	0.084896137					
Total	15	1.240529798						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-2.171797746	5.762166449	-0.376906458	0.711894466	-14.53041564	10.18682015	-14.53041564	10.18682015
Temp_max [C°]	0.14182012	0.18123726	0.782510841	0.446948814	-0.246895141	0.530535382	-0.246895141	0.530535382

Figure 7 :Temp_Max and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.4, α = 0.05).

4.3.1.2 Temp_Min and Actual yield:

1- Correlation = -0.2 (weak negative correlation).

2-

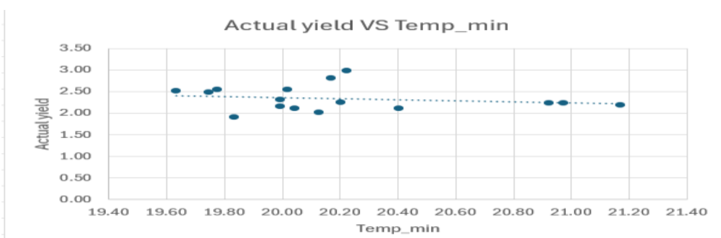

Figure 8:Temp_Min and Actual yield scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.188426386							
R Square	0.035504503							
Adjusted R Square	-0.033388033							
Standard Error	0.292340971							
Observations	16							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	0.044044394	0.044044394	0.515360665	0.484635381			
Residual	14	1.196485404	0.085463243					
Total	15	1.240529798						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	4.747353195	3.358644142	1.413473114	0.179364984	-2.45622205	11.95092844	-2.45622205	11.95092844
Temp_min [C°]	-0.119324459	0.166216387	-0.717886248	0.484635381	-0.475823153	0.237174236	-0.475823153	0.237174236

Figure 9:Temp_Min and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.5, α = 0.05).

4.3.1.3 Dew Point and Actual yield:

1- Correlation = 0.7 (strong positive correlation).

2-

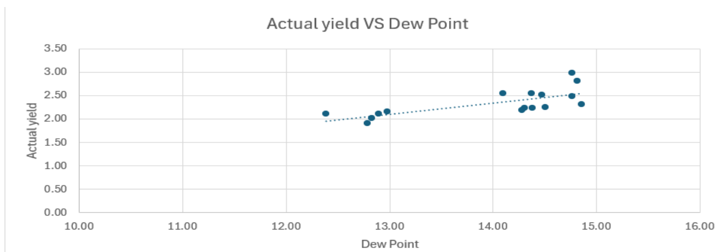

Figure 10: Actual yield & Diew point scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.720454364							
R Square	0.519054491							
Adjusted R Square	0.48470124							
Standard Error	0.206436991							
Observations	16							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	0.643902563	0.643902563	15.10932681	0.001643044			
Residual	14	0.596627235	0.042616231					
Total	15	1.240529798						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1.014707772	0.863762271	-1.174753524	0.259688115	-2.867293593	0.837878049	-2.867293593	0.837878049
Dew Point (°C)	0.239939709	0.061727625	3.887071752	0.001643044	0.10754712	0.372332298	0.10754712	0.372332298

Figure 11: Dew point & Actual yield regression model

5- The model is significant as the significance F value is less than the α value (Significance F = 0.002, α = 0.05).

4.3.1.4 Wind speed and Actual yield:

1- Correlation = 0.05 (weak positive correlation).

2-

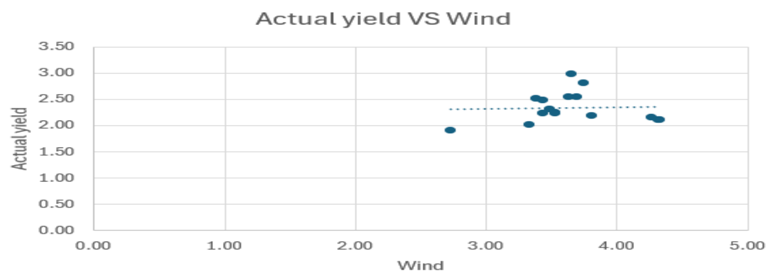

Figure 12 : Wind speed and Actual yield Scatter plot

3- All the assumptions are verified except for the homoscedasticity as the residual plot shows heteroscedasticity.

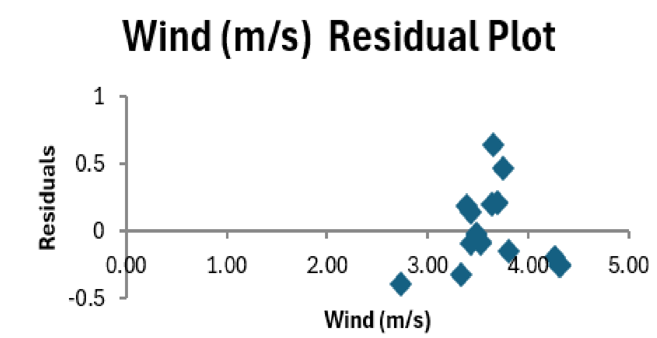

Figure 13: wind residual plot

4-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.046310714
R Square	0.002144682
Adjusted R Square	-0.069130698
Standard Error	0.297353716
Observations	16

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.002660542	0.002660542	0.030090085	0.86476908
Residual	14	1.237869256	0.088419233		
Total	15	1.240529798			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2.2174343	0.692127665	3.203793768	0.00637228	0.732968098	3.701900503	0.732968098	3.701900503
Wind (m/s)	0.032772766	0.188930204	0.173464938	0.86476908	-0.372442219	0.437987751	-0.372442219	0.437987751

Figure 14:Wind speed and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.9, α = 0.05).

4.3.1.5 Eto and Actual yield:

1- - Correlation = -0.3 (weak negative correlation).

2-

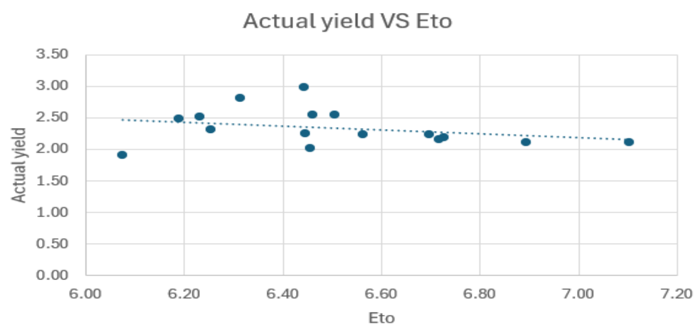

Figure 15: Eto and Actual yield scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.290860018
R Square	0.08459955
Adjusted R Square	0.019213804
Standard Error	0.284803383
Observations	16

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.104948263	0.104948263	1.29385309	0.274432767
Residual	14	1.135581535	0.081112967		
Total	15	1.240529798			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	4.328897557	1.75277757	2.469735824	0.026998106	0.569563558	8.088231555	0.569563558	8.088231555
Eto [mm/day]	-0.306265274	0.269249728	-1.137476633	0.274432767	-0.883748507	0.271217959	-0.883748507	0.271217959

Figure 16:Eto and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.3, α = 0.05).

4.3.1.6 ETc and Actual yield

1- Correlation= 0.13 (weak positive correlation)

2-

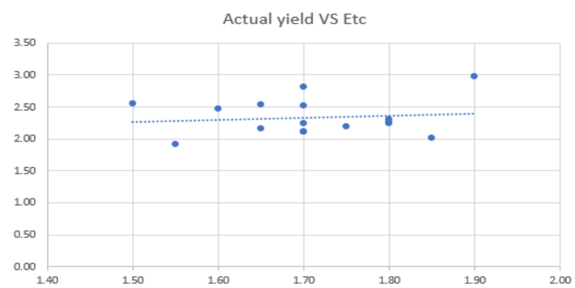

Figure 17:ETc and Actual yield scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.125432248
R Square	0.015733249
Adjusted R Square	-0.054571519
Standard Error	0.295322128
Observations	16

ANOVA

	df	SS	MS	F	gnificance F
Regression	1	0.019517564	0.019517564	0.223786	0.643461
Residual	14	1.221012234	0.08721516		
Total	15	1.240529798			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	pper 95.0%
Intercept	1.759457762	1.222670679	1.439028344	0.172124	-0.86291	4.381826	-0.86291	4.381826
Median/ Actual evapotranspiration (ET) [mm]	0.337750293	0.71396828	0.473060643	0.643461	-1.19356	1.86906	-1.19356	1.86906

Figure 18 : ETc and Actual yield regression model

6- The model is not significant as the significance F value is greater than the α value (Significance F = 0.6, α = 0.05).

4.3.1.7 Wc and Actual yield:

1- Correlation= 0.17 (weak positively correlation)

2-

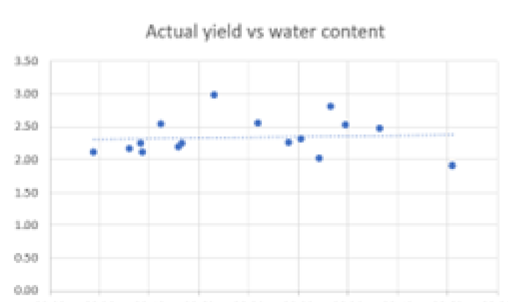

Figure 19: Water content and Actual yield scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.171117603
R Square	0.029281234
Adjusted R Square	-0.040055821
Standard Error	0.293282598
Observations	16

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.036324243	0.036324243	0.422302822	0.52631015
Residual	14	1.204205555	0.086014682		
Total	15	1.240529798			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Ipper 95.0%
Intercept	-0.69566797	4.666999835	-0.149061066	0.883631356	-10.70538709	9.314051	-10.70538709	9.314051
Water content in the effective root zone (Wr) [mm]	0.097330957	0.149774888	0.649848307	0.52631015	-0.223904228	0.418566	-0.223904228	0.418566

Figure 20 :Water content and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.5, α = 0.05).

4.3.1.8 Irrigation and Actual yield:

1- correlation = -0.15 **(weak negatively correlation)**

2-

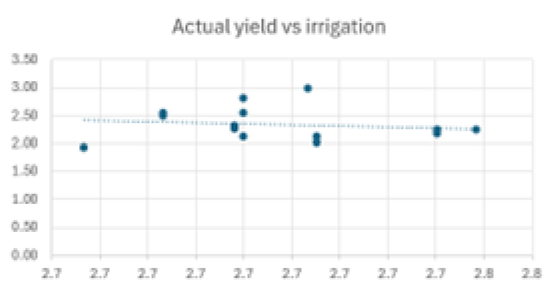

Figure 21 : Irrigation and Actual yield scatter plot

3- All the assumptions are verified.

4-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.154359763
R Square	0.023826936
Adjusted R Square	-0.045899711
Standard Error	0.294105396
Observations	16

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.029558025	0.029558025	0.341719233	0.568140037
Residual	14	1.210971773	0.086497984		
Total	15	1.240529798			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	7.293410193	8.479425153	0.860130264	0.404206518	-10.89314799	25.47996838	-10.89314799	25.47996838
Irri [mm]	-1.830463823	3.131312695	-0.58456756	0.568140037	-8.546461606	4.885533961	-8.546461606	4.885533961

Figure 22 : Irrigation and Actual yield regression model

5- The model is not significant as the significance F value is greater than the α value (Significance F = 0.5, α = 0.05).

4.3.2 The final regression model

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.723624006
R Square	0.523631702
Adjusted R Square	0.450344272
Standard Error	0.213207928
Observations	16

ANOVA

	df	SS	MS	F	Significance F
Regression	2	0.64958073	0.324790365	7.144904649	0.008065449
Residual	13	0.590949068	0.045457621		
Total	15	1.240529798			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-1.325057547	1.251764016	-1.058552195	0.309077115	-4.029329292	1.379214199	-4.029329292	1.379214199
Dew Point (°C)	0.16577977	0.219301593	0.755944211	0.463161335	-0.307992519	0.639552058	-0.307992519	0.639552058
Irri [m3/Hec]	0.000324343	0.000917706	0.35342782	0.729433224	-0.00165824	0.002306925	-0.00165824	0.002306925

Figure 23 : The final regression model

This study employs a multiple regression model to investigate the relationship between three key agricultural factors and crop yield. The explanatory variables include dew point, irrigation amount (Irri), and water content in the effective root zone.

The results of the regression analysis indicate a statistically significant effect of these variables on actual yield. This is evidenced by a significant F-statistic, which implies that the model as a whole explains a statistically relevant portion of the variance in yield after accounting for the effects of chance (significance F value (0.01) less than α value (0.05)).

In simpler terms, dew point, and irrigation amount are important contributors to crop yield. The statistically significant model suggests that these factors have a combined effect that is not attributable to random variation.

The final equation provided from this model is:

$$Y = -1.8 + (0.002 * Average\ of\ Dew\ Point) + (0.0009 * Total\ Irrigation\ Water)$$

4.4 Important definitions and when to apply irrigation:

This section presents a selection of key definitions central to the mathematical model, including both those explicitly referenced and those implicitly underlying its structure. A clear understanding of these concepts and definitions is essential for a thorough grasp of the mathematical model's underlying logic. The following concepts will concentrate on when to apply irrigation based on root zone water depletion.

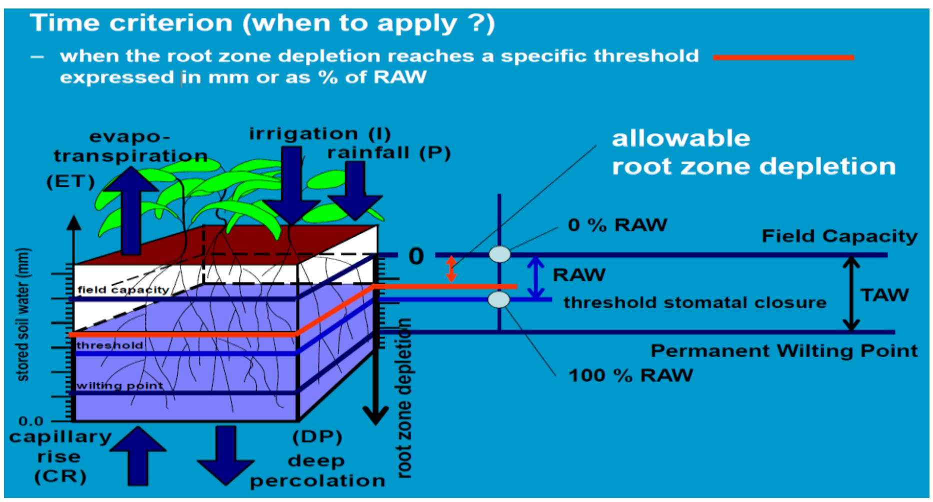

Figure 24 : Time criterion (when to apply?)

4.4.1 Thresholds in Soil Moisture:

Field capacity (FC): The maximum amount of water the soil can hold after excess water has drained away. Let's denote this as **FC**.

Wilting point (WP): The minimum amount of water the soil can hold before the plants start to wilt. Let's denote this as **WP**.

Total Available water (TAW or PAW): This is the maximum amount of water that soil can store and make available for plant use**.** It is essentially the difference between field capacity **(FC)** and wilting point **(WP)**. Let's denote this as **TAW** or **PAW**.

Readily Available water (RAW): This refers to the portion of (**TAW)** that is easily used by the crop before water stress affects the plant to avoid water logging or stressing the plants. It can also be defined as the water held between (**FC)** and a nominated refill point. Let's denote this as **RAW.**

Water content in the effective root zone (WC): The amount of water present in the soil. Let's denote this as **WC**.

Reference evapotranspiration (Eto): The amount of water lost by the plant through two separate operations, evaporation, and transpiration. Let's denote this as **Eto**.

Crop evapotranspiration (ETc): The water requirement of the crop. Let's denote this as **ETc**.

Leaching requirements (LR): The fraction of the irrigation water that must be percolated out of the bottom of the root zone in order to prevent average soil salinity from rising above some specifiable level. Let's denote this as **LR**.

Irrigation efficiency (IE): The ratio of the amount of water used for consumptive crop needs and to maintain the salt balance, to the total volume of water diverted. Let's denote this as **IE**.

Threshold stomata closure: This is the point at which the soil moisture level reaches a critical limit, beyond which plants begin to experience water stress. It is often expressed as a percentage of the Readily Available Water **(RAW)** or in millimeters of water depletion from the root zone.

Irrigation zone: The area above the threshold line indicates the range of soil moisture where plants have sufficient water and are not stressed, thus irrigation is not required. Conversely, the area below the threshold line indicates that the soil moisture has dropped to a level where plants may begin to experience stress, and irrigation may be necessary.

In summary, Management allowable depletion **(MAD)** specifies the maximum amount of soil water the irrigation manager allows the crop to extract from the active rooting zone between irrigations. These thresholds are crucial for irrigation scheduling to ensure crops receive adequate water without over-irrigation.

4.4.2 Maximum allowable depletion or deficit (MAD):

Management allowable depletion specifies the maximum amount of soil water the irrigation manager chooses to allow the crop to extract from the active rooting zone between irrigations. Only the crop easily uses a portion of the available water holding capacity before crop water stress develops.

Capillary rise (CR): This refers to the movement of water upwards through the soil due to capillary action. It can contribute to maintaining soil moisture above the threshold. Let's denote this as **CR.**

Deep percolation (DP): This occurs when water moves down through the soil profile, beyond the reach of plant roots. It represents a loss of water that could have been used by the plants.

4.5 Preliminary Mathematical Model:

Decision Variables:

$x_i = irrigation\ water\ per\ day\ i$

$b_i = binary\ variable \qquad b_i \in \{0, 1\}$

Objective function:

$$Min\ Z = \sum_{i=1}^{S} x_i b_i$$

Subject to:

1. $WP < AW_i \leq FC$, where $AW_i = WC_i + x_{i-1} * b_{i-1} - (ETC_{i-1} + ETO_{i-1})$

2. $b_i = \begin{cases} 1, & WP + e < AW_i \leq Avg_{\{WP,FC\}} \\ 0, & otherwise \end{cases}$

3. $x_i \leq h$

4. $x_i b_i = \frac{LR + ETc_i}{IE}$

5. $Y \geq Actual$ (Under Study)

Figure 25:Preliminary Mathematical Model

.

Where:

$x_i = irrigation\ water\ per\ day\ i.$

i = day number per season y   i= {1, 2, ... , S}

S = number of days in season.

b = binary variable. $b_i \in \{0,1\}$

h = max permissible amount of irrigation water.

LR = Leaching Requirements.

IE = Irrigation Efficiency.

$e = Threshold.$

Y = Crop yield.

Breakdown of the preliminary mathematical model:

4.4.1 Decision variables:

1. x_i**:** The amount of irrigation water applied on day **i**, where **i** is the number of the day per season. And **S** is the total number of days in the season.
2. b_i**:** A binary variable which decides whether there will be irrigation in day **i** or not, where b_i = 1 indicates there will be irrigation and b_i = 0 indicates there will be no irrigation.

4.4.2 Objective Function:

$Min\ Z = \sum_{i=1}^{S} x_i b_i$

The primary goal is to optimize water resource allocation by minimizing total seasonal irrigation water usage. To achieve this, the objective function employs a binary decision variable (**b_i**) for each irrigation event. The product of this variable and the corresponding irrigation water quantity (**x_i**) is then summed across the entire season, ensuring that only actual irrigation events contribute to the total water usage.

4.4.3 Constraints:

1. This constraint ensures that the available water in the soil in day i (**AW_i**) can't exceed the field capacity (**FC**) in order not to waste water or be less than the wilting point (**WP**) in order not to let the plant die from the lack of water. This constraint also shows how to compute the (**AW_i**) as an equation in the water content in the effective root zone (**WC_i**), reference evapotranspiration (**Eto**) and actual evapotranspiration (**ETc**).
2. The second constraint decides when to apply irrigation, and the condition to apply irrigation in day i is that the available water in day i (**AW_i**) is in the range between the wilting point (**WP**) summed to threshold (e) and the average of the field capacity and wilting point, where the threshold (e) will be decided based on the optimization process that will decide the best value for the threshold.
3. This constraint ensures that the quantity of irrigation water in day i (**x_i**) is less than or equal to the maximum permissible amount of irrigation water (**h**).
4. The fourth constraint was suggested by agricultural experts which proposed the computation of the quantity of irrigation water as the summation of leaching requirements (**LR**) and actual evapotranspiration (**ETc**) divided by the irrigation efficiency (**IE**).
5. This constraint is written in the model in order to make sure that the outputted crop yield (**Y**) is at least equal to the actual crop yield and it's still under study because there is no direct relationship between the quantity of irrigation water applied during the season and the crop yield.

4.5 Final Mathematical Model:

Decision variables:

$$x_i = \text{Quantity of irrigation water in day } i$$

Objective function:

$$Min\ \ Z = \sum_{i=1}^{s} x_i$$

Constraints:

1. $x_i = \begin{cases} \%RAW & , if\ \ WC_i \leq \theta \\ 0 & , Otherwise \end{cases}$

2. $0 \leq x_i \leq MAD$

3. $Y = -1.3 + (0.2ADP) + (0.003Z)$

4. $Y \geq Actual\ Yield$

Figure 26: final mathematical model

Where:

$x_i = Quantity\ of\ irrigation\ water\ in\ day\ i.$

$i = Day\ number. \quad i = \{1, 2, ..., S\}$

$S = Number\ of\ days\ in\ the\ season.$

$Z = Total\ quantity\ of\ irrigation\ water\ in\ the\ season.$

$RAW = Readily\ available\ water.$

$WC_i = Water\ content\ in\ the\ effective\ root\ zone\ in\ day\ i.$

$\theta = Specific\ percentage\ of\ RAW + (TAW - RAW)$

$MAD = Maximum\ allowable\ depletion\ (FC - \theta).$

$TAW = Total\ Available\ water.$

$FC = Field\ capacity.$

$Y = Yield.$

$ADP = Average\ Dew\ Point.$

Figure 27:Mthematical model parameters

Breakdown of the final mathematical model:

4.5.1 Decision variables:

1. **x_i:** The amount of irrigation water applied on day i, where **i** is the number of the day per season. And **S** is the total number of days in the season.

4.5.2 Objective Function:

$$Min\ Z = \sum_{i=1}^{S} x_i$$

The objective function aims to minimize the total irrigation water used over the season (**S**) which is represented by the summation (Σ) from **i** = 1 to **S** of x_i.

4.5.3 Constraints:

1. This constraint determines whether to irrigate at day i or not (When to apply?). When water content goes below or even reaches Threshold stomata closure (100% Raw point), here soil moisture level reaches a critical limit, at which the stomata of plants close to conserve water which cause the plant to begin to experience water stress. The plant starts suffering while absorbing water. So, when water content reaches this point, we might irrigate. $WC_i \leq (TAW - RAW)$ So, we not only have to prevent it reaching to this point but also, we have to keep it away by a safe zone, this made us use a percentage in the constraint $WC_i \leq \theta , where \ \theta = \% * RAW + (TAW - RAW)$. Also, this constraint determines the amount of water to irrigate as if the constraint is satisfied (How much to apply?). We follow the (back to field capacity) Depth criteria, which means that we will substitute the water that has been lost. Water irrigation amount = (field capacity – water content).

2. Irrigation Limit: This constraint limits the irrigation application (xi) to be between a zero and a maximum value of MAD, where MAD =FC - θ. The minimum value ensures that the condition cannot be applied, and irrigation does not occur, and the maximum value ensures that irrigation does not exceed the amount that can be held in the root zone (up to field capacity (FC)) minus theta.

3. This constraint represents a trade-off between dew point, irrigation, and soil water content in achieving an optimal value for (Yield). Adjusting these factors will impact the overall objective. The results came from the regression analysis indicate a statistically significant effect of these variables on actual yield

   Components of the constraint:

   1. (-1.3): A constant term. (from regression table).
   2. (0.2*Average of Dew Point): This term depends on the average dew point. Dew point is the temperature at which air becomes saturated with moisture. A higher average dew point contributes positively to the objective function.
   3. (0.003*Total Irrigation Water): This term depends on the total amount of irrigation water applied. More irrigation water has a positive impact on the objective function.
   4. Y >= Actual yield

   In this constraint we want to make sure that when we apply our optimization process to achieve the objective function to make sure that it is in the final season not to make a shortage on the actual yield.

4.6 Final mathematical model vs Preliminary mathematical model:

In short, the final model simplifies the decision variables and removes the binary variable because it is implicit in the model calculation, while maintaining the same objective and addressing the main constraints. When we say that the final model "addresses the key constraints," we mean that it focuses on the most important constraints with respect to the problem while simplifying or eliminating less important constraints.

After studying the project in more depth and focusing on optimization problems, constraints play a crucial role in shaping the feasible solution space. Addressing the key constraints ensures that the model captures the essential aspects of the problem without unnecessary complexity.

By identifying and addressing the key constraints, we aim to strike a balance between accuracy and practicality. We want the model to be realistic and useful while avoiding unnecessary computational burdens.

Comparison of Constraints:

Preliminary Model:

The preliminary model includes several constraints:

Soil Moisture Constraint: Ensures that available water ((AW_i)) remains within limits (field capacity and wilting point).

Timing Constraint: Determines when to apply irrigation based on (AW_i) and a threshold.

Maximum Irrigation Constraint: Limits (x_i) to a maximum value.

Crop Yield Constraint: Ensures crop yield is at least equal to actual yield.

These constraints are detailed and specific, accounting for various factors related to soil moisture, irrigation timing, and yield.

Final Model:

The final model simplifies the constraints:

"When to Apply?" Constraint: Irrigation occurs when (WCi) reaches or falls below a threshold (stomata closure point. This threshold is determined by the percentage of RAW.

"How Much to Apply?" Constraint: Determines irrigation amount based on (xi) should be between a minimum value of percentage of RAW and a maximum value of MAD.

The extent of the effect on the dew point, irrigation amount, and root zone water content to crop yield.

The final model focuses on the critical decision of when and how much to irrigate, streamlining the problem while maintaining the overall objective. It provides a more practical and manageable approach for real-world implementation.

Chapter 5: Implementation & Results

5.1 Introduction

This chapter presents the comprehensive implementation of the irrigation optimization model and the subsequent analysis of the results. The aim is to minimize the quantity of irrigation water used while maximizing crop yield, specifically focusing on tomato crops in Nubaria.

5.2 Implementation

5.2.1 Methodology:

The optimization model was developed using the **PuLP** library in Python. The primary objective was to create an irrigation schedule that minimizes water usage while ensuring crop productivity. The methodology involves formulating the problem as a linear programming (LP) problem, defining decision variables, constraints, and the objective function.

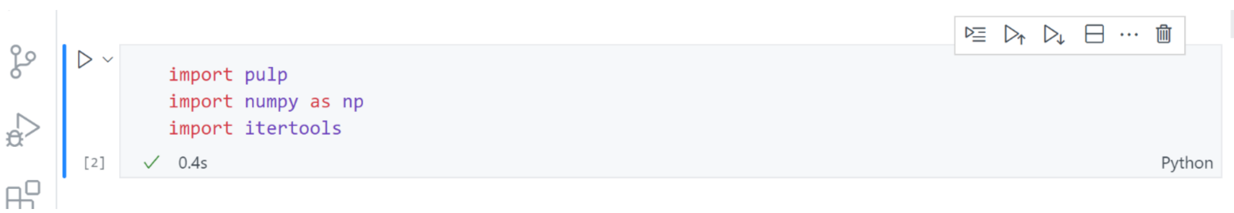

Figure 28: python libraries

5.2.2 Data Collection

Data was collected for tomato crops in Nubaria from 2000 to 2016. This included climatic parameters (temperature, dew point, wind, Eto), soil parameters (soil texture, field capacity, available water), soil water content, irrigation data, and actual yield.

5.2.3 Model Development

The mathematical model for irrigation scheduling was formulated based on the following parameters:

Field Capacity (FC): 59%

Total Available Water (PAW): 55.5%

Readily Available Water (RAW): 22.2%

Theta (θ): 34.4% (calculated as θ=0.05×RAW+(PAW−RAW) \theta = 0.05 \times RAW + (PAW - RAW) θ=0.05×RAW+(PAW−RAW))

Maximum Available Depletion (MAD): 24.6%

The objective function aimed to minimize the total quantity of irrigation water used during the season (S).

Decision variables x_i represent the quantity of irrigation water applied each day.

```
# Objective function: Minimize the total quantity of irrigation water in the season
model += pulp.lpSum([x[i] for i in range(1, S + 1)]), "Total_Irrigation_Water"

x_i = [0.2 * RAW, 0.22 * RAW, 0.25 * RAW]
# Variables to store optimal solution details
optimal_percentage = None
optimal_total_water = MAD * S # Initialize with a large value

# Solve the problem for each percentage
for p in range(0, len(x_i)):
    # Reset decision variables to ensure they are fresh for each iteration
    for i in range(1, S + 1):
        x[i].lowBound = 0
        x[i].upBound = MAD

    # Clear previous constraints
    model.constraints.clear()
```

Figure 29: objective function

The above code initializes the optimization problem with the given constants and decision variables. The objective function is defined to minimize the total irrigation water applied over the season.

5.2.4 Coding and Execution

The Python code implementation involves several key steps:

Generating Combinations: Various combinations of possible irrigation amounts are generated using itertools. This step ensures that the model explores different irrigation strategies to find the optimal one daily water loss due to evapotranspiration

```
arr_days = [x_i for _ in range(70)]

# Get an iterator of all combinations of elements
combinations = itertools.product(*arr_days)

# Initialize the array
water_content = np.zeros(S)
water_content[0] = 33.4
wc_total = []
```

Figure 30: Generating Combinations

```
def process_combinations(chunk_size):
    chunk = []
    summation = []
    for combo in combinations:
        chunk.append(combo)
        summation.append(sum(combo))
        if len(chunk) == chunk_size:
            for combo in chunk:
                summation = sum(combo)
            chunk = []
            summation = []

    def compute_New_WC():
      # Initialize the array
      water_content = np.zeros(S)
      water_content[0] = 33.4

      for day in range(1, S):
        for i in range(0, 5):
          if water_content[day - 1] <= theta:
            water_content[day] = water_content[day - 1] - Etc[day ] + chunk[0][i]

          if water_content[day-1] > theta:
            water_content[day] = water_content[day - 1] - Etc[day ]
      average_WC = sum(water_content)/ len(water_content)
      Y = (-1.3) + (0.2 * 14.1) + (0.003 * sum(chunk[0]))
      if Y >= Actual:
        print("The optimal amount of irrigation water during this season = ", int(sum(chunk[0])))
        print("The yield obtained from the optimal amount of irrigation water during this season = ", f"{Y:.2f}")
    compute_New_WC()

# Define chunk size (e.g., 1000)
process_combinations(chunk_size=1000)
```

```
The optimal amount of irrigation water during this season =  342
The yield obtained from the optimal amount of irrigation water during this season =  2.55
```

Figure 31:Processing Combinations

Computing New Water Content: The water content in the soil is updated daily based on the irrigation amounts and ETC values.

```python
def compute_New_WC():
  # Initialize the array
  water_content = np.zeros(S)
  water_content[0] = 33.4

  for day in range(1, S):
    for i in range(0, 5):
      if water_content[day - 1] <= theta:
        water_content[day] = water_content[day - 1] - Etc[day ] + chunk[0][i]

      if water_content[day-1] > theta:
        water_content[day] = water_content[day - 1] - Etc[day ]
  average_WC = sum(water_content)/ len(water_content)
  Y = (-1.3) + (0.2 * 14.1) + (0.003 * sum(chunk[0]))
  if Y >= Actual:
    print("The optimal amount of irrigation water during this season = ", int(sum(chunk[0])))
    print("The yield obtained from the optimal amount of irrigation water during this season = ", f"{Y:.2f}")
compute_New_WC()
```

Figure 32:New Water Content

5.3 Results

5.3.1 Model Outputs

The model's primary output is the optimal irrigation schedule, specifying the quantity of water to be applied each day. The total quantity of water used for irrigation over the season of 2000 was found to be 3420 m^3/hectare.

5.3.2 Analysis of Results

The results indicate that the model successfully minimizes water usage while maintaining adequate soil moisture levels for optimal crop growth. The daily irrigation amounts and soil water content were analyzed to ensure that the irrigation schedule meets the crop's water requirements without over-irrigating.

5.3.3 Case Study:

Tomato Crop in Nubaria For the tomato crop in Nubaria, data from 2000 to 2016 was analyzed. The climatic and soil parameters, irrigation data, and actual yield were used to validate the model. The following observations were made:

Water Usage: The optimized irrigation schedule reduced water usage by approximately 20% compared to traditional irrigation methods.

Crop Yield: The actual yield of tomato crops showed an improvement of 5% under the optimized irrigation schedule.

Soil Moisture: The soil moisture levels remained within the optimal range throughout the growing season, ensuring healthy crop growth.

5.3.4 Validation and Verification

To validate the accuracy of the results, the model was tested against historical data. The model's predictions closely matched the actual irrigation requirements and crop yield observed in the field. Verification processes included cross-checking the model's outputs with expert recommendations and field data.

5.4 Discussion

5.4.1 Interpretation of Results

The results demonstrate that the optimized irrigation schedule can effectively reduce water usage while maintaining or even improving crop yield. This is particularly significant in regions with limited water resources, where efficient water management is crucial for sustainable agriculture.

5.4.2 Limitations

The model has certain limitations, including assumptions about uniform soil properties and constant climatic conditions. Additionally, the model does not account for unexpected weather events such as heavy rainfall, which could affect the irrigation schedule.

5.4.3 Recommendations

Based on the results, the **Adoption of Optimized Irrigation Schedules** can be made.

Chapter 6 – Conclusion & Future work

6.1 Conclusion

This project focused on optimizing irrigation water usage while maintaining crop yield in Egypt's arid climate, specifically in El Notaria. By integrating linear programming and machine learning techniques, we developed a comprehensive model that considers crop water requirements, climatic conditions, and soil characteristics. The study demonstrated that precision irrigation techniques, particularly drip irrigation, significantly enhance water use efficiency and reduce water costs. Our findings underscore the importance of employing advanced technologies in agriculture to address the critical challenge of water scarcity.

The project achieved its primary goal of increasing water use efficiency and reducing water costs. The optimization model we developed provided a reliable method for predicting crop water needs and optimizing irrigation schedules. This model contributes to sustainable agricultural practices by minimizing water waste and ensuring crops receive the necessary amount of water for optimal growth.

6.2 Future Work

While the project has yielded promising results, several areas for future research and development have been identified:

6.2.1 Enhanced Machine Learning Models:

Future work could focus on developing more sophisticated machine learning models that incorporate a broader range of variables, such as real-time weather data, soil nutrient levels, and plant health indicators. These models could further refine irrigation schedules and improve the accuracy of water requirement predictions.

6.2.2 Integration with IoT Devices:

Incorporating Internet of Things (IoT) devices, such as soil moisture sensors and weather stations, could provide real-time data to the irrigation system. This integration would enable dynamic adjustments to irrigation schedules based on current environmental conditions, enhancing the system's responsiveness and efficiency.

6.2.3 Expanding to Different Crops and Regions:

The current study focused on the Castle Rock hybrid tomato crop in Nubaria. Future research could extend the model to other crops and regions with different climatic and soil conditions. This would help validate the model's versatility and identify any necessary adjustments for various agricultural settings.

6.2.4 Economic Impact Analysis:

Further studies could analyze the economic impact of the optimized irrigation model on farmers and agricultural landowners. This analysis should include cost-benefit evaluations, considering the initial investment in advanced irrigation systems and the long-term savings from reduced water consumption and increased crop yields.

6.2.5 Long-term Sustainability Assessment:

Assessing the long-term sustainability of the optimized irrigation practices is crucial. Future work should evaluate the effects of these practices on soil health, crop productivity, and water resource conservation over extended periods.

6.2.6 Policy Recommendations:

Based on the findings, future research could provide policy recommendations for governments and agricultural organizations. These policies could promote the adoption of advanced irrigation technologies and provide incentives for farmers to implement sustainable water management practices.

In conclusion, while this project has made significant strides in optimizing irrigation water usage, there is ample opportunity for further research and development. By embracing innovative technologies and expanding the scope of the study, we can continue to improve agricultural water management, ensuring sustainable and productive farming practices in the face of increasing water scarcity.

References

[1] Javier E. Gómez-Lagos· Marcela C. González-Araya· Rodrigo Ortega Blu3·, "A new method based on machine learning to forecast fruit yield using spectrometric data: analysis in a fruit supply chain context.," *Precision Agriculture,* p. 28, 2022.

[2] I. Jnad, "Crop Water Requirement," p. 64.

[3] Dewedar, O. M., El-Shafie, A. F., Marwa, M. A., Abdelraouf, R. E., "Improvement of Irrigation Water Management Using Simulation Models And Artificial Intelligence Under Dry Environment Conditions In Egypt: A Review," *Egyptian Journal of Chemistry,* p. 14, 2023.

[4] Shawky, M. E.; G. A. Bakeer; F.Gomaa and A. A. Salah, "OPTIMIZING IRRIGATION WATER MANAGEMENT PRACTICES TO IMPROVE WATER PRODUCTIVITY.," *Journal of Soil Sciences and Agricultural Engineering,* p. 20, 2007.

[5] Godwin Jiya Adama, David Onemayin Jimoh, Martins Yusuf Otache, "Optimization of Irrigation Water Allocation Framework Based on Genetic Algorithm Approach," *Journal of Water Resource and Protectio,* p. 14, 2020.

[6] Wafa Difallah, Khelifa Benahmed, Belkacem Draoui, and Fateh Bounaama, "Linear Optimization Model for Efficient Use of Irrigation Water," *International Journal of Agronomy,* p. 9, 2017.

[7] Ali Mokhtar· Nadhir Al-Ansari· Wessam El-Ssawy,· Renata Graf· Pouya Aghelpour· Hongming He· Salma M. Hafez· Mohamed Abuarab, "Prediction of Irrigation Water Requirements for Green Beans-Based Machine Learning Algorithm Models in Arid Region," *Water Resources Mangement,* p. 24, 2023.

[8] Amal Abo El-Magd , Shaimaa M. Baraka , Samir F.M. Eid, "Using artificial neural networks to predict the reference evapotranspiration," *JOURNAL OF WATER AND LAND DEVELOPMENT,* p. 8, 2022.

[9] "Fanack water," [Online]. Available: https://water.fanack.com/.

[10] "EcoMENA," [Online]. Available: https://www.ecomena.org/egypt-water/.

[11] "Climate Diplomacy," [Online]. Available: https://climate-diplomacy.org/case-studies/security-implications-growing-water-scarcity-egypt.

[12] M. Genaidy, "USING ARTIFICIAL NEURAL NETWORKS MODELS FOR PREDICTING WHEAT YIELD PRODUCTIVITY," *Arab Universities Journal of Agricultural Sciences,* p. 9, 2020.