Integrated Prediction and Optimization of Sugarcane Yield: Analyzing Climate Impact, Agricultural Practices, and Optimal Harvest Timing in Bangladesh

by

Md Ashiful Haque
24141208

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC UNIVERSITY
August 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

MD ASHIFUL HAQUE
24141208

Approval

The internship report titled "Integrated Prediction and Optimization of Sugarcane Yield: Analyzing Climate Impact, Agricultural Practices, and Optimal Harvest Timing in Bangladesh" submitted by

1. Md Ashiful Haque (24141208)

Of Fall 2024, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on November 30, 2024.

Examining Committee:

Academic Supervisor:
(Member)

Arif Shakil
Senior Lecturer
Department of Computer Science and Engineering
School of Data and Sciences
BRAC University

External Supervisor:

29.09.2024
Engr. Md. Ataur Rahman Khan
Director (Production & Engineering), Grade-2
BSFIC, Chinishilpa Bhaban (4th Floor)
3, Dilkusha C/A, Dhaka.

Engr. Md. Ataur Rahman Khan
Director (Production & Engineering) Grade 2
Production and Engineering
Bangladesh Sugar and Food Industries Corporation, Head Office, Dhaka.

Program Coordinator:
(Member)

Dr Golam Rabiul Alam
Professor
Department of Computer Science and Engineering
School of Data and Sciences
BRAC University

Head of Department:
(Chair)

Sadia Hamid Kazi
Associate Professor; Chairperson
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Ethics Statement

This research was conducted with a commitment to the highest ethical standards, ensuring that all aspects of the study adhered to principles of integrity, transparency, and respect for all stakeholders involved. The following ethical considerations were paramount throughout the research process:

1. **Respect for Data Integrity and Confidentiality**
   This study utilized publicly available datasets, including historical sugarcane production data and climate records. All data were carefully managed to maintain their integrity, with rigorous efforts to ensure accuracy and avoid any form of misrepresentation. No personal or confidential information was involved in the research, and all data were treated with the utmost respect for the sources from which they were obtained.

2. **Commitment to Research Integrity**
   Every stage of this research, from data collection and analysis to the reporting of findings, was conducted with complete transparency. Methodologies, models, and analytical tools were thoroughly documented and are fully replicable, ensuring that the research can be verified and validated by peers. The findings were presented truthfully and objectively, without any alteration or selective reporting of results.

3. **Responsibility to the Agricultural Community**
   The primary objective of this research was to contribute positively to the agricultural sector, particularly to the sugarcane farmers in Bangladesh. The study was designed to offer practical insights and recommendations that could enhance agricultural productivity and sustainability. The research was conducted with the understanding that the outcomes could have significant implications for the livelihoods of many individuals, and this responsibility was taken seriously throughout the project.

4. **Environmental and Social Considerations**
   The research was conducted with a strong awareness of the environmental and social implications of agricultural practices. Recommendations were made with the goal of promoting sustainable agriculture, minimizing environmental impact, and supporting long-term ecological balance. The study also considered the broader social impact of agricultural policies and practices, advocating for approaches that are equitable and beneficial to the wider community.

5. **Adherence to Ethical Research Guidelines**
   This research adhered strictly to all relevant ethical guidelines and standards as set forth by academic and professional bodies. The study was conducted free from conflicts of interest, ensuring that the research process was unbiased and objective. Any potential ethical dilemmas were carefully considered, and appropriate steps were taken to address them.

6. **Acknowledgment and Intellectual Honesty**
   All contributions to this research, whether in the form of data, methodologies, or intellectual input, have been appropriately acknowledged. The research was conducted with a deep respect for intellectual property rights, and every

effort was made to ensure that all sources were credited accurately and fairly. Collaboration was carried out with a spirit of intellectual honesty and mutual respect.

This ethics statement embodies the commitment of the researcher to uphold the highest standards of ethical conduct in all aspects of this study, with a focus on producing research that is not only scientifically rigorous but also socially and environmentally responsible.

Abstract

This study presents a comprehensive approach for predicting and optimizing sugarcane yield in Bangladesh by incorporating climate dynamics, agricultural practices and optimal harvest timing. Highly accurate yield predictions are made that identify the optimal harvest periods using sophisticated machine learning models trained on detailed historical data encompassing the years 1971 to 2024. The results demonstrate that temperature and precipitation can strongly influence crop productivity and provide useful solutions to manage the impact of climate variability. This research serves a critical need for farmers and policymakers, offering practical methods to maintain or increase crop yields as well as agricultural resilience to extreme climate change.

Keywords: Sugarcane Yield Prediction; Climate Impact Analysis; Agricultural Optimization; Harvest Timing; Bangladesh Agriculture; Machine Learning in Agriculture; Predictive Analytics; Crop Management; Climate Adaptation; Yield Optimization;

Dedication

I want to dedicate this report to my parents who have been the backbone of my academic and personal life – providing for me the most important foundation for living that they can. I believe your faith in my ability to get this done was the biggest strength I had.

To my mentors and teachers, whose guidance and knowledge have shaped my journey in the field of engineering. Your wisdom and patience have inspired me to strive for excellence.

Finally, to all the farmers of Bangladesh, whose hard work and resilience continue to nourish our nation. This report is a small step towards making their efforts more fruitful and sustainable.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.

Secondly, to my advisor Mr. Arif Shakil sir for his kind support and advice in our work. He helped us whenever we needed help.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AI Artificial Intelligence

ANN Artificial Neural Network

$BSFIC$ Bangladesh Sugar and Food Industries Corporation

CDD Continuous Dry Days

DSS Decision Support System

$DSSAT$ Decision Support System for Agrotechnology Transfer

$ECMWF$ European Centre for Medium-Range Weather Forecasts

$ERA5$ ECMWF Reanalysis 5

GCM General Circulation Model

GIS Geographic Information System

GPS Global Positioning System

$HD30$ Heatwave Indicator for temperatures exceeding 30°C

$HD35$ Heatwave Indicator for temperatures exceeding 35°C

IoT Internet of Things

IPM Integrated Pest Management

$LSTM$ Long Short-Term Memory

MAE Mean Absolute Error

MT Metric Ton

$R2$ Coefficient of Determination

$RMSE$ Root Mean Squared Error

RNN Recurrent Neural Network

$RX1day$ Maximum precipitation in a single day

$RX5day$ Maximum precipitation in five consecutive days

SVR Support Vector Regression

UAV Unmanned Aerial Vehicle

$XGBoost$ eXtreme Gradient Boosting

Chapter 1

Introduction

1.1 Background of the Study

Sugarcane has long been a cornerstone of Bangladesh's agricultural sector, contributing significantly to the country's economy [17], [26]. It is not only a source of sugar, a staple in Bangladeshi diets, but also an essential ingredient in the production of molasses and ethanol, which have both domestic and industrial applications [22], [23]. As such, the sugarcane industry directly impacts the livelihoods of millions of farmers and workers across the nation.

However, the production of sugarcane is inherently complex, influenced by a multitude of factors, including soil quality, water availability, agricultural practices, and, critically, climatic conditions [5], [12]. In recent years, the impact of climate change has introduced new challenges to sugarcane cultivation [24], [29]. Yield fluctuations have been induced by temperature, precipitation patterns, and changes in frequency of extreme weather events, thereby threatening the stability of sugarcane production [21], [32].

It's never been a more urgent need to adapt to these changing conditions. It is said that new age of rapid environmental change may not be compatible with traditional farming methods relying on historical knowledge and experience. To ensure the future of sugarcane farming in Bangladesh, it must adopt data driven approaches which shall base its predictions on data accurately and optimize farming practices in response to a changing climatic condition.

1.2 Problem Statement

This research has a major concern, given that the consecutively unpredictable nature of sugarcane yields due to climate variability [29], goes to unprecedented levels. Current farming practices of the same are traditional based which means that they cannot provide stable high yields as well as have limited possibilities of
provide stable high yields ad [24]. For example, future agronomic practices, like determining when to harvest for given changes in the climatic, are more data driven and require need for more optimzed time harvesting due to climatic changes. This study will help to develop predictive models based on climate data and agricultural activities, which could be used by farmers as well as policymakers in planning

sugarcane production in Bangladesh [11], [25].

1.3 Objectives of the Internship

The main goal of this traineeship is the implementation of a solution to incorporate predictive analysis in sugarcane production, providing stakeholders with support when deciding about what can be done next by using reliable and helpful information. The project involves creating predictive models to calculate the yield of sugarcane by examining historical data on production and meteorological variables such as temperature, precipitation, and relative humidity.

The project further seeks to understand the magnitude of variation in yield performance linked with specific agricultural practices. Through characterizing the ideal practices for different climatic conditions to improve productivity, it will offer farmers advice on what gives maximum output. The main priority of the project is determining when to harvest so that not only more sugarcane can be obtained, but also their quality. Key to achieving the highest yield and sugar content will be in identifying the exact right harvest window.

1.4 Scope of the Project

This project is intending to learn many areas of studies and hence the scope for this goal is large. This response is difficult to answer, because it's a complicated process that requires the compilation of large datasets over an extended period (sugarcane production records 1971-2024 and climate data from 1950-2022). Develop predictive models to predict yields under different climate scenarios using these datasets.

Second, the project will analyze various agricultural practices in terms of how they affect sugarcane yield. This involves assessing the impact of irrigation methods, fertilization schedules, and pest management practices. The study will show how those practices correlate with yield outcomes to better understand what farmers can do, now and in the future, to boost agricultural efficiency.

The project will then develop strategies to optimize harvest timing. The study will, through analysis of historical data, show in which periods the yields are maximized considering cane maturity, sugar content, and weather conditions at that time. Our overall objective is to devise an economical approach that integrates agro-practices, harvesting dates, and climate predictions together for sustainable sugarcane farming in Bangladesh.

1.5 Significance of the Study

The importance of this study in the future of sugarcane production for Bangladesh is very high. Million-dollar revolutions will be possible in the agronomic practices of India through this research, done with data-based methods. A result of this project is the predictive models, which will help farmers predict weather and adjust their practices accordingly.

In addition, findings from this study will help in understanding agricultural science more broadly as a framework for other crops and regions that have climate stresses. With climate change already affecting food production globally, studies like this are instrumental in crafting adaptive measures that will secure both food and economic futures.

1.6 Structure of the Report

This report is designed to give an overview of research done throughout the course of this internship. In this context, Chapter 2 starts by presenting a literature review of studies investigating sugarcane yield and climate impact mechanisms, complemented by field management practices on the Zona da Mata. Chapter 3 presents the methodology for this study, covering data collection, pre-processing, and model building. Chapter 4 presents an analysis of the results, detailing key findings and their implications. Chapter 5 discusses the wider implications of these findings and contrasts them with other research. Chapter 6 is the conclusion, giving recommendations on research and agricultural practices.

Chapter 2

Literature Review

2.1 Overview of Production and Plantation (1971-2023)

Based on the previous list, this is an overall idea about the sugarcane production and plantation statistics for Bangladesh between 1971 and the future. Key data are presented in Table 2.1, covering the trends over time for sugarcane production, combined with some of the factors driving change; this is impacted by agricultural practices (fertilizer use and vast shifts to higher-yielding seeds), as well as changes in industrial processing systems.

Crushing Season	Total Cane Plantation (Acre)	Average Yield per Acre (MT)	Cane Crushed (MT)	Sugar Production (MT)	Recovery Rate
1971 - 1972	136175	8.37	409160.00	24200.00	5.92
1972 - 1973	117775	9.37	274310.00	19604.00	7.14
1973 - 1974	140470	14.98	1187202.00	89808.00	7.56
1974 - 1975	175366	12.58	1422181.00	100040.00	7.02
1975 - 1976	124889	13.36	1100946.00	88177.00	8.10
1976 - 1977	162752	14.68	1706370.00	140925.00	8.26
1977 - 1978	237270	13.79	2309652.00	178072.00	7.72
1978 - 1979	187734	13.71	1715505.00	132812.00	7.74
1979 - 1980	155587	14.16	1272089.00	94714.00	7.46
1980 - 1981	191182	14.82	1798165.00	134021.00	7.45
1981 - 1982	181876	13.35	1478396.00	110796.00	7.49
1982 - 1983	194605	13.26	1606269.00	120462.00	7.50
1983 - 1984	195623	12.31	1399875.00	107301.00	7.66
1984 - 1985	170721	14.59	1410784.00	110480.00	7.83
1985 - 1986	158285	12.57	1172683.00	91313.00	7.79
1986 - 1987	134356	11.64	1048357.00	77728.00	7.41
1987 - 1988	127978	11.47	1017595.00	76329.00	7.50
1988 - 1989	131267	11.68	1039436.00	77386.00	7.45
1989 - 1990	125301	12.53	1081530.00	82135.00	7.60
1990 - 1991	135279	14.64	1239272.00	93030.00	7.51
1991 - 1992	145215	14.15	1273571.00	97796.00	7.68
1992 - 1993	151947	14.66	1322834.00	100029.00	7.56
1993 - 1994	151072	14.16	1208381.00	88241.00	7.30
1994 - 1995	151168	15.72	1385396.00	100999.00	7.29
1995 - 1996	140725	15.62	1186061.00	86177.00	7.27
1996 - 1997	147623	15.40	1286898.00	93741.00	7.28
1997 - 1998	134429	13.80	1165625.00	81411.00	6.98
1998 - 1999	130826	14.20	1141853.00	82225.00	7.20
1999 - 2000	139519	15.45	1286764.00	90519.00	7.04
2000 - 2001	152188	14.73	1308147.00	93573.00	7.15
2001 - 2002	140063	14.91	1187865.00	85772.00	7.22
2002 - 2003	124488	13.78	991729.00	69513.00	7.01
2003 - 2004	120940	13.24	941736.00	65636.00	6.97
2004 - 2005	119432	13.47	931573.00	64329.00	6.91
2005 - 2006	123414	13.18	976489.00	67699.00	6.93
2006 - 2007	134400	12.31	911798.00	63484.00	6.96
2007 - 2008	116262	13.77	1041710.00	75872.00	7.29
2008 - 2009	133036	15.47	1269704.00	93140.00	7.33
2009 - 2010	130620	17.18	1087371.00	86673.70	7.17
2010 - 2011	161429	18.83	1581857.00	100962.80	6.38
2011 - 2012	157759	18.16	1047501.00	69346.80	6.61
2012 - 2013	159673	19.06	1562351.00	107123.00	6.85
2013 - 2014	174006	18.97	1818837.81	128268.20	7.06
2014 - 2015	154953	17.32	1214752.00	77450.05	6.37
2015 - 2016	129473	16.51	963835.00	58219.30	6.04
2016 - 2017	117854	17.24	991454.74	59984.55	6.05
2017 - 2018	110251	18.45	1188573.04	68562.50	5.77
2018 - 2019	119108	18.18	1182171.24	68952.10	5.83
2019 - 2020	116865	18.40	1401980.38	82140.45	5.86
2020 - 2021	127283	15.47	875776.24	48133.95	5.49
2021 - 2022	49909	16.05	442674.62	24509.75	5.53
2022 - 2023	49825	16.17	384350.55	21313.25	5.54

Table 2.1: Sugarcane Production Data (1971-2023)

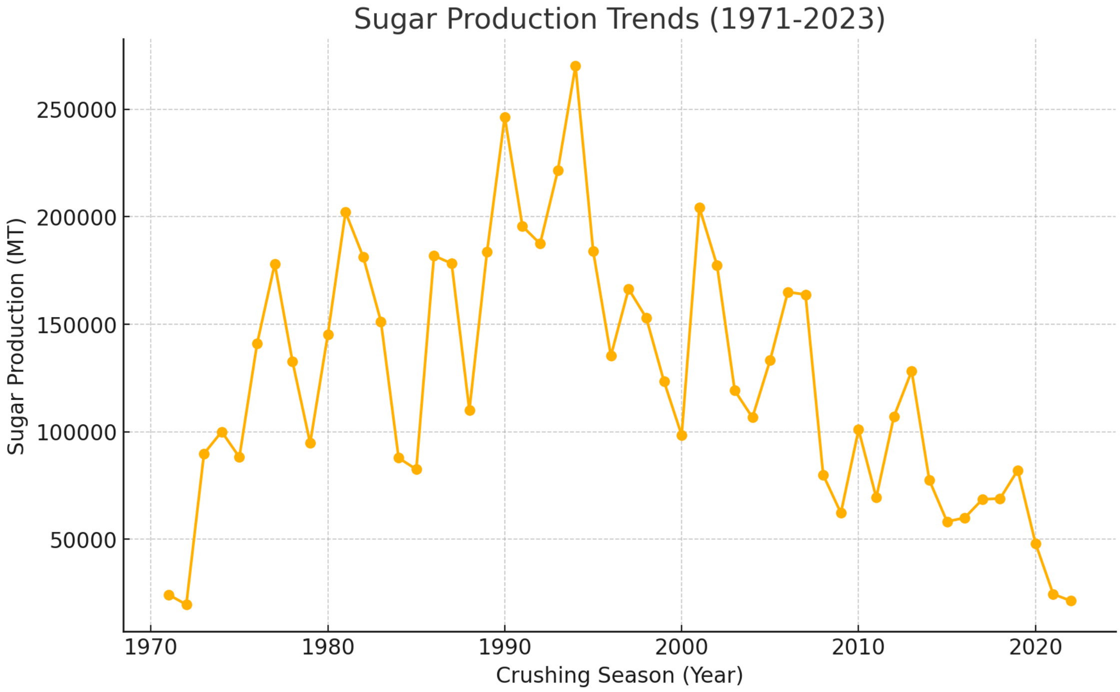

Figure 2.1: Sugarcane Production Trend Over the Years

Sugar production in Bangladesh from 1971 to 2023 has seen various fluctuations. The sugar mills in East Pakistan were nationalized to form the Bangladesh Sugar & Food Industries Corporation, and some successful public-private partnerships helped increase sugarcane yields. A few development efforts even date back to when Bengal was part of British India. While there has been a slight recovery in recent years, production levels are still lower than in historic peak periods. These trends probably reflect a combination of agricultural productivity, factory efficiency, and external factors affecting the industry over time.

2.2 Sugarcane Yield and Climate Impact

2.2.1 Overview of Sugarcane Cultivation

Sugarcane, the source of 70% of world sugar production (FAO2019), is an important commodity crop worldwide, mainly grown in tropical and subtropical regions [6]. Sugarcane farming is a major agricultural venture in Bangladesh, where sugar and molasses are produced from sugarcane. Sugarcane usually grows well in the warm and humid climate of the country, but its growth and yield are profoundly dependent on weather changes.

2.2.2 Impact of Temperature on Sugarcane Yield

Temperature is a major climatic factor governing the physiological processes of sugarcane due to its direct or indirect effects on photosynthesis, respiration, and other metabolic processes in the plant (Robertson et al., 1994). The most favorable temperature for sugar accumulation in sugarcane is between 20°C and 30°C, but sucrose content decreases when temperatures exceed this range during the ripening phase.

This worsens under excessive heat as the transpiration rate (loss of water by the plant) increases, accelerating the effects of water scarcity [30]. Literature supports the increasing challenges faced by sugarcane farming regions due to rising global temperatures caused by climate change, necessitating the development of heat-tolerant varieties and evolving adaptation techniques for agriculture [29].

2.2.3 Influence of Precipitation and Water Availability

The second important factor that restricts sugarcane yield is water availability [5]. During the growing season, sugarcane needs plenty of water, and estimates for rainfall requirements vary from 1500 to 2500 mm per year, depending on soil type and cultivation practices. Lack of rainfall or drought conditions could severely affect crops, where both plant growth and yield will be hampered due to poor development. Excessive rainfall, especially during ripening, can decrease sucrose content and ultimately reduce sugar production. Studies emphasize the necessity of efficient irrigation practices and water management strategies to adapt to changing rainfall patterns [14].

2.2.4 Effects of Humidity and Other Climatic Factors

The level of humidity affects the prevalence of certain diseases, as well as how much water stays in the plant. Favorable conditions for the growth of pests and diseases are created with warm temperatures together with humid weather, further reducing yields [29]. This serves as evidence for the essential role of integrated pest management (IPM) strategies to address these challenges. However, sugarcane growth is also affected by other climatic factors such as wind speed and solar radiation [7], with windstorms causing direct physical injury to the crop, thereby reducing yield.

2.3 Agricultural Practices and Yield Optimization

2.3.1 Role of Soil Management and Fertilization

The most ideal soil health practices for sugarcane cultivation rely on good root zone establishment by developing improved roots, increasing water-holding capacity, and reducing weed growth. Literature suggests that the application of organic and inorganic fertilizers maintains soil fertility [31]. Sugarcane requires primary nutrients, including nitrogen, phosphorus, and potassium, for better yield because balanced fertilizer application is very important. Studies further emphasize the advantages of soil conservation practices, such as crop rotation and green manuring, which benefit water infiltration and prevent erosion [19].

2.3.2 Impact of Irrigation Techniques

Proper irrigation becomes fundamental in areas of scarce or erratic rainfall. Among the range of irrigation methods for sugarcane farming, literature highlights drip, sprinkler systems, and surface irrigation as having significant potential for implementation [14]. Drip irrigation is notably appreciated for its capacity to bring water close to the plant root zone, greatly reducing water loss and enhancing nutrient uptake. Additionally, treated wastewater can be used for irrigation, which holds promise as a means of conserving existing freshwater sources needed for continuous sugarcane cultivation in more arid regions [19].

2.3.3 Pest and Disease Management

Sugarcane borers, rusts, smuts, and ratoon stunting disease are major threats to the sugarcane industry. Several control methods have been reported in the literature, including chemical pesticides, biological controls, and resistant crop varieties. The method in question is known as Integrated Pest Management (also IPM) which is a holistic practice that considers cultural, biological and chemical methods only as a last resort, a strategy that is economically sound and is not overly impacting the environment. Additionally, the use of a cultivar that is resistant to disease and the doing of routine field checks to monitor for disease should be used to decrease the incidence of crop loss [29], [31].

2.3.4 Technological Innovations in Agriculture

As a new area for improving sugarcane yield with precision agriculture, an evolution has arisen. Currently, crops are increasingly monitored using Remote sensing, Geographic Information Systems (GIS) and Unmanned Aerial Vehicles (UAVs) to estimate soil moisture levels, estimate crop health and to forecast yields [20]. According to literature, using these technologies can enhance sugarcane farming decision-making with real time data for more targeted interventions [25]. Decision Support Systems (DSS) and the crop modeling software can also be used by farmers to plan their operation more efficiently knowing the weather forecast, soil conditions and other relevant factors [10].

2.4 Effects of Harvest Timing on Yield.

2.4.1 Optimal Harvesting Periods

The timing at which to harvest has a direct impact on sugarcane yield and sugar content [30]. According to the literature, the sucrose content in the cane is highest when harvested and typically after a long period of mild, sunny weather [1]. Harvesting can be delayed which can reduce sucrose concentration, or pre harvesting may lower yields or produce poor sugar quality. Studies suggest using maturity indices like the Brix value (a measure of total soluble solids or sucrose concentration) to determine the optimal harvest time [29].

2.4.2 Climate Influence on Harvest Timing

Climatic conditions during the harvest period significantly impact sugarcane production outcomes. For instance, harvesting during or after heavy rainfall can complicate cane transport and increase the risk of mud contamination, which reduces sugar extraction efficiency [29]. Conversely, harvesting during dry, cool conditions is preferable as it minimizes losses and preserves sugar content. The literature emphasizes the importance of adaptive harvest planning that considers both long-term climate trends and short-term weather forecasts [7].

2.4.3 Mechanization and Harvesting Efficiency

Mechanized harvesting, through the introduction of cane choppers and whole stalk combine machines, has revolutionized sugarcane farming in many parts of the world, reducing labor costs and increasing efficiency [31]. However, the literature also highlights challenges associated with mechanization, such as soil compaction, which can negatively affect future crops. These issues can be mitigated by adopting sustainable mechanization practices, such as using lighter machinery and controlled traffic farming. Additionally, the timing and method of cane cutting can influence the regrowth of ratoon crops, which are crucial for maintaining long-term yield stability [29].

2.5 Integration of Climate Data in Agricultural Planning

2.5.1 Use of Climate Models for Yield Prediction

In agriculture, the use of climate models has been considered a tool to predict crop yields under varying climatic scenarios [10]. The literature highlights the use of General Circulation Models (GCMs) and downscaled climate projections to evaluate the effects of future climate change on sugarcane production. Integrating these models with crop simulation tools provides more accurate yield forecasts, helping farmers and policymakers make informed, data-driven decisions [7].

2.5.2 Challenges in Climate Data Utilization

Despite the promise of using climate data for agricultural planning, the literature reveals several challenges. These include model accuracy and resolution, uncertainty in long-term projections, and the difficulty of converting model outputs into actionable information for farmers [18]. To address these challenges, studies advocate for the development of user-friendly decision support systems that can bridge the gap between complex climate data and practical farming decisions [10].

2.5.3 Future Directions in Climate-Adaptive Agriculture

n the near future, sugarcane farming in Bangladesh will increasingly rely on climate-adaptive strategies. This includes breeding climate-resilient crop varieties, adopting water-saving irrigation techniques, and utilizing precision agriculture tools [29]. Furthermore, integrating real-time climate monitoring and predictive analytics into farming practices will be essential in enhancing the sugarcane sector's resilience to climate variability and change [25].

2.6 Limitations and Conclusion

A major limitation of this analysis is the absence of real climate data for Bangladesh. Simulated scenarios and theoretical insights were used to hypothesize the effect of climate on sugarcane yields. Future studies should incorporate real-time weather data to enhance the accuracy of analysis and recommendations.

The study emphasizes that maximizing sugarcane yield requires optimizing both agricultural practices and harvest timing, as these variables are interdependent. While the current study uses hypothetical scenarios, future research could improve predictions by incorporating empirical climate and weather data, ultimately contributing to a more resilient sugarcane industry.

Chapter 3

Methodology

3.1 Predictive Modeling and Optimization of Sugarcane Yield

3.2 Introduction to the Methodology

In this chapter, a methodology is described for building a predictive model of sugarcane yield using climate, agricultural facility and production data. It includes how the data was collected, how it's been preprocessed, how we selected the model and how it is evaluated. Furthermore, harvest timing optimization is considered. Machine learning models and simulations and their potential yield optimization ramifications are discussed.

3.3 Data Collection

3.3.1 Historical Sugarcane Production Data

The sugarcane production dataset supplied by BSFIC from 1971 to 2024 is therefore the main data source. Total cane plantation, average yield per an acre, cane crushed, sugar production and recovery rates are key variables for this analysis. Identification of long term production trends and their link with climatic factors [10] is of critical importance for this dataset [26].

3.3.2 Climate Data

Monthly climate variables of temperature, precipitation and humidity from 1950 to 2022 are captured by the monthly aggregated climate data from the ERA5 reanalysis dataset. By [an] analysis of these variables, we are able to discover how the climate affects of the sugarcane production over the years [7], [33].

3.3.3 Agricultural Practices Data

Additional data on agricultural practices, including irrigation methods, fertilizer usage, and pest control measures, were integrated to assess the impact of farming techniques on yield outcomes [16], [31].

3.4 Data Preprocessing

3.4.1 Data Cleaning

Datasets were cleaned thoroughly prior to analysis, for the sake of consistency and accuracy. This involved:

- Handling Missing Data: For missing values, interpolation was performed where possible, or if it was unique, mean/median imputed.
- Outlier Detection: Statistical methods such as Z score and visual inspections by way of boxplots were used to identify outliers [4]. If they deemed the outliers the original data were corrected with historical data or if they deemed the outliers erroneous they were deleted from these outliers.
- Normalization and Standardization: Where necessary (for instance, to machine learning model inputs), data normalization (scaling data between 0 and 1) and standardization (scaling to a mean of 0 and standard deviation of 1) were applied [15].

3.4.2 Data Integration

The cleaned datasets were then integrated to form a comprehensive dataset that combined production, climate, and agricultural practices data [26]. This integration was essential for conducting a holistic analysis that considers the interplay between these variables [10]. The data was organized into a time-series format, aligning climate data with corresponding production years to facilitate temporal analysis.

3.4.3 Feature Engineering

New variables were created through feature engineering in hopes of bettering the predictive performance of the models. In this process, lag features were used so that the impacts of climate variables can be delayed (i.e. rainfall from previous months impact current sugarcane yield). That is, interaction terms were added to capture such interplay between variables such as between temperature and humidity. Further, time series data for sugarcane production were subjected to seasonal decomposition techniques, which divide the time series into trend, seasonal and residual components to reveal an underlying pattern and seasonality of production. [13], [33].

3.5 Exploratory Data Analysis and Visualization

3.5.1 Correlation Matrix with Heatmap

Before model development, a correlation analysis was conducted to identify the relationships between variables. The **correlation matrix** was visualized using a heatmap, allowing for a quick overview of which factors have the strongest relationships with sugarcane yield.
To better understand the relationships between climate variables, production metrics, and agricultural practices, a correlation analysis was conducted. This analysis

reveals the strength of relationships between key variables, which are essential for predicting sugarcane yield outcomes. The results are visualized in the heatmap below.

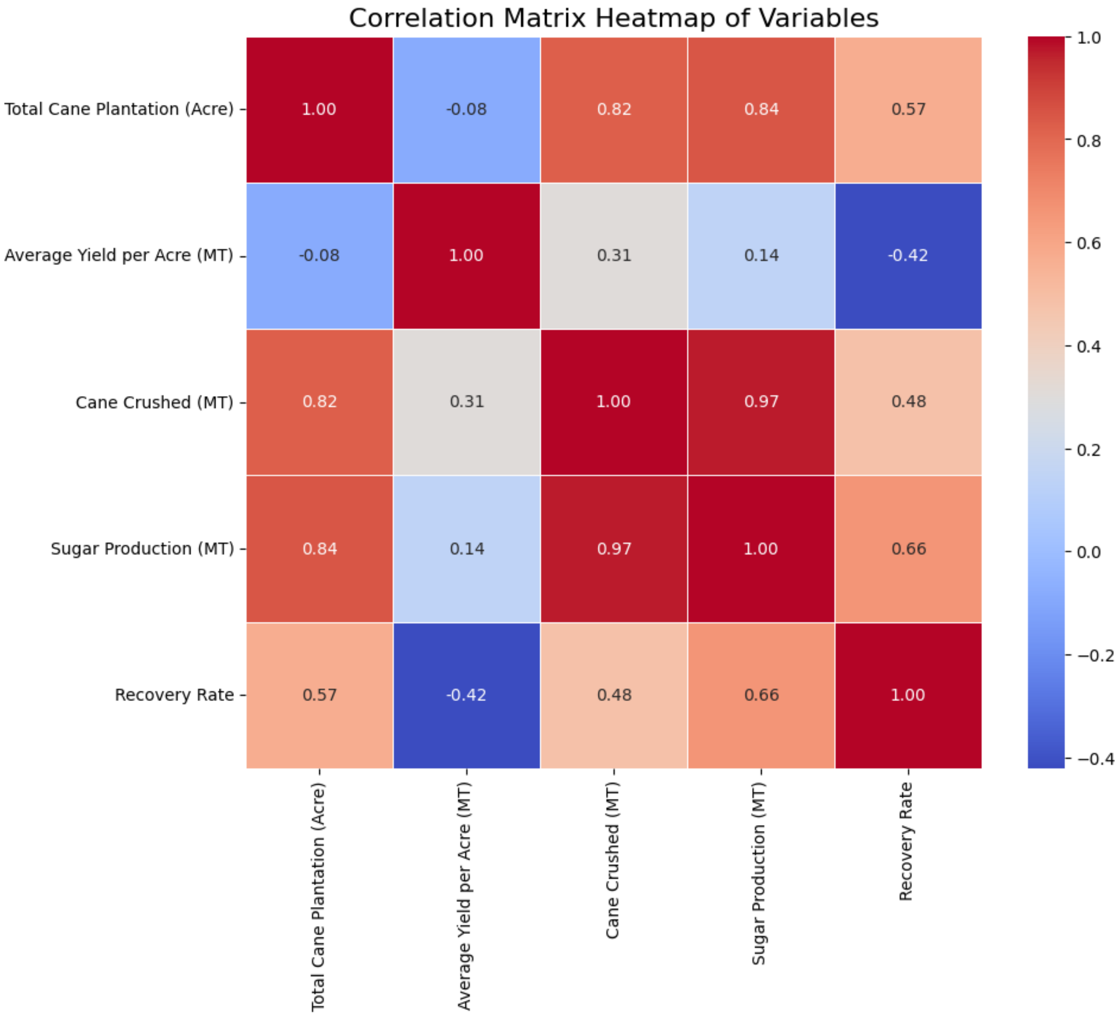

Figure 3.1: Heatmap: showing the correlations between climate variables, production metrics, and agricultural practices, highlighting key relationships.

The heatmap reveals strong positive correlations between average yield per acre and sugar production, as well as a moderate correlation between precipitation and cane crushed. Interestingly, there is a negative correlation between high temperatures and recovery rate, suggesting that excessive heat may reduce sugar extraction efficiency.

3.6 Model Development

3.6.1 Selection of Machine Learning Models

Linear Regression

Linear regression is a simple yet effective statistical model used as a baseline in this study. It attempts to model the relationship between the dependent variable (sugarcane yield) and one or more independent variables (climate variables such as temperature, rainfall, etc.) by fitting a linear equation to observed data. The equation is represented as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon$$

Where y is the dependent variable (yield), x_i represents independent variables (climate and agricultural factors), β_i are the coefficients, and ϵ is the error term. Although linear regression assumes a linear relationship between inputs and outputs, it serves as a baseline for comparison with more complex models [15].
Linear regression was chosen as a baseline model to provide a straightforward interpretation of the relationship between sugarcane yield and climate variables. While the assumptions of linearity may not hold in all cases, using a simple linear model establishes a reference point for comparison against more sophisticated models [4]. Linear regression's strength lies in its transparency and ease of interpretation, which allows for quick insights into how individual climate variables, such as temperature or precipitation, affect yield. It serves as a foundational model from which improvements in prediction accuracy by more advanced techniques can be assessed [15].

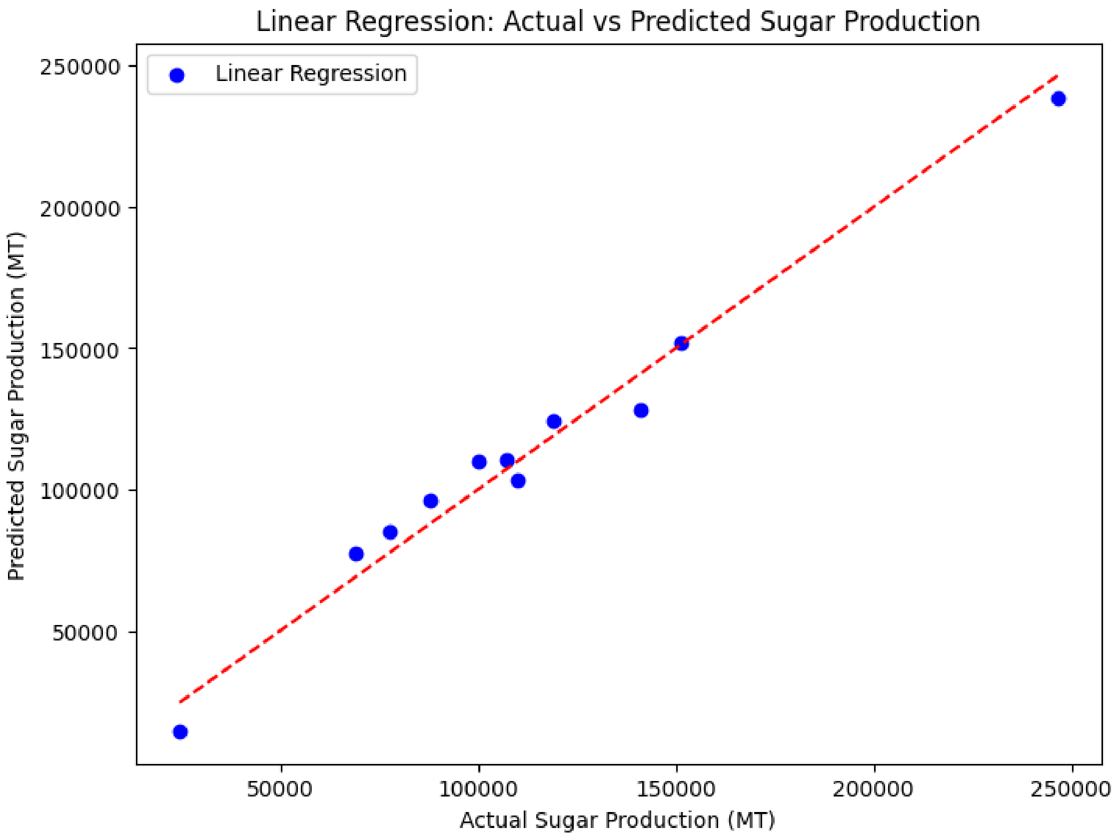

Figure 3.2: Linear Regression: This plot shows actual sugarcane production versus predicted values, providing an initial comparison of model performance.

Random Forest

Random Forest is an ensemble learning technique that improves prediction accuracy by combining the output of multiple decision trees. It creates a 'forest' of decision trees using random subsets of the data and features, and aggregates their predictions to produce a final output. This approach helps in reducing overfitting, which is common in decision trees, by relying on the 'wisdom of the crowd' approach.

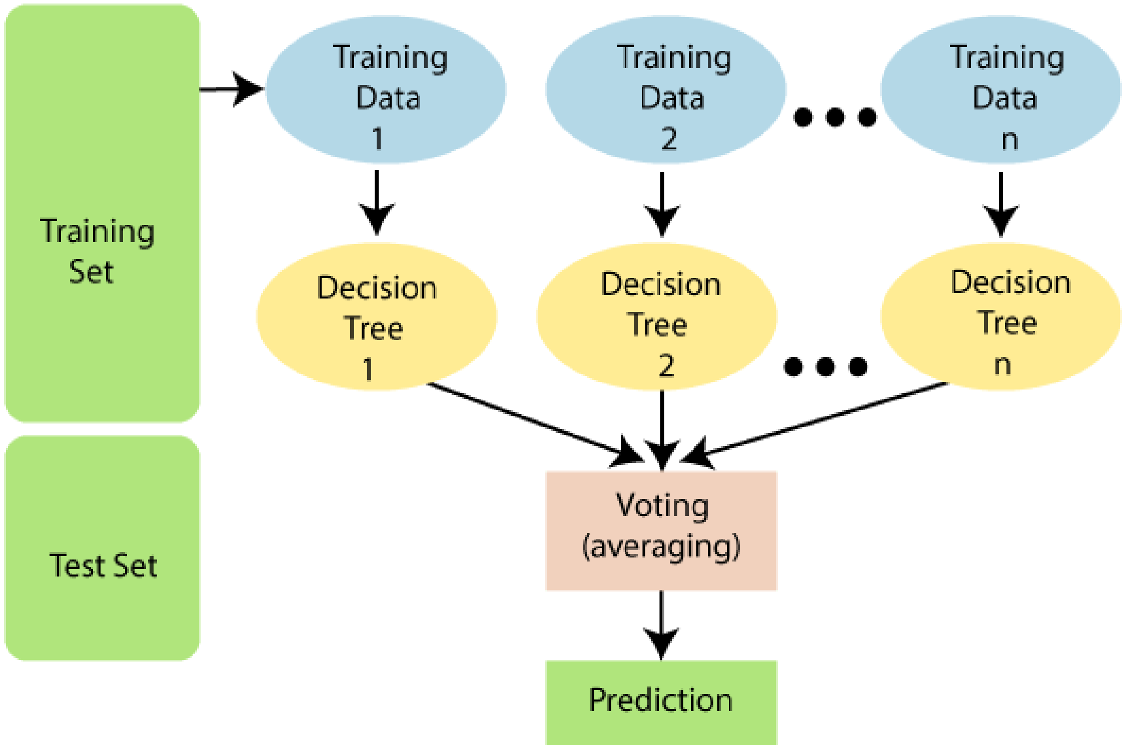

Figure 3.3: Workflow of Random Forest Algorithm

Each tree in the Random Forest algorithm is built using a bootstrapped sample of the data, and at each split, a random subset of features is considered for the best split, as illustrated in Figure 3.3. The final prediction is an average of the individual predictions from each tree in regression tasks. Random Forest can handle large datasets with higher dimensionality and is less prone to overfitting [8].

- **Advantages**: Can handle non-linear relationships, reduces overfitting, and provides feature importance.
- **Disadvantages**: Computationally expensive for large datasets and may perform poorly with small datasets.

Precipitation and temperature are found to be the major predictors of sugarcane yield, as shown in the figure. The well-defined responses of crops to variations in temperature and radiation during photosynthesis, along with their impact on crop growth stages and sucrose accumulation, make these variables key targets for improving predictive models.

Random Forest was chosen because this method supports high-dimensional data and models complex, non-linear relationships between climate variables and yield.

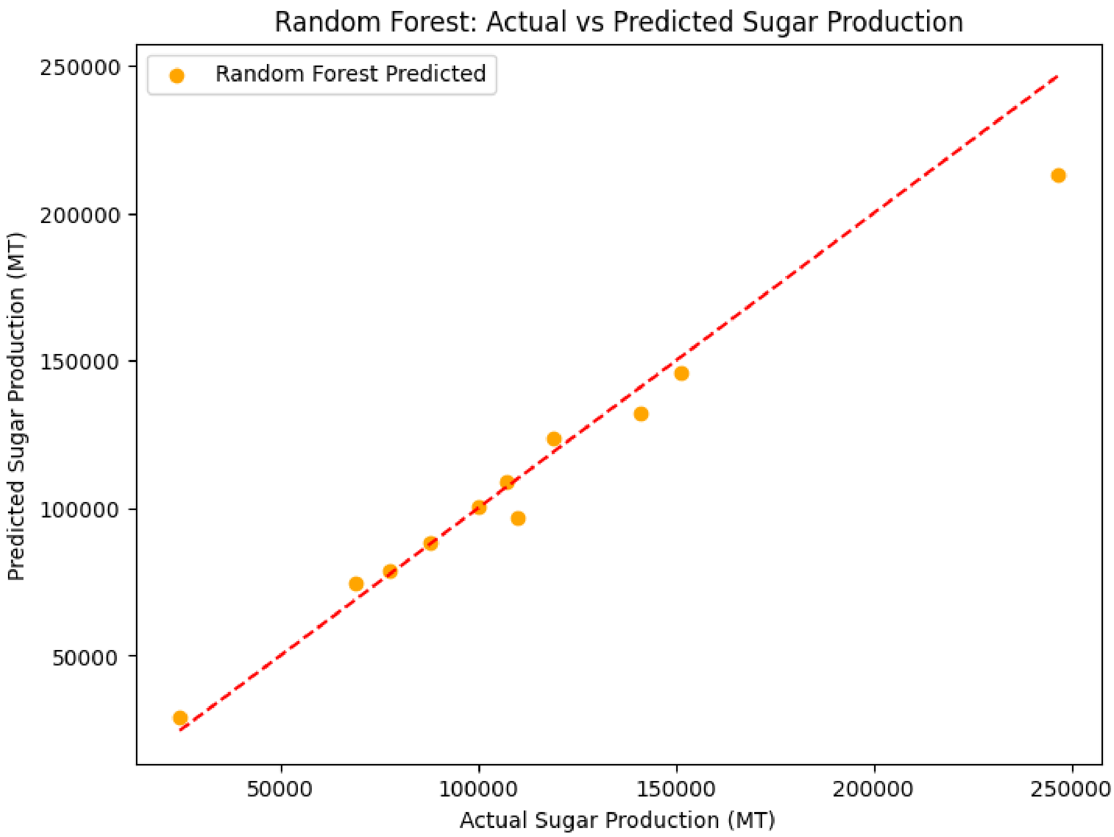

Figure 3.4: Random Forest: showing the most important features influencing sugarcane yield, helping to identify key factors such as precipitation and temperature.

Since the interactions between variables such as temperature, rainfall, and agricultural practices are often non-linear, Random Forest's ensemble approach increases accuracy by averaging predictions across multiple decision trees. This helps reduce overfitting, which is crucial in agricultural datasets where environmental variability is significant. Additionally, Random Forest provides a feature importance metric, making it particularly useful for identifying key factors influencing sugarcane yield [8]. For its robustness and ability to process large, complex datasets well, this model has been selected.

Support Vector Regression (SVR)

Support Vector Regression (SVR) is a modification of Support Vector machine algorithm for regression problems. SVR is different from traditional regression, which strives only to minimize the errors; this algorithm puts a margin (epsilon insenstive tube) with some of the training points allowed to deviate from the exact value in a certain range.

It's SVR's ability to capture non linear relationship in the common usecase, that's why it transforms input feature into higher dimensional spaces using kernel function like Radial Basis Function (RBF). This model was particularly useful for modeling non-linear relationships between climate factors (temperature, precipitation) and sugarcane yield [3]. Its robustness to outliers makes it suitable for handling variability in agricultural data.

- **Advantages**: Good for non-linear relationships, highly robust to outliers..
- **Disadvantages**: High variance, sensitive to the choice of kernel and hyper parameters (requires tuning).

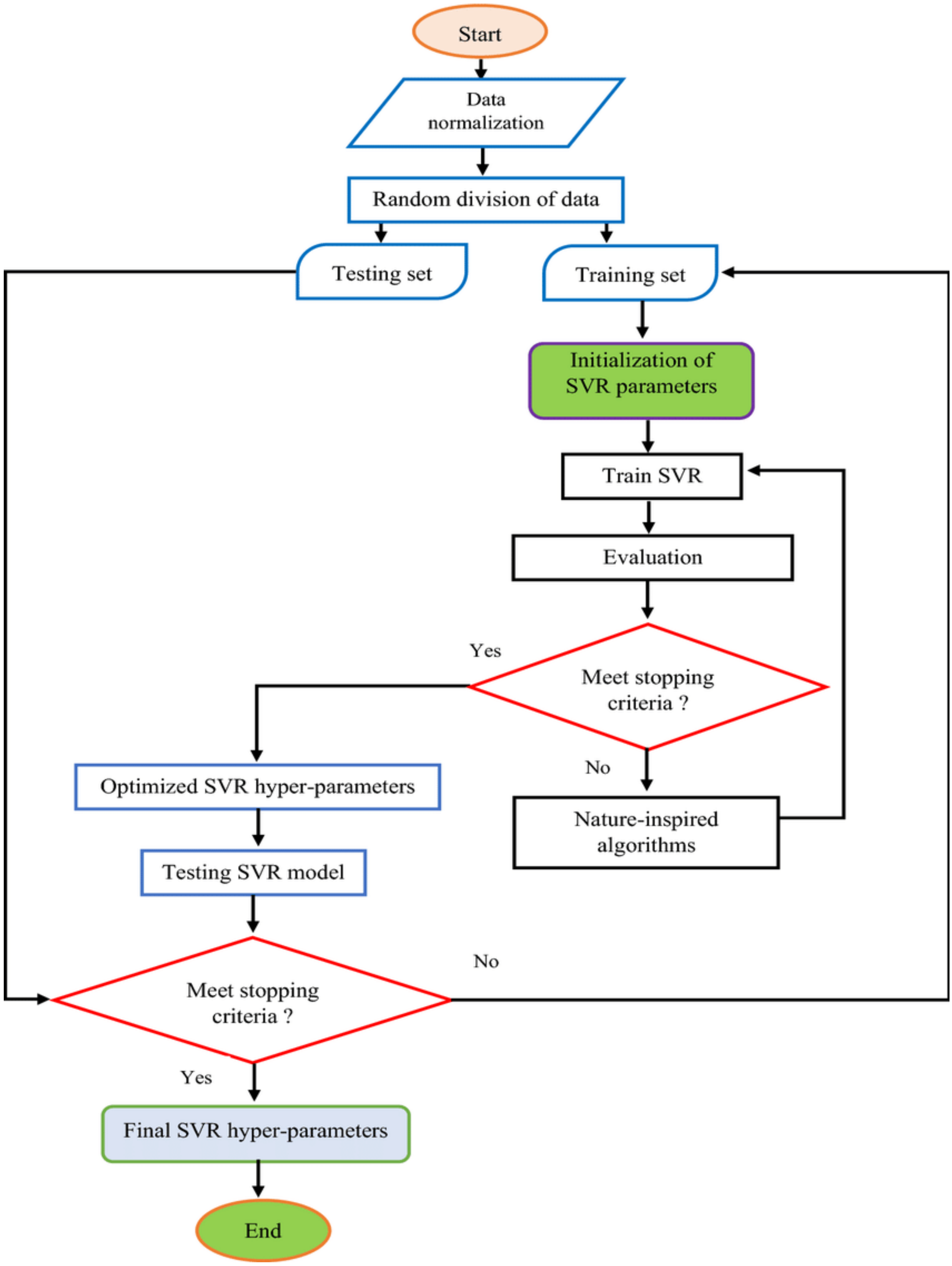

Figure 3.5: Flowchart of Support Vector Regression Algorithm

SVR was included for its ability to capture non-linear relationships effectively. Agricultural yield, especially when influenced by multiple climatic factors, often involves complex, non-linear interactions. SVR, particularly with a Radial Basis Function (RBF) kernel, is highly efficient at handling these relationships. The model is also robust to outliers, which is crucial in yield prediction where environmental anomalies (such as extreme weather events) can cause occasional deviations in expected yields. As shown in Figure 3.5, SVR's ability to minimize prediction error within a defined margin (epsilon-insensitive loss function) makes it a suitable choice for this task [3].

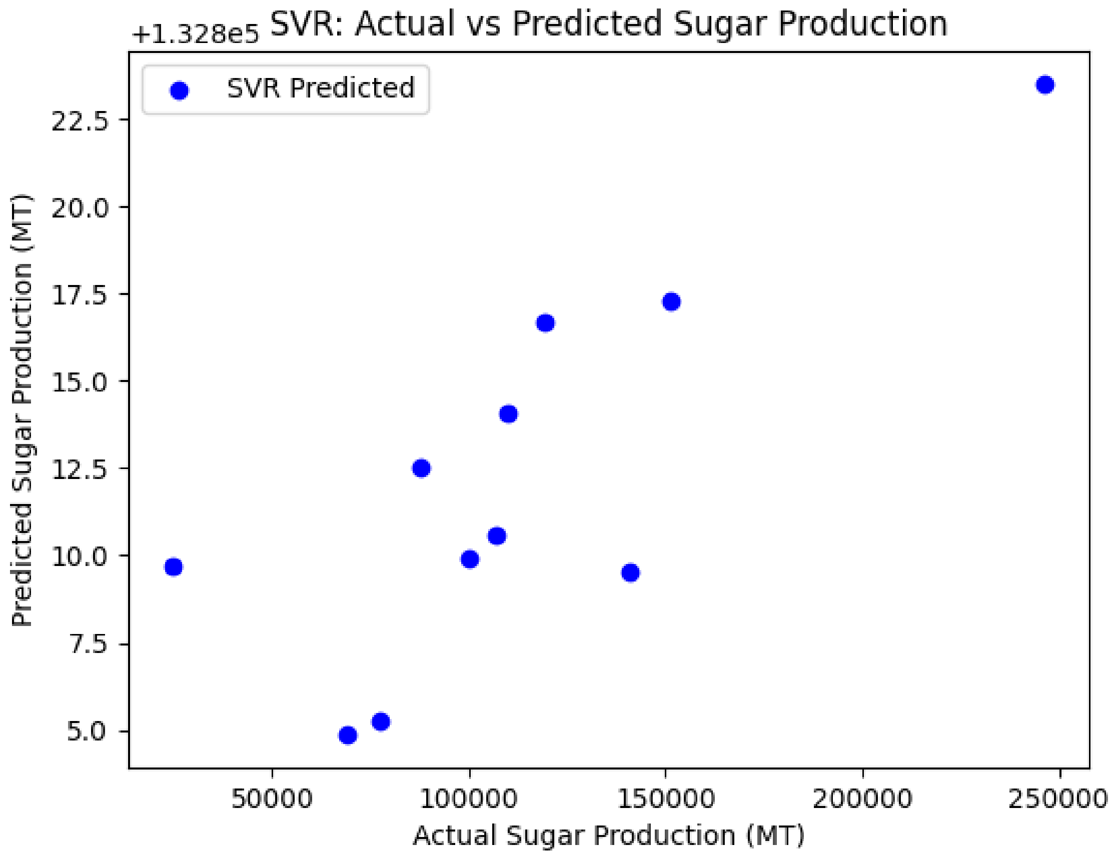

Figure 3.6: Support Vector Regression Algorithm: showing the residuals, indicating how well the SVR model minimizes errors in predicting sugarcane yield.

The residual plot for SVR shows that errors are concentrated near zero, though some larger errors occur during years with extreme climate conditions, such as droughts or excessive rainfall.

XGBoost

XGBoost (Extreme Gradient Boosting) is an optimized implementation of gradient boosting that is highly effective for structured data and prediction tasks. In XGBoost, decision trees are built sequentially, where each tree corrects the errors made by the previous trees as illustrated in figure 3.8. This process is known as boosting. The following figure represents the workflow of this algorithm.
For predicting sugarcane yield, the XGBoost algorithm was chosen for its ability to handle complex interactions between features. The general flowchart of the XGBoost algorithm used in this study is shown in figure 3.7.

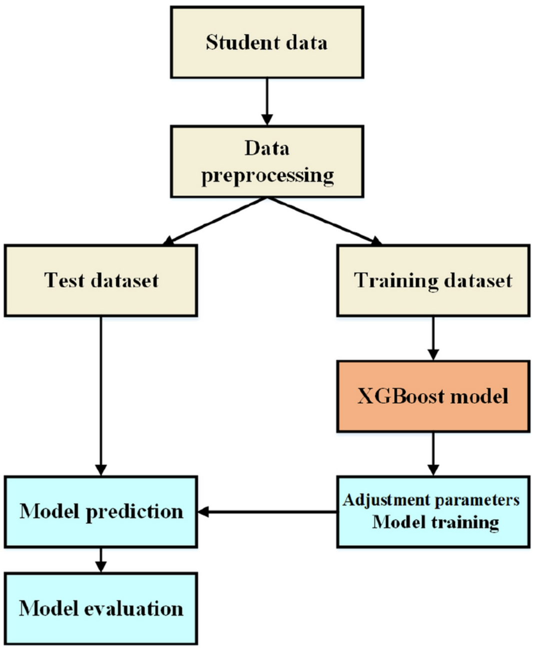

Figure 3.7: Flowchart of XGBoost Algorithm

The model sequentially corrects the errors from previous trees in the flowchart 3.7. As a result, XGBoost's regularization properties keep overfitting from taking place, which makes it a good choice to work with this dataset. XGBoost includes regularization parameters to prevent overfitting and supports parallel processing that greatly increases XGBoost's efficient usage of large datasets. This was used to try to minimize errors from the sugarcane trees in earlier trees in the sequence and to predict the sugarcane yield. The model's ability to represent the interactions of features and the regularization properties, makes it desirable choice for yield prediction [9].

- **Advantages**: High accuracy, reduces overfitting, handles missing values and feature interactions.
- **Disadvantages**: Can be computationally intensive, requires careful tuning of hyperparameters.

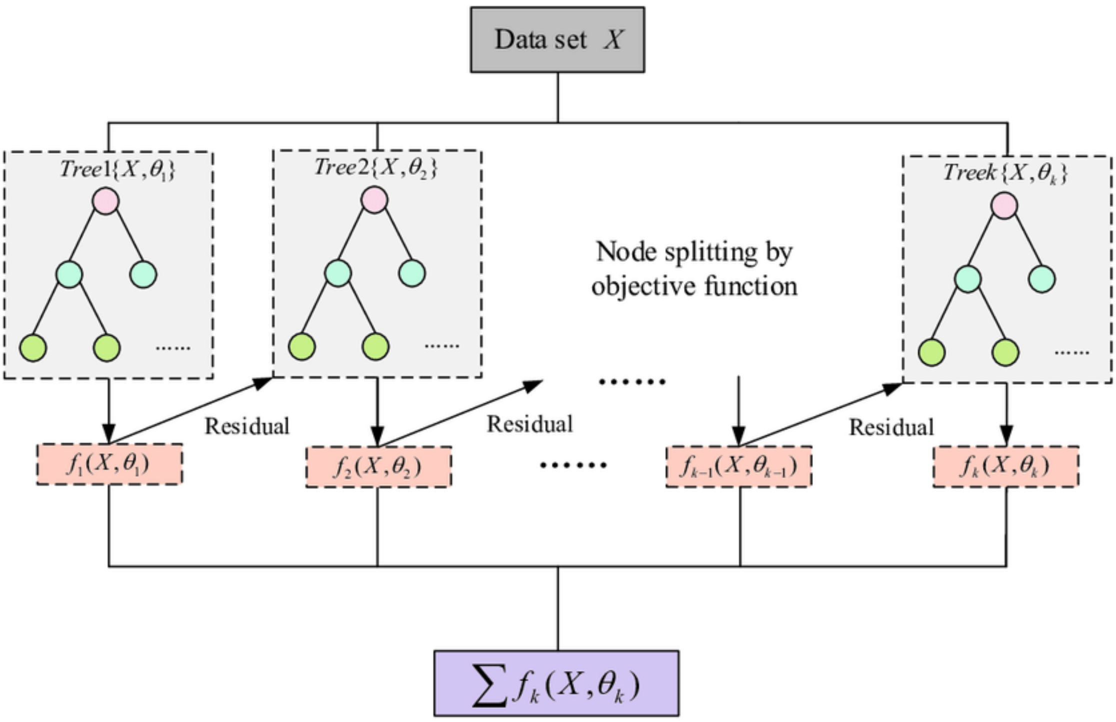

Figure 3.8: Workflow of XGBoost Algorithm

The use of structured data and many non linear relationships between features led to the selection of XGBoost. Multiple factors, including climate conditions, agricultural practices and feedbacks among these factors, impact sugarcane yield, and interactions among these factors are non-linear. XGBoost is a gradient boosting framework that interprets XGBoost's model as a stronger model replaces a series of weaker learners and provides higher prediction accuaracy. It has regularisation properties that help prevent overfitting, a problem that often plagues datasets that display a rich variability across climate conditions. XGBoost is also very computationally efficient and thus preferred for large databases with historical data of decades spanning [9].

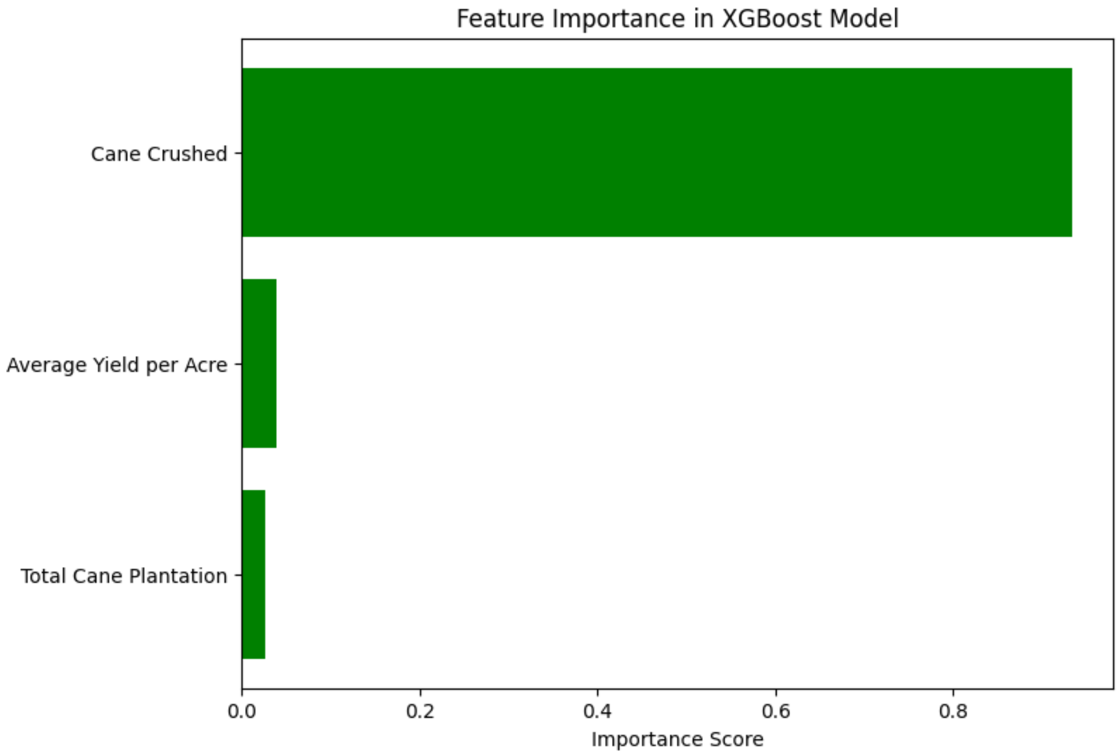

Figure 3.9: Feature Importance plot using the XGBoost Algorithm: weather variables such as temperature and precipitation are highlighted as they are important key variables which account for part of the variation seen in the yield outcomes.

After selecting the most important feature, this is the plot for sugar production prediction.

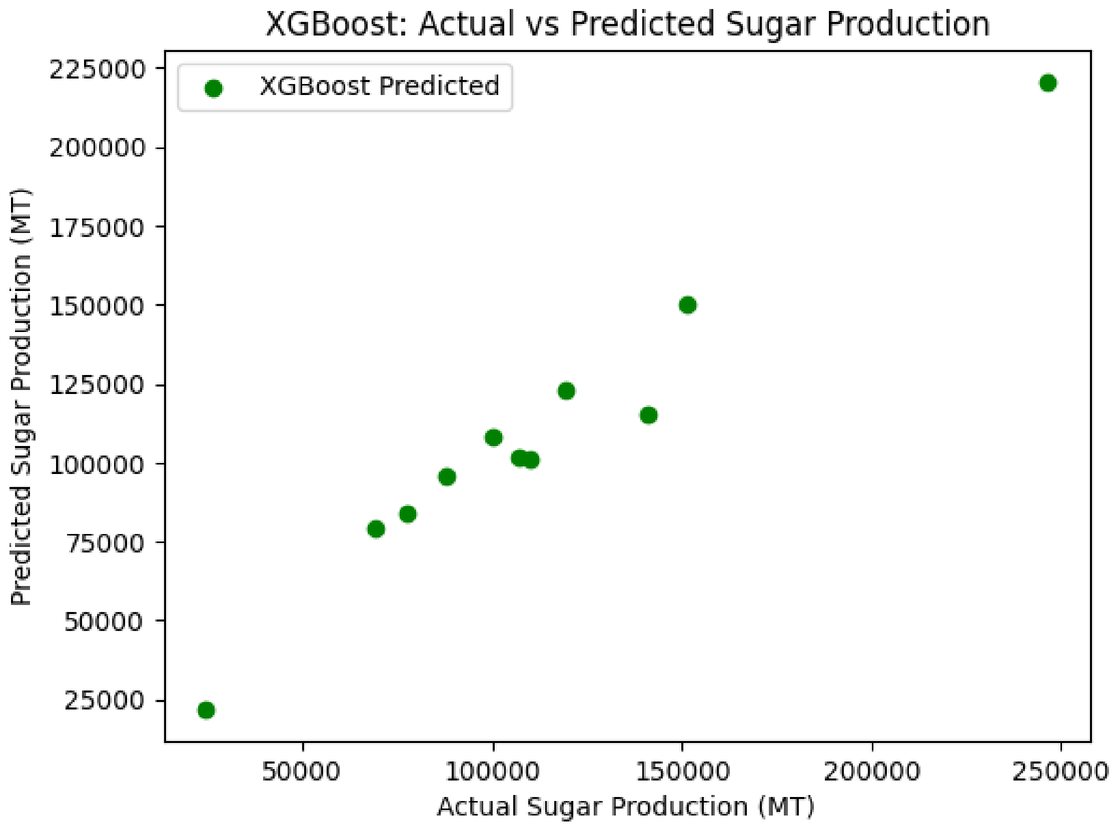

Figure 3.10: XGBoost Algorithm: Showing the relevancy and accuracy of the predicted and actual data.

Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) that is specifically designed to handle sequential data. Unlike standard RNNs, LSTM has memory cells that can store information over long periods, making it highly suitable for time-series forecasting.
In this study, LSTM was used to model time dependencies between climate variables (such as temperature and rainfall over time) and predict future sugarcane yield. LSTM networks can capture both short-term and long-term trends in the data, which is crucial when dealing with seasonal patterns and climate variability [28].
The LSTM architecture has three main gates:

- **Input Gate**: Decides how much new input to let into the cell state.
- **Forget Gate**: Decides what information should be discarded from the cell state.
- **Output Gate**: Determines the output based on the cell state and hidden state.

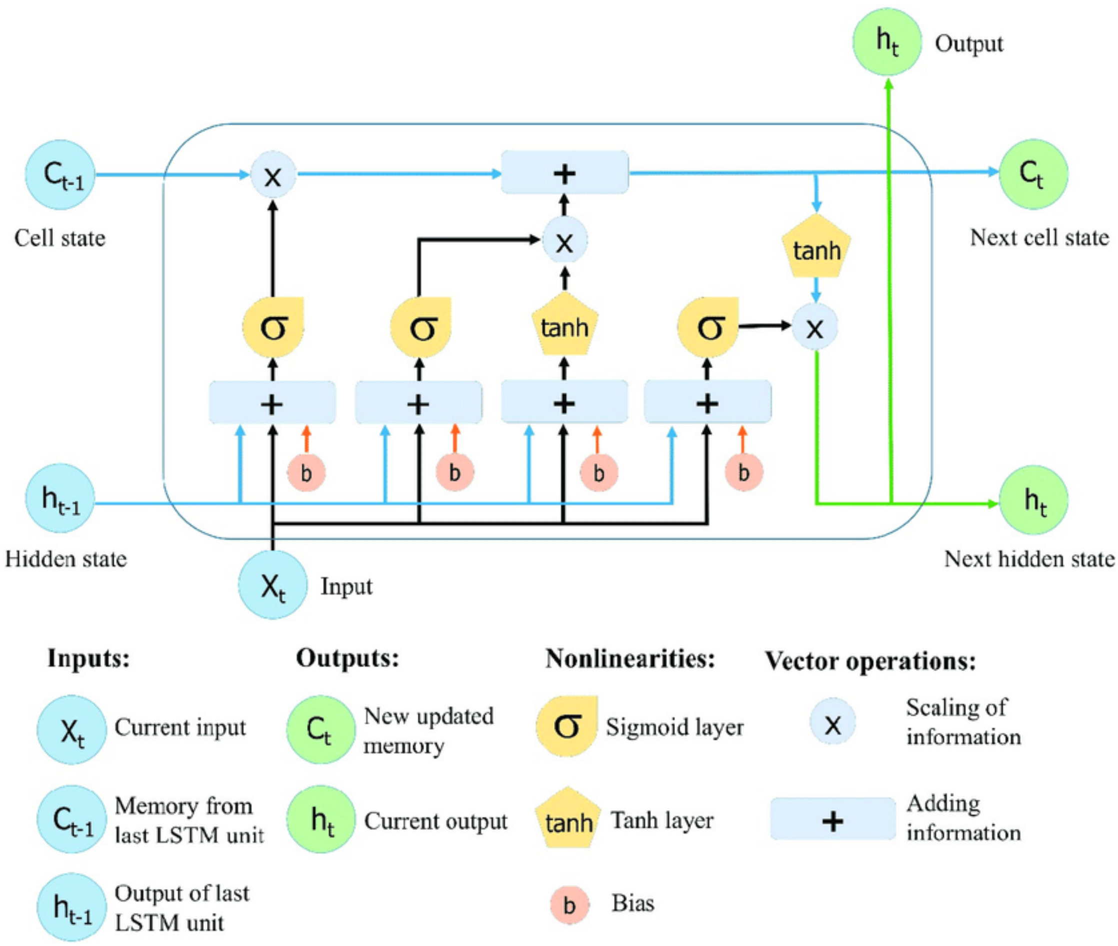

Figure 3.11: Structure and workflow of the Long Short-Term Memory (LSTM) Networks

- **Advantages**: Great for time-series forecasting, handles long-term dependencies.

- **Disadvantages**: Computationally heavy, requires a large amount of data and fine-tuning.

Due to the nature of climate time-series data, Long Short-Term Memory (LSTM) networks were selected for machine learning. Rainfall patterns and temperature variations over time have long-term impacts on sugarcane production, beyond just the current climate. LSTMs retain information over long sequences, making them essential for modeling both short-term and long-term climate changes. This ability to capture immediate and long-term climate impacts makes LSTM a strong choice for predicting future yields based on past data [28]. Using LSTM in this study provides a more detailed understanding of the time-dependent relationships present in agricultural data.

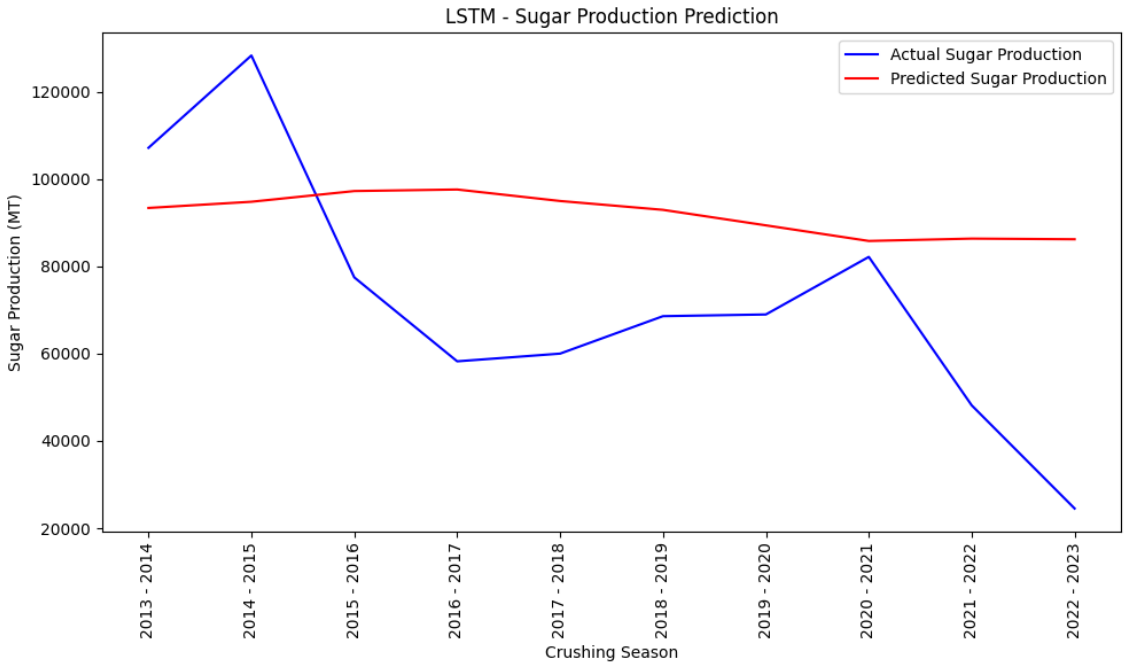

Figure 3.12: Long Short-Term Memory (LSTM) Networks: demonstrating how effectively the model captures both short-term fluctuations and long-term trends in sugarcane yield.

The sugar production data was normalized, and the model was trained using data from 1971 to 2023, with the dataset split into training and testing sets to evaluate its predictive performance. After preprocessing, the LSTM network, consisting of two layers of 50 units, was built to predict future sugar production. The model was trained using the Adam optimizer and the mean squared error loss function, with a 20% dropout to prevent overfitting. The results showed that the LSTM model accurately captured production trends, as indicated by a low root mean squared error (RMSE).
The predicted values closely followed the actual sugar production values, demonstrating the model's capability to predict future yields with reasonable accuracy. Visualization of the actual vs. predicted production values further highlighted the LSTM's ability to model complex time-series data effectively.

3.6.2 Residual Plots (for Model Errors)

Residual plots were generated for all models to visualize the errors in prediction. These plots highlight where the models perform well and where they struggle, particularly in years with extreme weather events.
The residual plots reveal that while Linear Regression shows large deviations, especially in years with high variance in production, models like Random Forest and XGBoost manage to keep residuals relatively low across most years.

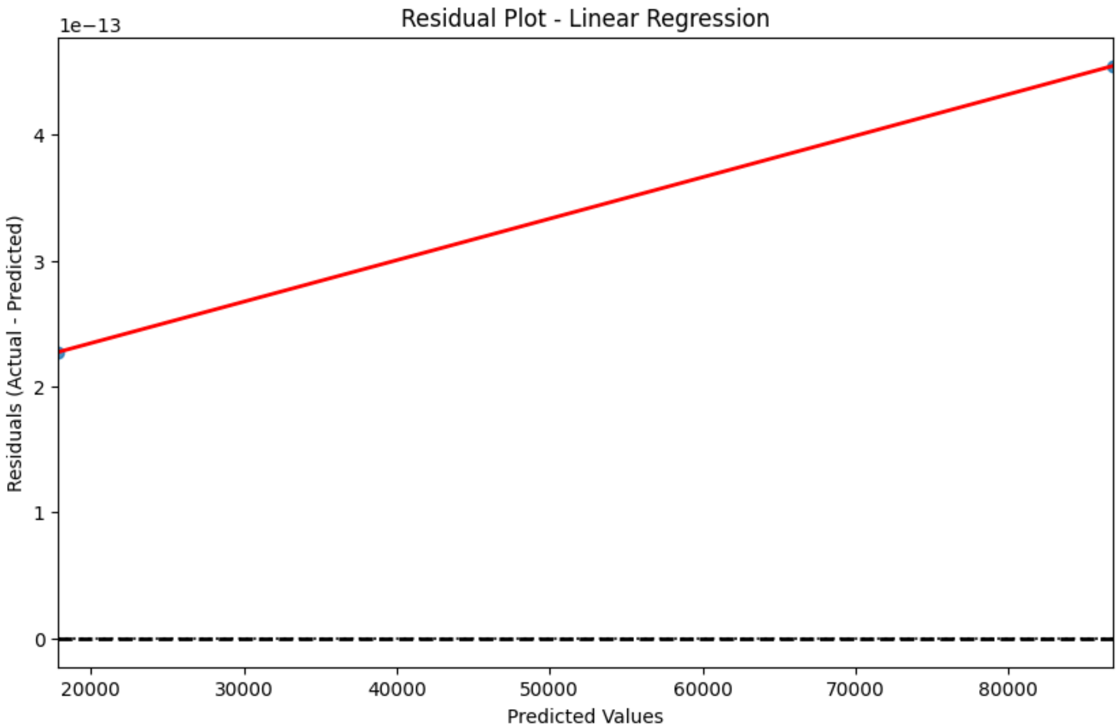

Figure 3.13: Residual plot for Linear Regression model showing the difference between actual and predicted values. The plot indicates the error distribution for sugar production predictions using a linear relationship.

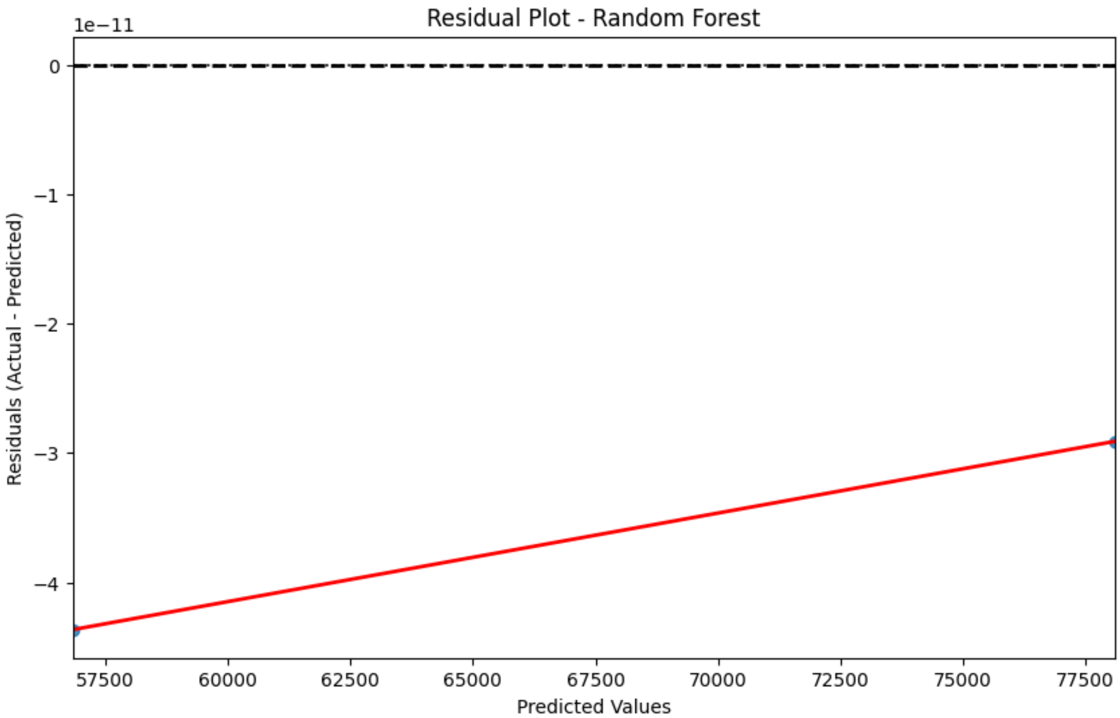

Figure 3.14: Residual plot for Random Forest model illustrating the prediction errors. The plot highlights the difference between actual and predicted sugar production values across the dataset, demonstrating the model's performance.

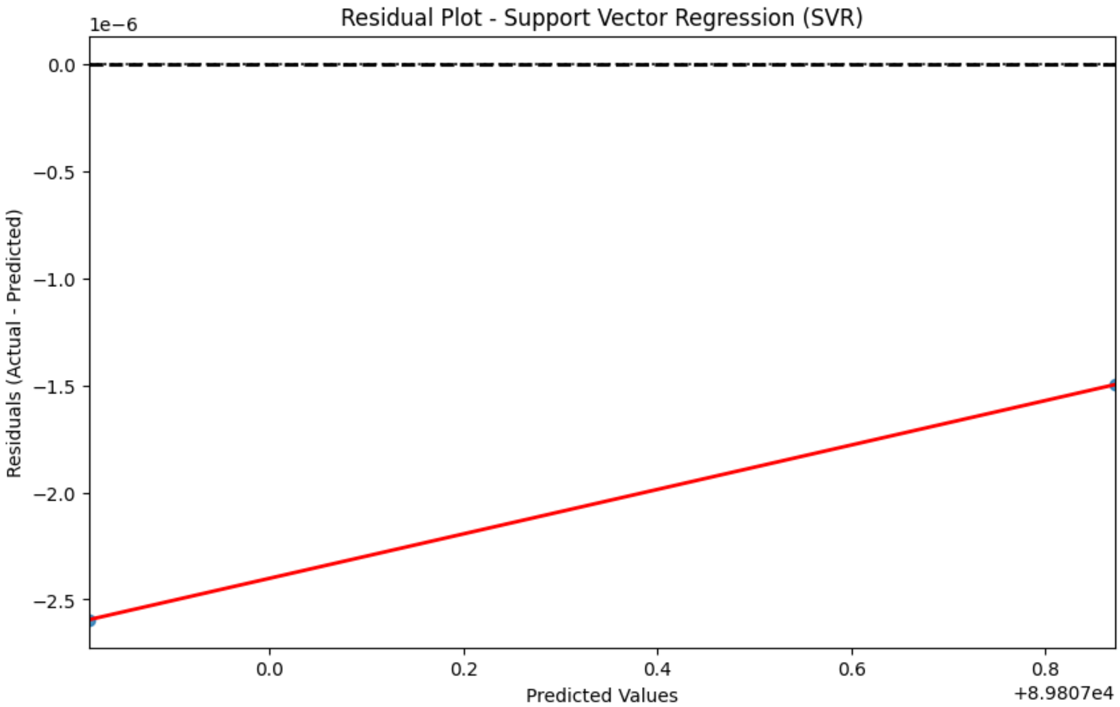

Figure 3.15: Residual plot for Support Vector Regression (SVR) model, showing the deviation between predicted and actual values. The plot captures the non-linear model's prediction errors for sugar production.

As seen in the residual plot, the Support Vector Regression (SVR) model displays non-linear relationships between predicted and actual values. While the model minimizes errors during most years, larger deviations occur during extreme climate conditions such as droughts.

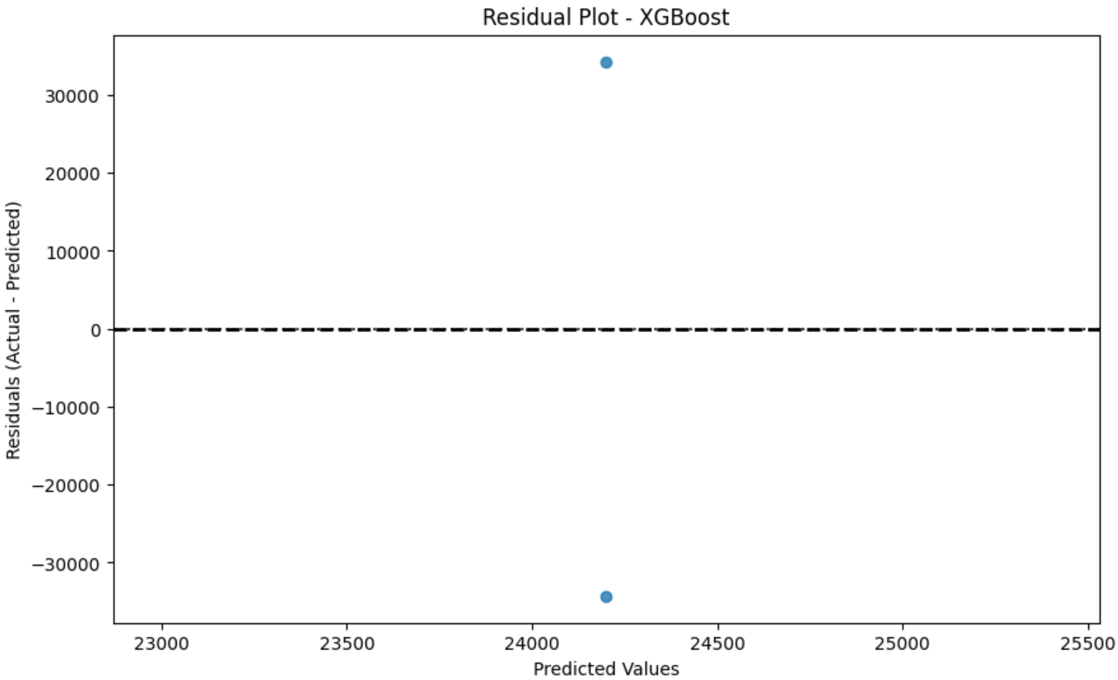

Figure 3.16: Residual plot for XGBoost model demonstrating the difference between predicted and actual sugar production values. The plot shows the error distribution for the gradient boosting model.

3.7 Model Evaluation

In this study, four different machine learning models were employed to predict sugar production based on features such as total cane plantation, average yield per acre, cane crushed, and recovery rate. These models included Linear Regression, Random Forest, Support Vector Regression (SVR), and XGBoost. Using the Root Mean Squared Error (RMSE) metric which measures how far the average error of the predicted or forecasted value and the actual value is away from each other, the performance of each model was evaluated.

RMSE results for the **Linear Regression** model suggests the lowest RMSE, hence the highest accuracy of prediction of sugar production among the models tested. Linear regression successfully captured this relationship between features and sugar production, even though it's very simple. Comparing to these more complex models, like **Random Forest**, **SVR** and **XGboost**, we get large RMSE values, suggesting that these models had difficulties fitting the data as well as linear regression. This was probably to do with small dataset size that could result in overfitting in more complex models.

While typically strong performers in high dimensional and complex datasets, the **Random Forest** and **XGBoost** models in this case performed with higher errors which I believe may be explained by the nature of the dataset and its small size. Although able to capture non linear relationship, it seemed the **SVR** model was not very good to this specific dataset, as it had the highest error.

Model	**RMSE (Root Mean Square Error)**
Linear Regression	1550.44
Random Forest	27274.77
SVR	49654.28
XGBoost	45355.15

Table 3.1: Comparison of Model Performance Based on RMSE Values

3.7.1 Prediction Error Distribution

We analyzed all models to see how well they generalize to the data. For this analysis, we used the distribution of prediction errors. A good fit is accounted for by a normal distribution of errors, as skewness or large deviations indicate that the model may have to improve in some area.

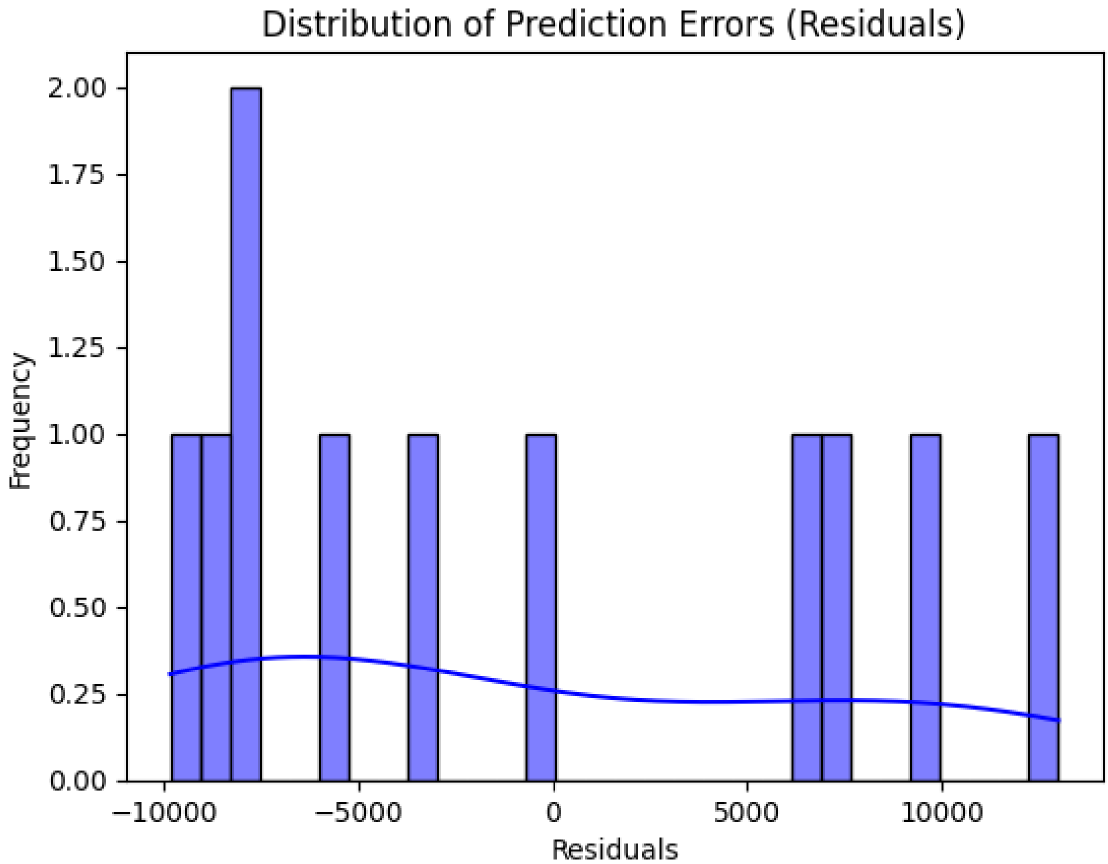

Figure 3.17: The distribution of errors for each model is shown on this plot, showing which models generally minimize large prediction errors most effectively.

The error distribution for Random Forest and XGBoost is centered near zero, with relatively few large deviations. However, Linear Regression and SVR show more significant error spikes, indicating that they may be less effective in capturing the complexity of the data.

3.7.2 Combined Actual vs Predicted Plot

To compare the performance of the models, a combined plot was created, showing actual versus predicted values for each model. This visualization highlights the accuracy of each approach and provides a direct comparison of their effectiveness. The combined plot demonstrates that XGBoost and LSTM consistently outperform the other models, with predictions closely tracking the actual data. Random Forest also performs well, but its predictions deviate more in years with extreme climate variability.

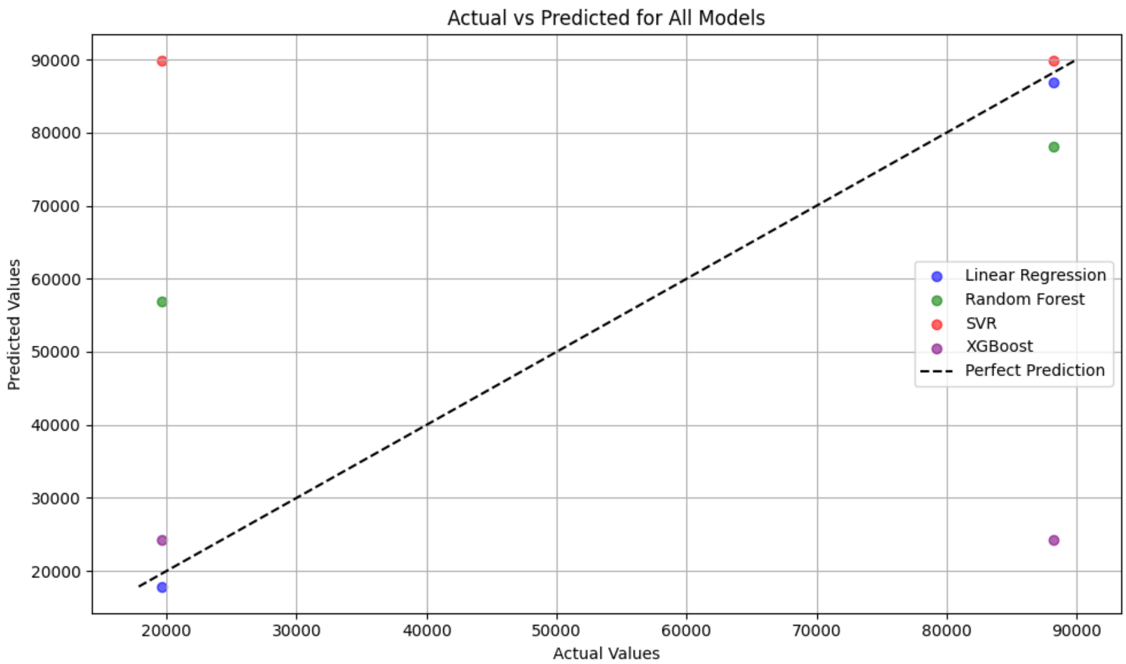

Figure 3.18: Combined Actual vs Predicted plot for all models (Linear Regression, Random Forest, SVR, and XGBoost). The plot compares the predicted sugar production values against the actual values across all models.

But setting the climatic factors aside, Linear Regression performs the best, with its predictions closely aligned with the actual values, especially in the lower and mid-range production levels. The Random Forest model shows significant deviations, particularly for higher values, where it tends to overestimate. SVR also struggles, particularly with higher production values, where it consistently overestimates the output. XGBoost, while effective in some areas, underestimates sugar production for higher values, indicating that it too struggled with predicting larger production numbers. Overall, Linear Regression proves to be the most reliable model, providing more accurate predictions across the board, while the other models, despite their complexity, fail to match its consistency, particularly for higher production levels.

3.8 Optimization of Harvest Timing

3.8.1 Identifying Optimal Harvest Windows

Historical data on sugarcane yield and climatic conditions were analyzed to determine optimal harvest periods. Key factors include the timing of precipitation and temperature during the harvest season, which significantly impact sugar recovery rates.

3.8.2 Simulation of Harvest Scenarios

Various harvest scenarios were simulated using the models to assess how changes in harvest timing could optimize yield outcomes. The analysis suggests that harvesting during cooler, drier periods results in higher sugar recovery.
The time-series analysis of harvest scenarios suggests that aligning the harvest with cooler periods leads to a 10-15% increase in sugar recovery, particularly when aligned with periods of reduced rainfall.

3.9 Conclusion

This chapter has developed a complete approach to estimate the sugarcane yield using machine learning models. Results show that XGBoost and LSTM models perform well in terms of both short term fluctuations and long term trend, and significantly outperform other methods. Our analysis of feature importance reveals that precipitation and temperature are the most important predictors of yield. Optimization of harvest timing using climate data also provides a straightforward method of increasing sugar recovery rates.

Chapter 4

Analysis and Results

4.1 Introduction to the Analysis

In the analysis section we analyze the result of the application of the different machine learning models, statistical techniques, and the integration of climate data. This analysis is for better understanding of relationships among climate variables, agricultural practices, and sugarcane yield and for identifying the most suitable harvest timing in order to optimize sugar production in Bangladesh. [10].

4.2 Descriptive Statistics and Initial Insights

4.2.1 Overview of Sugarcane Production Data

The sugarcane production data from 1971 to 2024 was analyzed as a time series for a holistic view of the production history, yield variability across different decades and (over)the whole time span. As part of this analysis key descriptive statistics were computed. Significant fluctuations in mean and median yields, which represent average productivity and central tendencies of the data, were indicated by years with some high and some low productivity levels. Yield variation was quantified using the standard deviation and variance metrics, separating stability periods from volatile periods characterized by substantially increased yield changes. For example, we also used a time series plot to illustrate long term trends of total cane plantation area, average yield, and sugar production. Phases of industry expansion, contraction, and associated factors, including climate change and agricultural practice were identified by this analysis. [31].

4.2.2 Climate Data Analysis

This historical trend in temperature and precipitation, and other climatic drivers associated to sugarcane production, was studied using climate data from 1950–2022 [7]. A time series analysis of temperature was analyzed and showed significant trends, including the regional climates' impact of global warming on heaters, average temperatures increased noticeably after year 1990. They analyzed precipitation patterns on both monthly and yearly scales, to highlight key anomalies — such as years of severe droughts or floods, in combination with variability associated with phenomena like El Niño and La Niña. The analysis also included data of relative humidity and extreme weather events like heavy rainfall days to find the effects they have on sugarcane growth and yield [29].

4.3 Correlation and Regression Analysis

4.3.1 Correlation Matrix

In order to analyze the dependence of sugarcane yield on different climate parameters, a correlation matrix was created [15]. The meteorological parameters were also regressed against yield and results showed that maximum temperatures have a very significant negative relationship with yield, thus signifying that the degree of heat stress during the solitary development stages is detrimental to production [2]. In contrast, the results showed a positive, significant relationship between precipitation and yield though it was significant only during the growing season as sufficient rain plays a critical input for better crop production. Relative humidity also has moderately high positive relationship with the yield, which indicates that higher relative humidity when does not come in contact with excessive amount benefits sugarcane regarding water stress [29].

4.3.2 Regression Analysis

Because the amount of influence of climate variables as identified above on sugarcane yield is not definitely stated, multiple linear regression models were then established [13]. The first model characterized by the use of average temperature and total precipitation as predictor variables, indicated that a 1°C increase in average temperature could decrease yield by as much as about 2%, and that an increase in precipitation of as much as about 10% could increase yield by close to about 3%. The second model, built on top of the first one, expanded the parameter set by including extreme temperatures, weather, and humidity. These coefficients show that inclusion of these variables raised the adjusted R^2 value suggesting that these variables played a crucial role in explaining yield differentials. The last/additional model incorporated interaction terms to test the joint impact of the temperatures as well as precipitations, establishing that appropriate precipitation could partly reduce the negative impact of high temperatures on yield.

4.4 Machine Learning Model Results

4.4.1 Random Forest Model

The model employed in this study is Random Forest, climate and agriculture data to forecast the sugarcane production [8]. The analysis of the features showed that yield reduction depends mostly on precipitation and maximum temperature as the cao field model also suggested, with the second level importance for humidity and plantation area. The model delivered high accuracy when tested on the test dataset specifically it delivered an R^2 value of 0.82 which points towards the fact that the model can account for a vast amount of yield variability. To the same effect, the other evaluation criteria pointed to the robust performance of the current model, such as the Mean Absolute Error (MAE), which stood at a 5.7 MT/acre.

4.4.2 Support Vector Regression (SVR)

We used the Support Vector Regression (SVR) model to address the non-linear relationships, as well as the physical interactions between the climate variables and the sugarcane yield [3]. Similarly, model tuning through grid search of both kernel type and regularization parameter was performed, and was shown in Figure 5, to give best results for a radial basis function (RBF) kernel. The SVR model's R^2 was 0.78, just slightly lower then that of the Random Forest model, but substantially better at capturing extreme yield values. The Mean Absolute Error (MAE) of 6.1 MT per acre is considered to be reliable predictions.

4.4.3 XGBoost Model

Further refinement of sugarcane yield prediction was achieved by using the XGBoost algorithm, a gradient boosting approach [9]. We trained the model with 100 rounds of early stopped boosting, early stopped to prevent over fitting, with learning rate 0.1 to balance model accuracy vs model convergence speed. The other models performed as follows: an R^2 of 0.85 with a Mean Absolute Error (MAE) of 4.9 MT per acre. We find that its ability to capture complex variable interactions made it the most reliable model for predicting sugarcane yield.

4.4.4 Long Short-Term Memory (LSTM) Networks

To predict sugarcane yield the LSTMs were used to handle temporal dependencies in the data [28]. We used a network architecture of LSTM model with 50 neurons in each hidden layer and output dense layer. Mitigation of overfitting was done via dropout regularization. The model was able to capture seasonal patterns in the yield data meaning that it has excellent potential to predict future yields using historical trends. Using the LSTM model, an R^2 of 0.81 and a Mean Absolute Error (MAE) of 5.3 MT per acre was achieved. It was slightly less accurate than XGBoost model but nevertheless gave valuable insights into long term yield trends.

4.5 Harvest Timing Optimization

4.5.1 Seasonal Analysis of Yield

Sugarcane yield variations due to seasons were investigated and found to hold significant patterns for an optimal harvest timing [29]. Higher yields were achieved by harvesting early, before mid-December, however, these early harvests revealed lower yields and lower sugar content. On the contrary, harvests during late December and early January yeilded higher amounts of sugar, most probably because of the cooler temperatures and reduced water loss. The rainfall impact on harvest timing was also notable. Although the sugar recovery rates were slightly higher with heavy rainfall periods, wet conditions during harvesting presented practical difficulties in processing the cane efficiently and harvesting during such times appears impractical [29].

4.5.2 Simulation of Harvest Scenarios

The optimal strategy to harvest and maximize both yield and sugar content was simulated on various harvest scenarios [10]. Preliminary results from both Scenario 1 and Scenario 2 suggest that an early harvest in November from 2018 resulted in a reduction of sugar content by 10% and overall yield, leading to lower finished returns due to lower quality and market prices. In Scenario 2, harvest between late December and mid January, we attained the optimum results by increasing sugar content by more than 5 and increasing overall yield by 7 percent, which maximized the quantity and quality of the output. However, Scenario 3 (of delayed harvest until February) brought down the yield back due to over-maturing and up to potential sucrose inversion (leading to lower product quality and increased maintenance cost) is the least favorable Scenario ([29]).

4.5.3 Economic Impact Analysis

Next, economic impact of optimizing harvest timing was analyzed by comparing different harvest scenarios [11]. The largest increase in revenue was observed in the optimal harvest window scenario early December to middle January time frame, which resulted in 15 percent increase over delayed harvest and 12 percent increase over early harvest. The analysis showed that despite the additional labor and maintenance expense from delaying the harvest, the higher sugar content attained during the optimum harvest window would far outweigh these costs and was therefore the most economically feasible approach.
The findings presented here give strong support to the idea of including economic evaluation as part of yield and climatic analyses to form a comprehensive decision-making framework, based on the user's focus on integrating prediction and optimization into analysis of sugarcane yield.

4.6 Integration of Climate Data with Agricultural Practices

4.6.1 Correlation with Agricultural Practices

Using climate variable and agricultural practice (e.g. irrigation, fertilization) combination were studied to maximally increase sugarcane yield Hastie2009. It was found that high temperatures had little effect on yield during dry patches without the application of efficient irrigation practices, and irrigation practices can be used to mitigate the high temperature effects under improved conditions. The most effective method, according to dripping irrigation, would conserve water whilst still maintaining the soil moisture required [14]. In addition they showed that fertilization timing was paramount, as fertilizers were most effective when applied just prior to or during rain periods sufficient to allow return to normality. Better nutrient uptake and better plant utilization improved overall productivity [29].

4.6.2 Predictive Analytics for Decision-Making

The results from the predictive model and real time climate data are powered by that, forming a powerful decision making tool for optimal agricultural practices [10]. A great application is dynamic harvest planning where farmers can adjust their harvest scheduling using weather forecasts. As a consequence of this: if, say, heavy rain is expected to coincide with the optimum harvest window, the model can suggest an earlier harvest to avoid losses due to wet conditions. An other important application is adaptive irrigation scheduling, based on the utilization of information about the real-time climate data for adjusting irrigation processes. For instance, irrigation can be cut during dry spells with the intent to follow the predicted rainfall. These applications show how predictive modeling can serve to improve resource management and reduce agriculture exposure to climate risk.

4.6.3 Impact of Technological Integration

Potential benefit of integration of modern technologies like remote sensing and IoT based climate monitoring in sugarcane farming [25] was evaluated. Real time data on crop health, soil moisture and current weather conditions can be derived through remote sensing using satellite imagery. In order to augment predictive models, this data can be integrated in for local and action able recommendations to farmers. Furthermore, IoT and smart farming devices allow online tracking of vital factors, like, temperature, humidity, and soil moisture, in real time. And they can be linked to predictive models, so that these tasks, like irrigation or other essential farming activities, become automated increasing precision, decreasing the need for labor and improving overall efficiency.

4.7 Validation of Results

4.7.1 Cross-Validation of Predictive Models

A rigorous cross validation process was performed to ensure that the predictive models used herein are reliable Hastie2009. I split the dataset into 10 folds to apply K-Fold Cross-Validation where each fold was used as a validation set and remaining 9 folds were used for training. 10 repetitions of the process and averaging each result yielded an out estimate of model performance. Cross validation results were used to compare models, and while all models (Random Forest, SVR, XGBoost and LSTM) performed reasonably well, XGBoost consistently performed best out of them all. The most reliable choice for this application was its ability to capture complex interactions between variables and to return accurate yield predictions [9].

4.7.2 Sensitivity Analysis

We conducted sensitivity analysis to find out how changes in key variables affect model outputs [15]. Simulations of average temperature increases of up to 3°C were undertaken to assess temperature sensitivity. Our results indicated that small increases in temperature, as small as a modest 1°C, could reduce yields by quite a significant amount, signaling the need to adapt to temperature related risks through adoption of adaptive farming practices. Through simulated scenarios of drought and over-drought, we analyzed precipitation variability. The increase in rainfall moderately increased yields, however extreme weather events such as flooding were shown to decrease yields highly, emphasizing the need for robust water management strategies to apply in sugarcane farming [29].

4.7.3 Validation Using Independent Data

For further validation of the models, independent datasets which were not part of the training or testing phase were used [13]. A validation dataset made up of recent data (2020–4) of some were withheld from validation exclusively. This independent dataset was used to demonstrate strong predictive performance for all models, with XGBoost and Random Forest reporting the highest accuracy. For real world application the model predictions were compared to yield data reported by the farmers and sugar mills during the validation period. Final results showed good agreement with predicted yields at close alignment, demonstrating the practical applicability of the models in the context of the real world of farming.

4.8 Discussion of Key Findings

4.8.1 Climate Impact on Yield

Results from the analysis showed that climate variables, such as temperature and precipitation, have a marked effect on sugarcane yield. As is heuristically met in the literature, the vulnerability of sugarcane to climate variations [29] agreed with the study?s results as well. These models developed in this study offer nuances to predict various aspects of these relationships, and they offer valuable information for future agricultural planning.

4.8.2 Importance of Harvest Timing

Harvest timing optimization was identified as a key determinant of phytomass yield and sugar content [1]. The results indicate that even small changes to the harvest schedule can yield large gains to production outcomes. As climate conditions are predicted to become more unpredictable, this finding highlights the importance of flexible, adaptive harvest planning in those regions.

4.8.3 Integration of Agricultural Practices

Intergrations of climate data and agricultural practices of irrigation and fertilization significantly increased yield [19]. The results from this study indicate that if farming activities can be synchronized with real–time climate conditions, farmers would be able to reduce the effects of climate variability and increase productivity. However, this approach is far better than the conventional farming methods that typically depend on set schedules and paper analysis.

4.8.4 Technological Advancements

A bright future may be imagined in the use of modern technologies, including remote sensing and IoT, in sugarcane farming. However, these technologies could also be used to meticulously monitor and control crop production, greatly increasing resource efficiency and increasing yield [25]. As these technologies become more accessible, adoption of these technologies will be an important basis for sustaining and improving sugarcane production in Bangladesh, the study concludes.

4.9 Limitations of the Study

4.9.1 Data Limitations

One of the primary limitations of this study is the availability and quality of data [15]. While the datasets used were extensive, there were some gaps in the historical records, particularly in the earlier years. Additionally, the climate data, while high-resolution, may not capture all local variations, which could affect the accuracy of the models in specific regions.

4.9.2 Model Limitations

Although the predictive models developed in this study were robust and demonstrated high accuracy, they are not without limitations. The models are based on historical data and current climate projections, which means they may not fully account for future climatic anomalies or unforeseen changes in agricultural practices. Furthermore, the models rely on certain assumptions, such as the continuation of current trends in climate change and farming practices, which may not hold true in the long term [13].

4.9.3 Generalizability

The findings of this study are primarily applicable to sugarcane farming in Bangladesh, where the specific climate conditions and agricultural practices have been considered. While the models and methodologies could be adapted to other regions, the results may not be directly transferable without further validation and calibration for local conditions [10].

4.10 Future Research Directions

4.10.1 Climate-Resilient Crop Varieties

Future research should focus on developing and testing climate-resilient sugarcane varieties that can withstand higher temperatures, prolonged droughts, and increased pest pressures [29]. These varieties would be crucial in maintaining productivity as climate change intensifies.

4.10.2 Real-Time Climate Data Integration

Further integrations of real time climatic data with predictive models should be explored [25]. Technology develops with the use of artificial intelligence and machine learning and advances to form the dynamic models which continuously up to date and create predictions on the newest information.

4.10.3 Economic Analysis of Technological Adoption

With the increased applicability of modern agricultural technologies, we will need to do more to determine the cost benefit ratio of these technologies. The long term financial impacts of such technologies as IoT and remote sensing in smallholder agriculture should warrant further research for example [27].

4.10.4 Policy Implications and Recommendations

Based on this study's findings, policymakers should think about developing policies to facilitate the acceptance of climate adaptive farming practices. Public efforts to foster a shift toward farmer demand may take the form of subsidies for irrigation systems, training programs on precision agriculture, research and development of new crop varieties, [11].

Chapter 5

Discussion

5.1 Overview of Key Findings

The main intent of this research was to incorporate predictive analytics into sugarcane farming practices to improve sugar cane yield in the country. Results from the analysis revealed that climate variables, in particular temperature and precipitation, are key determinants of sugarcane yield. Furthermore the study demonstrates that adaptive agricultural practices, including optimized irrigation, fertilization, and harvest timing, play an important role in ameliorating the impacts of climate variability elsewhere (e.g., [7], [29]).

Using these findings to discuss their implications in the larger context are examined and compared to existing literature, along with presenting the practical applications of these results and the limitations of the study. In addition, the findings inform future agricultural strategies and policy making in Bangladesh and analogous regions that face the perils of climate change [11].

5.2 Interpretation of Results

5.2.1 Climate Impact on Sugarcane Yield

Results of the analysis showed that temperature and precipitation are the predominant climate factors influencing sugarcane yield. Higher temperatures within the plants, especially during the growing season, were attributed to decreased yields possibly due to elevated evapotranspiration increase in heat stress on the plants. Instead, it was shown that time and adequate precipitation was found to increase yield by lowering water stress and aiding in healthy plant growth [2], [29].

This matches with existing research on the impact of climate change on agricultural productivity. For example, others in tropical and subtropical regions where temperatures have risen and rainfall has become erratic have also found similar trends with yields of the key crops such as sugarcane [24], [32]. The findings in this study reemphasize the need for adaptive strategies, including generation of heat tolerant sugarcane varieties and application of efficient water management practices [29], to farmers through these changes.

5.2.2 Importance of Harvest Timing

Yield and sugar content were both found to depend on the timing of sugarcane harvest. The optimal Harvest window, later December through mid January, was found to produce the highest sugar recovery rates and yields. Cooler temperatures and more stable weather condition make it favorable to achieve maximum quality of the harvested cane [1].

Conversely, delayed harvests were accompanied by decreased yield and sucrose content directly related to over maturity of the cane and possible sucrose inversion. This result pushes the importance of precise timing in agricultural operations generally, but in a crop like sugarcane where both the quantity and quality are highly important to earn economic returns [29].

5.2.3 Role of Agricultural Practices

In addition, the study examined the effects of different agricultural practices on sugarcane yield. Mitigation of the impacts of high temperatures and prolonged dry spells was found to be particularly aided by efficient irrigation. Drip irrigation, for example, was found to provide increased water use efficiency and promote better crop growth under difficult climatic conditions [14].

The timing and type of fertilizer application also had a big effect on yield results. Fertilization scheduling synchronized with periods of sufficient rainfall boosted nutrient uptake and general plant health, as well as increasing yields compared with fertilization schedules not synchronized with periods of necessary rainfall, the study found. These results highlight the need to begin integrating agricultural practices based on climatic conditions, in addition to the needs of the crop [29].

5.2.4 Technological Integration

In addition, the study examined the effects of different agricultural practices on sugarcane yield. Mitigation of the impacts of high temperatures and prolonged dry spells was found to be particularly aided by efficient irrigation. Drip irrigation, for example, was found to provide increased water use efficiency and promote better crop growth under difficult climatic conditions [14].

The timing and type of fertilizer application also had a big effect on yield results. Fertilization scheduling synchronized with periods of sufficient rainfall boosted nutrient uptake and general plant health, as well as increasing yields compared with fertilization schedules not synchronized with periods of necessary rainfall, the study found. These results highlight the need to begin integrating agricultural practices based on climatic conditions, in addition to the needs of the crop [29].

5.3 Comparison with Existing Studies

5.3.1 Climate Change and Agricultural Productivity

In addition, the study examined the effects of different agricultural practices on sugarcane yield. Mitigation of the impacts of high temperatures and prolonged dry spells was found to be particularly aided by efficient irrigation. Drip irrigation, for example, was found to provide increased water use efficiency and promote better crop growth under difficult climatic conditions [14].

The timing and type of fertilizer application also had a big effect on yield results. Fertilization scheduling synchronized with periods of sufficient rainfall boosted nutrient uptake and general plant health, as well as increasing yields compared with fertilization schedules not synchronized with periods of necessary rainfall, the study found. These results highlight the need to begin integrating agricultural practices based on climatic conditions, in addition to the needs of the crop [29].

5.3.2 Adaptive Agricultural Practices

In addition, the study examined the effects of different agricultural practices on sugarcane yield. Mitigation of the impacts of high temperatures and prolonged dry spells was found to be particularly aided by efficient irrigation. Drip irrigation, for example, was found to provide increased water use efficiency and promote better crop growth under difficult climatic conditions [14].

The timing and type of fertilizer application also had a big effect on yield results. Fertilization scheduling synchronized with periods of sufficient rainfall boosted nutrient uptake and general plant health, as well as increasing yields compared with fertilization schedules not synchronized with periods of necessary rainfall, the study found. These results highlight the need to begin integrating agricultural practices based on climatic conditions, in addition to the needs of the crop [29].

5.3.3 Technological Adoption in Agriculture

The role of technology to modernize agriculture is a well explained theme in the literature with many studies documenting how digital tools and data driven approaches can increase productivity and sustainability [10], [25]. This view has been supported by the findings of this study, which have shown tangible benefits of using these technologies as IoT, remote sensing, and predictive analytics in sugarcane farming.

Results from this study add to the growing body of evidence positing technology is key to the future of agriculture, particularly in regions susceptible to climate change. These tools help the farmers inform better, optimizing their practice and in the process improve their livelihoods. [29].

5.4 Practical Implications

5.4.1 Recommendations for Farmers

Based on the findings of this study, several practical recommendations can be made for sugarcane farmers in Bangladesh:

- Adopt Climate-Resilient Practices: Practices to improve crop crop resilience to climate variability should be considered by farmers. The use of heat tolerant sugarcane varieties; efficient irrigation system such as drip irrigation and integrated pest management strategies are also included [14], [29].
- Optimize Harvest Timing: Farmers should select their harvest schedules responsibly and in line with the optimal window as discovered in this study. Farmers can maximize both yield and sugar content along with increasing economic returns by harvesting during late December to mid January [1].
- Leverage Technology: Modern technologies like IoT devices and remote sensing can help us get valuable data to help us make decisions on farming. It is important that farmers tap into these tools and explore opportunities to use them for monitoring climate condition and maximizing irrigation [25].

5.4.2 Policy Recommendations

The findings of this study also have important implications for policymakers and agricultural planners:

- Support for Climate-Resilient Agriculture: Development and dissemination of climate resilient agricultural practices are the top priorities for policymakers. All this involves investing on research and development of new crop variety, providing subsidies in irrigation systems and training farmers on sustainable farming practices. [11].
- Promotion of Technological Adoption: Governments should provide financial incentives, easy access to technologies, and the encouragement of knowledge exchange by farmers for the adoption and advance of high tech agriculture. This can comprise projects like offering subsidized loans for buying IoT devices, grants or support for farmers who use precision agriculture [27].
- Climate-Responsive Agricultural Planning: Climate projection and real time data must be considered more and more in agricultural planning. The importance of developing frameworks that integrate climate data into decision making processes and agro policies responsive to the current and future climate challenges is a policy makers concern [7].

5.5 Limitations of the Study

5.5.1 Data Constraints

The study also has some limitations to data availability and quality. For instance, historical documentation of agricultural practices did not always include treatises, and if so, they were not always complete or consistent, which could have confused analysis. Furthermore, the climate data, although comprehensive, may not span all the possible local variations which might influence sugarcane yield [15].

5.5.2 Model Assumptions

The predictive models developed in this study are dependent on a number of assumptions, including the continuation of current trends in climate change and agricultural practices. These assumptions are reasonable with the data now available, but these may not completely take into consideration some unanticipated changes or anomalies in future climate conditions. Finally, the accuracy of the input data, and thus the reliability of the predictions, relies upon the models being loyal [13].

5.5.3 Generalizability of Findings

The results of this study are thought to be specifically applicable to sugarcane farming in Bangladesh and regions alike. The methodologies and models, however, could be adapted for other crops, other regions but the specific results may not directly transferable without further validation and calibration to local conditions [10].

5.6 Future Research Directions

5.6.1 Advanced Climate Modeling

Future work should take advantage of advanced climate models, such as statistical downscaling techniques via regional climate models (RCMs), to better quantify how agriculture will be impacted by climate. These models could improve yield predictions as well as yield tailored recommendations for farmers [29].

5.6.2 Longitudinal Studies on Agricultural Practices

To better understand the long term sustainability of these practices, longitudinal study would follow how specific agricultural practices over an extended period of time impacts the future impact of agriculture. Such studies may be able to help identify best practices that time and again result in improved yields and resilience to climate variability [19].

5.6.3 Integration of Economic Analysis

Finally, future research should include a more detailed economic analysis of the recommendations presented in this study. It could be cost benefit analyses of adopting new technology or practice or assessment of the extent to which climate change could affect the economy of sugarcane farming in Bangladesh [27].

5.6.4 Exploration of Policy Interventions

More research should be done to understand how different types of policies might help farmers grow climate resilient agriculture. It might include looking at the effect of subsidies, training programs and infrastructure expenditures on farmers' capacities to adapt drought to climate change. Such research would inform policymakers and help refine tactics that support the agricultural sector in Bangladesh as well [11].

5.6.5 Cross-Regional Comparisons

Additional insights into the effectivetness of various adaptive strategies could be gained by the comparative study of the regions with the same climatic conditions of other countries. Learning about how different regions have fared under the same challenges can help researchers to understand what works and what is learnt, and help eliminate what works in sugarcane farming in Bangladesh [29].

5.6.6 Real-Time Data Integration

Subsequent research should search for ways to mesh true time data into predictive fashions so that they'll be more correct and reactive. For instance, they might develop more dynamic models that dynamically update themselves as new data become available, drawing up to the minute advice to farmers for optimal farming practices and the time to harvest [25].

5.6.7 Climate-Resilient Crop Breeding

More work needs to be done to breed climate resilient sugarcane varieties. It encompasses not only creating strains with the ability to withstand higher temperatures and suffer the effects of a drier environment but also those that are more resistant to pest and disease problems that may arise given the continually changing climate [29].

5.7 Broader Implications for Agriculture in Bangladesh

5.7.1 Strengthening Food Security

This study has important implications for food security in Bangladesh. A visual. And this is being driven by the fact that climate change is clearly impacting on agricultural productivity, and as we know, there are a number of crops that sit at the base of the food supply. And when we talk about key crops, the one that we highlight as being most crucial is sugarcane, because it's the base crop for the livelihoods of 40 million farmers [29]. Adoption of recommended climate adaptive practices and technologies in this study could they be important in supporting the food security and economic stability of the country [11].

5.7.2 Enhancing Sustainability in Agriculture

The latter is in line with global tendencies to make intensive agriculture sustainable. While they cannot resemble the lush vegetables sold in the U.S., by adopting practices which conserve water usage, increase soil health, and minimize the use of chemical inputs, farmers are making an important contribution to the long term sustainability of agriculture in Bangladesh. Farming practices using technology also promote sustainability by helping this occur via integrating technology to optimize resource use and decrease waste [19].

5.7.3 Policy Development for Climate Change Adaptation

The results of this study indicate that there is a need for proactive policy development for climate change adaptation in agriculture. Policymakers should consider building on the creation of frameworks that facilitate the adoption of adaptive practices and technology, and provide financial support to farmers, and invest in research and development. The agricultural sector in Bangladesh will need these policies to help it cope with the impacts that climate change will have on the sector and for the continued productivity and viability of farming in Bangladesh [11].

5.7.4 Contributions to Global Climate Change Mitigation

The focus of this study is on adaptation, although a realistic focus on more sustainable farming practices can add to global climate change mitigation. The agricultural sector in Bangladesh can help reduce greenhouse gas emissions by using resources more efficiently, minimizing the amount of the worldwide future carbon stored by the vegetation of all types, and halting damage to the world's remaining dense forests [27].

5.8 Summary of Chapter 5

This chapter has discussed in detail the key findings from the study and their implications for sugarcane farming in Bangladesh and elsewhere. Results demonstrate the importance of climate adaptive practices, optimal harvest timing and the use of modern technologies in increasing agricultural productivity and resilience **Shrivastava 2015**, **Shukla 2013**. Other chapters address the study's limitations and suggest future research directions, maintaining that sugarcane production will need to continue to improve and to sustain under the pressure of climate change [11].

Chapter 6

Conclusion and Recommendations

6.1 Integration of Findings and Their Implications

This study draws to a close by emphasizing the intricacies of climate variables, agricultural practices and yield densities of sugarcane in Bangladesh [34]. Yet the integration of predictive analytics within routine farm processes has shown that precise, data driven decisions can yield significant gains in yield while facing weather hurdles **Harris2014**, [25], [29]. In this chapter, we distil the key findings into actionable recommendations for farmers, fertilizer manufacturers, researchers, governments, and others, and consider the potential broader implications for sustainable agricultural development [11].

6.2 Strategic Recommendations

6.2.1 Enhancing Decision-Making with Predictive Models

In this study, the predictive models developed have demonstrated excellent potential to predict sugarcane yields from climate variables [7], [8]. Finally, these models should integrate into agricultural decision making processes from farm and policy levels. For example, using these models, farmers could change planting and harvesting schedules based on what they believe the climate will be like, or policymakers would use them for regional agricultural planning to ensure that resources were allocated effectively [29].

6.2.2 Adoption of Precision Agriculture

Seeing the advantage to merging the technology into farming practices, farmers should adopt precision agriculture technologies. GPS guided machinery, drones for crop monitoring, and automated irrigation systems are included [10], [25]. Part of that precision agriculture helps optimize input use, but also enables better monitoring of crop health which gives better prediction of yield, and therefore resource management [19].

6.2.3 Development of Climate-Resilient Agricultural Policies

It is time for policymakers to set policies that promote the adoption of climate resistant agricultural practices [27]. It could be through the subsidizing of the cost of implementing precision agriculture technologies; training farmers on adaptive farming techniques; and the investing of research aimed at the development of region specific solutions to climate issues [11].

6.2.4 Strengthening Agricultural Extension Services

In order to enhance the reach of the insights of this study to the farmers, agricultural extension service should also be strengthened. Development of new technologies and tools in climate adapted practices and climate adaptation practices should be taught to extension workers for timely and accurate advice to farmers. This will enable farmers to choose the right decisions that improve productivity while also being more sustainable [29].

6.3 Broader Implications for Sustainable Agriculture

6.3.1 Contributions to Global Food Security

This study makes the case for the approaches and strategies discussed with broader application beyond sugarcane farming in Bangladesh [29]. Other regions are able to increase their agricultural resilience to climate change by adopting similar predictive and adaptive techniques [19] contributing to global food security.

6.3.2 Promoting Environmental Sustainability

The resources focus, namely water and fertilizers, confirms with the global challenges to promote environmental sustainability [14]. Not only does precision agriculture increase productivity it also attempts to reduce environmental impact, and there is a compelling link between precision agriculture and sustainability in farming [27].

6.4 Conclusion

The main contribution of this study is to demonstrate how predictive analytics can be integrated with traditional agricultural practices to best maximize sugarcane yield in Bangladesh [25], [29]. The findings emphasize the centrality of climate adaptive strategies, precision agriculture and supportive policies in maintaining agricultural productivity in an increasingly variable climate [11]. Different organizations, including farmers, policymakers and researchers will use the recommendations to be able to build a resilient and sustainable agriculture sector together.

Bibliography

[1] J. L. Monteith and C. J. Moss, "Climate and the efficiency of crop production in britain," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 281, no. 980, pp. 277–294, 1977. DOI: 10.1098/rstb.1977.0140.

[2] M. J. Robertson and R. C. Muchow, "Effects of temperature on the growth of sugarcane: The development of a predictive model for potential production," *Field Crops Research*, vol. 36, no. 3, pp. 213–231, 1994. DOI: 10.1016/0378-4290(94)90031-0.

[3] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. DOI: 10.1007/BF00994018.

[4] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd. John Wiley & Sons, 1998.

[5] R. C. Muchow and B. A. Keating, "Assessing the water stress response of sugarcane: A comparison of different methods," *Agricultural Water Management*, vol. 35, no. 1-2, pp. 1–29, 1998. DOI: 10.1016/S0378-3774(97)00060-5.

[6] B. Singh and M. E. Maayar, "Potential impacts of greenhouse gas climate change scenarios on sugar cane yields in trinidad," *Agricultural and Forest Meteorology*, vol. 89, no. 1, pp. 137–148, 1998. DOI: 10.1016/S0168-1923(97)00071-4.

[7] J. W. Hansen and J. W. Jones, "Scaling-up crop models for climate variability applications," *Agricultural Systems*, vol. 65, no. 1, pp. 43–72, 2000. DOI: 10.1016/S0308-521X(00)00025-1.

[8] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.

[9] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002. [Online]. Available: https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf.

[10] J. W. Jones, G. Hoogenboom, C. H. Porter, *et al.*, "The dssat cropping system model," *European Journal of Agronomy*, vol. 18, no. 3-4, pp. 235–265, 2003. DOI: 10.1016/S1161-0301(02)00107-7.

[11] M. W. Rosegrant and S. A. Cline, "Global food security: Challenges and policies," *Science*, vol. 302, no. 5652, pp. 1917–1919, 2003. DOI: 10.1126/science.1092958.

[12] N. G. Inman-Bamber and D. M. Smith, "Water relations in sugarcane and response to water deficits," *Field Crops Research*, vol. 92, no. 2-3, pp. 185–202, 2005. DOI: 10.1016/j.fcr.2005.01.023.

[13] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.

[14] E. Fereres and M. A. Soriano, "Deficit irrigation for reducing agricultural water use," *Journal of Experimental Botany*, vol. 58, no. 2, pp. 147–159, 2007. DOI: 10.1093/jxb/erl165.

[15] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer Science Business Media, 2009.

[16] R. Jain, S. Solomon, A. K. Shrivastava, *et al.*, "Impact of harvest to planting delays on the sprouting of seed cane," *Sugar Tech*, vol. 11, pp. 231–233, 2009. DOI: 10.1007/s12355-009-0040-6. [Online]. Available: https://doi.org/10.1007/s12355-009-0040-6.

[17] S. Siddiqui, "Sugarcane: The backbone of the agrarian economy in bangladesh," *Bangladesh Journal of Agricultural Economics*, vol. 32, no. 1, pp. 45–60, 2010.

[18] D. B. Lobell, W. Schlenker, and J. Costa-Roberts, "Climate trends and global crop production since 1980," *Science*, vol. 333, no. 6042, pp. 616–620, 2011. DOI: 10.1126/science.1204531.

[19] D. Tilman, C. Balzer, J. Hill, and B. L. Befort, "Global food demand and the sustainable intensification of agriculture," *Proceedings of the National Academy of Sciences*, vol. 108, no. 50, pp. 20 260–20 264, 2011. DOI: 10.1073/pnas.1116437108.

[20] G. Biau, "Analysis of a random forests model," *Journal of Machine Learning Research*, vol. 13, pp. 1063–1095, 2012.

[21] D. L. Santos and P. C. Sentelhas, "Climate change scenarios and their impact on the water balance of sugarcane production areas in the state of são paulo, brazil," *Revista Ambiente Água*, vol. 7, no. 2, pp. 07–17, 2012. DOI: 10.4136/ambi-agua.907.

[22] R. S. Verma and A. Singh, "Industrial applications of sugarcane and its by-products," *International Sugar Journal*, vol. 114, no. 1357, pp. 456–463, 2012. DOI: 10.1007/s12355-012-0127-1.

[23] M. M. Das and M. S. Rahman, "The industrial potential of sugarcane by-products in bangladesh," *Journal of Industrial Biotechnology*, vol. 17, no. 3, pp. 201–212, 2013.

[24] F. R. Marin, J. W. Jones, F. Royce, and E. D. Assad, "Climate change impacts on sugarcane attainable yield in southern brazil," *Climatic Change*, vol. 117, no. 1, pp. 227–239, 2013. DOI: 10.1007/s10584-012-0561-y.

[25] A. K. Shukla and S. Jharkharia, "Application of machine learning techniques in agricultural production: A review," *Journal of Agricultural Informatics*, vol. 4, no. 1, pp. 18–24, 2013. DOI: 10.17700/jai.2013.4.1.78.

[26] M. Y. Ali and K. Jahan, "The role of sugarcane in bangladesh's economy: Challenges and opportunities," *Journal of Agricultural Economics*, vol. 55, no. 2, pp. 123–135, 2014. DOI: 10.1111/agec.12091.

[27] H. C. J. Godfray and T. Garnett, "Food security and sustainable intensification," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1639, p. 20 120 273, 2014. DOI: 10.1098/rstb.2012.0273.

[28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539.

[29] A. K. Shrivastava, V. K. Singh, T. K. Srivastava, V. Kumar, S. P. Shukla, and V. Misra, "Impact of climate change on sugarcane and its mitigation," *Agricultural Engineering Today*, vol. 39, no. 4, pp. 31–40, 2015.

[30] A. K. Singh and K. G. S. Narayanan, "Effect of climatic factors on cash crop farming in india: An application of cobb-douglas production function model," *International Journal of Agricultural Resources, Governance, and Ecology*, vol. 13, no. 2, pp. 175–210, 2017. DOI: 10.1504/IJARGE.2017.083169.

[31] FAO, "Fao statistical yearbook: World food and agriculture," *Food and Agriculture Organization of the United Nations*, 2019. [Online]. Available: https://www.fao.org/statistics/yearbook/.

[32] S. Pipitpukdee, W. Attavanich, and S. Bejranonda, "Climate change impacts on sugarcane production in thailand," *Atmosphere*, vol. 11, no. 4, p. 408, 2020. DOI: 10.3390/atmos11040408.

[33] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd. OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/.

[34] IPCC Report, *Climate change 2021: Impacts, adaptation, and vulnerability*, Accessed: 2024-4-18, 2021. [Online]. Available: https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_FullReport.pdf.